SIAM J. CONTROL OPTIM.
Vol. 58, No. 6, pp. 3389–3412

A GAUSS–SEIDEL TYPE METHOD FOR DYNAMIC NONLINEAR COMPLEMENTARITY PROBLEMS*

SHU-LIN WU†, TAO ZHOU‡, AND XIAOJUN CHEN§

Abstract. The dynamic nonlinear complementarity problem (DNCP) consisting of a nonlinear differential system and a complementarity system has been used to formulate and study many dynamic problems. In a Gauss–Seidel type method for DNCPs, by first guessing a solution of the differential system, we can solve the complementarity system and then with the computed solution we can solve the differential system to update the guess. Upon convergence at the current time point we can move to the next one. The idea can be easily generalized to a multipoint version: instead of doing iterations at each single time point, we can do iterations for a number of time points, say J time points, all at once. Despite its simplicity and easy implementation, convergence of this method is not justified so far. In this paper, we present interesting convergence theorems for this method. We show that the method with a fixed length of time interval converges *superlinearly* and the convergence rate is robust with respect to the step-size h. Moreover, we show that the method with a fixed number of time points converges with a rate $\mathcal{O}(h)$. Since at each iteration the differential system and the complementarity system are solved separately, many existing solvers are directly applicable for each of these two systems. It is notable that we can solve the complementarity system at all the J time points in parallel. Numerical results of the method to solve the 4-diode bridge wave rectifier with random circuit parameters and the projected dynamic systems are given to support our findings.

Key words. dynamic nonlinear complementarity problems, iterative methods, convergence analysis, nonsmooth circuit systems, projected dynamic systems

AMS subject classifications. 65M55, 65M12, 65M15, 65Y05

DOI. 10.1137/19M1268884

1. Introduction.

We are interested in solving the following dynamic nonlinear complementarity problem (DNCP) with initial value $x(0) = x_0$:

$$\dot{x}(t) = F(t, x(t), y(t)), \quad 0 \le y(t) \perp G(t, x(t), y(t)) \ge 0,\ t \in (0, T), \tag{1.1}$$

where $x(t) \in \mathbb{R}^m$, $y(t) \in \mathbb{R}^n_+$, $F : \mathbb{R}_+ \times \mathbb{R}^m \times \mathbb{R}^n_+ \to \mathbb{R}^m$, and $G : \mathbb{R}_+ \times \mathbb{R}^m \times \mathbb{R}^n_+ \to \mathbb{R}^n$. The nonnegativity and perpendicularity in (1.1) are explained in the component sense. Applications of differential complementarity problems and other closely related models, such as the differential variational inequalities [12, 31], can be found in many places; we refer the reader to the excellent monographs [13, 16] and the survey papers [26, 29]. An important subclass is the following differential semiaffine systems [2, 16, 26]:

$$\dot{x}(t) = F(t, x(t), y(t)), \quad 0 \le y(t) \perp Nx(t) + My(t) + g(t) \ge 0,\ t \in (0, T), \tag{1.2}$$

*Received by the editors June 17, 2019; accepted for publication (in revised form) August 6, 2020; published electronically November 17, 2020.
https://doi.org/10.1137/19M1268884
Funding: The work of the first author is supported by NSF of China (under grant 11771313) and NSF of Sichuan Province (under grant 2018JY0469). The work of the second author is partially supported by NSF of China (under grants 11822111 and 11688101), Science Challenge Project (TZ2018001), and Youth Innovation Promotion Association (CAS). The work of the third author is partially supported by Hong Kong Research Grant Council for grant PolyU153001/18P.

†Corresponding author. School of Mathematics and Statistics, Northeast Normal University, Changchun 130024, China (wushulin84@hotmail.com).
‡LSEC, Institute of Computational Mathematics and Scientific/Engineering Computing, AMSS, Chinese Academy of Sciences, Beijing, 100190, China (tzhou@lsec.cc.ac.cn).
§Department of Applied Mathematics, The Hong Kong Polytechnic University, Kowloon, Hong Kong (xiaojun.chen@polyu.edu.hk).

where $N \in \mathbb{R}^{n\times m}$, $M \in \mathbb{R}^{n\times n}$, and $g : \mathbb{R}_+ \to \mathbb{R}^n$. A special case of (1.2) is the dynamic *linear* complementarity problem (DLCP) [4, 5, 6, 9, 16, 18, 20, 26, 28, 30, 32, 33]:

$$\dot{x}(t) = Ax(t)+By(t)+f(t), \quad 0 \le y(t) \perp Nx(t)+My(t)+g(t) \ge 0,\ t \in (0,T), \tag{1.3}$$

where $A \in \mathbb{R}^{m\times m}$, $B \in \mathbb{R}^{m\times n}$, and $f : \mathbb{R}_+ \to \mathbb{R}^m$.

For a DNCP, the exact solution is not available in general, and quantitative study of these problems mainly relies on numerical computation. The time-stepping method is widely used [1, 4, 9, 10, 11, 13, 16, 18, 20, 26, 33]. Applying the Backward-Euler method to (1.1) gives

$$0 \le y_j \perp G(t_j, x_j, y_j) \ge 0, \quad x_j = x_{j-1} + hF(t_j, x_j, y_j),\ \ j = 1, 2, \ldots, N_t, \tag{1.4}$$

where $h = \frac{T}{N_t}$ is the step-size and N_t is a positive integer.[1] Clearly, the major computation cost of solving (1.1) lies in solving the nonlinear system (1.4) at each time point, and the goal of this paper is to establish an efficient method to solve such a nonlinear system.

1.1. The existing methods. There are two mainstream methods for solving (1.4) at each time point. The first one is a direct method, which is for DLCP (1.3). The nonlinear system (1.4) for DLCP is

$$0 \le y_j \perp Nx_j + My_j + g_j \ge 0, \quad x_j = x_{j-1} + hAx_j + hBy_j + hf_j, \quad j = 1, 2, \ldots, N_t. \tag{1.5}$$

From the discretized ODE system x_j can be given as

$$x_j = h(I - hA)^{-1}By_j + \tilde{f}_j \text{ with } \tilde{f}_j := (I - hA)^{-1}(x_{j-1} + hf_j), \tag{1.6a}$$

and then we substitute x_j into the linear complementarity system (LCS) in (1.5):

$$0 \le y_j \perp \tilde{g}_j + M_h y_j \ge 0 \text{ with } M_h := hN(I - hA)^{-1}B + M, \quad \tilde{g}_j := g_j + N\tilde{f}_j. \tag{1.6b}$$

Solving (1.6b) gives y_j and then by substituting y_j into (1.6a) we get x_j. Method (1.6a)–(1.6b) is extensively studied in the literature; cf. [1, 4, 5, 10, 11, 18, 20, 25, 26]. Problems exist for this approach in two aspects. First, even though the LCS in (1.5) has a unique solution in a certain sense, there is no guarantee that this is also true for (1.6b). For example, if M is a P-matrix[2] the LCS in (1.5) has a unique solution at each time point t_j, but this is often not the case for (1.6b) because the matrix M_h may not be a P-matrix, unless the step-size h is sufficiently small. Second, for large-scale problems, e.g., the DLCPs arising from the parabolic Signorini problems [14, 32], it would be difficult to compute the matrix M_h. For example, suppose $B = [b_1, \ldots, b_n]$; then we have to solve n linear systems $\{(I - hA)^{-1}b_l\}_{l=1}^n$. The massive computer memory and computation time required would be serious problems if A and B are large size matrices.

The second mainstream method for solving (1.4) is the semismooth Newton method [7, 9]. Under some suitable assumptions, the complementarity system has

[1] We assume that the time points $\{t_j\}_{j=0}^J$ are equally spaced, i.e., $\{t_j = jh\}_{j=0}^{N_t}$, but this is not a restrictive assumption since all the results obtained in this paper also hold for arbitrarily chosen time points.

[2] A matrix M is called a P-matrix if all the principal minors of M are positive.

a unique solution denoted by $\mathcal{Y}(x_j)$. Then, by substituting this expression into the differential system we get a nonlinear system concerning x_j: $\mathcal{F}(x_j) = 0$, where

$$\mathcal{F}(z) := z - x_{j-1} - hF(t_j, z, \mathcal{Y}(z)).$$

Applying the semismooth Newton method to this problem results in the following iterations:

$$V_j^k \Delta x_j^k = -\mathcal{F}(x_j^k), \quad x_j^{k+1} = \Delta x_j^k + x_j^k, \quad k = 0, 1, \ldots, \tag{1.7}$$

where x_j^0 is the initial guess and V_j^k is the Clarke generalized Jacobian matrix of $\mathcal{F}(z)$ at x_j^k. The semismooth Newton method is *locally* convergent, and therefore, the initial guess x_j^0 must be very close to x_j [9]. Moreover, it would be difficult to get the Clarke Jacobian matrix V_j^k for DNCPs of large size.

1.2. New idea. The goal of this paper is to avoid the aforementioned problems for the direct method (1.6a)–(1.6b) and the semismooth Newton method (1.7), by solving (1.4) iteratively in a Gauss–Seidel fashion:

$$0 \le y_j^{k+1} \perp G(t_j, x_j^k, y_j^{k+1}) \ge 0, \quad x_j^{k+1} = x_{j-1} + hF(t_j, x_j^{k+1}, y_j^{k+1}), \tag{1.8}$$

where $k \ge 0$ is the iteration index and x_j^0 is an initial guess of x_j. Upon convergence, we have $x_j^\infty = x_j$ and $y_j^\infty = y_j$. Since the differential and complementarity systems are solved separately, for each of these two systems many existing methods can be used without changes. Precisely, the nonlinear function F that models the differential system is often a smooth function, and therefore, x_j^{k+1} can be obtained by using the *classical* Newton method [15]. For the complementarity system, we can solve it by many mature solvers, e.g., the iterative method based on some *linearization* technique [27] and the PATH solver [17].[3]

Of particular interest is the semiaffine problems (1.2), i.e.,

$$0 \le y_j^{k+1} \perp Nx_j^k + My_j^{k+1} + g_j \ge 0, \quad x_j^{k+1} = x_{j-1} + hF(t_j, x_j^{k+1}, y_j^{k+1}) \tag{1.9}$$

for which we can solve the LCS via existing optimization solvers. For example, if M is a Z-matrix,[4] we can obtain y_j^{k+1} via solving the following linear programming problem [9]:

$$\min \|y\|_1, \text{ s.t. } y \ge 0, \ My + Nx_j^k + g_j \ge 0.$$

Moreover, for DLCPs, i.e.,

$$0 \le y_j^{k+1} \perp Nx_j^k + My_j^{k+1} + g_j \ge 0, \quad x_j^{k+1} = x_{j-1} + hAx_j^{k+1} + hBy_j^{k+1} + hf_j, \tag{1.10}$$

we need to solve only one linear system to get x_j^{k+1}, i.e., $(I - hA)x_j^{k+1} = x_{j-1} + hBy_j^{k+1} + hf_j$. Suppose after k^* iterations of (1.10) the error arrives at the prescribed tolerance; then we solve k^* linear systems in total. In practice, e.g., for the examples studied in section 4, k^* is much less than the number of linear systems needed to form the matrix M_h for the direct method (1.6a)–(1.6b) and the Clarke generalized Jacobian matrix V_j^k for the semismooth Newton method (1.7). For large-scale problems, this is an important advantage.

[3]The most recent PATH solver can be downloaded for free from http://pages.cs.wisc.edu/∼ferris/path.html.

[4]A matrix M is a Z-matrix if its off-diagonal elements are nonpositive.

Instead of applying (1.8) to each single time point, we can generalize the idea to a *multipoint* version as follows. First, we divide $\{t_1, t_2, \ldots, t_{N_t}\}$ into P groups:

$$\{t_1, t_2, \ldots, t_J\}, \{t_{J+1}, t_{J+2}, \ldots, t_{2J}\}, \ldots, \{t_{(P-1)J+1}, t_{(P-1)J+2}, \ldots, t_{PJ}\},$$

where $t_{PJ} = t_{N_t} = T$ and $J = \frac{N_t}{P} \geq 1$ is an integer. Then, we apply (1.8) to each of these P groups of time points one-by-one. Without loss of generality, we assume $P = 1$ (i.e., $J = N_t$) and then we get a multipoint version of (1.8) as

$$0 \leq y_j^{k+1} \perp G\left(t_j, x_j^k, y_j^{k+1}\right) \geq 0, \quad x_j^{k+1} = x_{j-1}^{k+1} + hF\left(t_j, x_j^{k+1}, y_j^{k+1}\right), \tag{1.11}$$

where $j = 1, 2, \ldots, J$. The quantities $\{y_j^{k+1}\}_{j=1}^J$ are independent of each other, and therefore, the computation of the complementarity system at all the J time points is in parallel. If the computation of the complementarity system is much more expensive than that of the ODE system, e.g., for DNCP (1.1) with $n \gg m$, such a parallelism can save considerable computation time.

The iterative algorithm (1.8) is the basis of this paper, but for completeness we will make a convergence analysis for the multipoint version (1.11) and the convergence properties of (1.8) can be directly deduced (cf. Remark 2.2). We will show the convergence of the iterative method (1.11) when G is a uniform P-function with respect to y (cf. (2.1) for definition). We prove that the method has two different convergence properties depending on whether we use it for a fixed length of time interval or we use it for a fixed number of time points. For the first situation, we prove *superlinear* convergence with a rate independent of the step-size h. For the second situation, we prove that the method converges with a rate $\mathcal{O}(h)$ and thus a smaller step-size h accelerates the convergence speed. For the case when G is a linear function of y,

$$G(t, x(t), y(t)) = My(t) + \tilde{G}(t, x(t)), \tag{1.12}$$

we prove that these results hold if M is a Z-matrix or positive semidefinite matrix.

The rest of this paper is organized as follows. In section 2, we present the convergence analysis of the new iterative method for the case when G is a uniform P-function. In section 3, we consider the case (1.12) for M being a Z-matrix or positive semidefinite matrix. In section 4, we show applications together with numerical results of the new iterative method for differential complementarity systems arising in three different fields. This includes a direct application of the method to a 4-diode bridge wave rectifier consisting of a nonlinear resistor and a capacity with random value, and a modified application of the method to a projected dynamic system arising from the spatial price equilibrium problem. The numerical results show that the new iterative method is superior to the existing methods, with respect to robustness, complexity, and computation time. We conclude this paper in section 5.

Remark 1.1. An iterative method of Gauss–Seidel style for solving DLCPs has been studied in [32]. The main idea is to express the exact solution $x^{k+1}(t)$ via Laplace inversion transform:

$$x^{k+1}(t) = \frac{1}{2\pi i} \int_\Gamma e^{-st}(sI - A)^{-1}(x_0 + B\widehat{y}^{k+1}(s) + \widehat{f}(s))ds, \tag{1.13}$$

where $\widehat{y}^{k+1}(s)$ and $\widehat{f}(s)$ denote the Laplace forward transforms of $y^{k+1}(t)$ and $f(t)$ and Γ denotes a contour in the complex plane, which is a simple, closed, positively oriented curve enclosing the spectrum of A. This method is only applicable to DLCP (1.3). For the nonlinear case or the linear case with time-dependent coefficient matrix $A(t)$, we cannot represent $x^{k+1}(t)$ via Laplace inversion, and thus the method proposed in [32] is entirely not applicable to DNCPs.

2. Convergence analysis in the uniform P-function case. In this section, we consider the case that the function $G(t,x,y)$ is a uniform P-function of y in $\mathbb{R}^n_+$, i.e., there exists a constant $L_0>0$ such that (cf. [26, section 5])

$$\max_{1\le l\le n}(\bar{y}_l-\tilde{y}_l)(G_l(t,x,\bar{y})-G_l(t,x,\tilde{y}))\ge L_0\|\bar{y}-\tilde{y}\|_2^2, \tag{2.1}$$

which holds for all $(t,x)\in\mathbb{R}_+\times\mathbb{R}^m$ and $\bar{y},\tilde{y}\in\mathbb{R}^n_+$, where G_l denotes the lth component of the function G. Moreover, we assume that there exists a constant $L_G>0$ such that

$$\|G(t,x,\bar{y})-G(t,x,\tilde{y})\|_2\le L_G\|\bar{y}-\tilde{y}\|_2\ \forall(t,x)\in\mathbb{R}_+\times\mathbb{R}^m,\ \bar{y},\tilde{y}\in\mathbb{R}^n_+. \tag{2.2}$$

The property of uniform P-function, together with the Lipschitz condition, for the function G implies the following lemma.

LEMMA 2.1 (Theorem 5.1 in [26]). *Suppose the function $G(t,x,y)$ satisfies* (2.1)–(2.2). *Then, the nonlinear complementarity problem $0\le y\perp G(t,x,y)\ge 0$ has a unique solution $\mathcal{Y}(x)$ for any $(t,x)\in\mathbb{R}_+\times\mathbb{R}^m$, which is a Lipschitz continuous function of x at a fixed t. Particularly, there exists a constant $\eta_t>0$ such that $\|\mathcal{Y}(\bar{x})-\mathcal{Y}(\tilde{x})\|_2\le\eta_t\|\bar{x}-\tilde{x}\|_2$.*

Remark 2.1. The property of uniform P-function resembles the notion of "pointwise strong regularity" discussed by Pang and Shen in [24]. The assumption of pointwise strong regularity is weaker than the assumption of uniform P-function property, and a local Lipschitz solution can result in a specific NCP solution, which is similar to the statements of Lemma 2.1.

We next assume the following Lipschitz conditions for $F(t,x,y)$:

$$\begin{aligned}&\langle F(t,\bar{x},y)-F(t,\tilde{x},y),\bar{x}-\tilde{x}\rangle\le L_1\|\bar{x}-\tilde{x}\|_2^2\ \forall(t,y)\in\mathbb{R}_+\times\mathbb{R}^n_+,\ \bar{x},\tilde{x}\in\mathbb{R}^m,\\ &\|F(t,x,\bar{y})-F(t,x,\tilde{y})\|_2\le L_2\|\bar{y}-\tilde{y}\|_2\ \ \forall(t,x)\in\mathbb{R}_+\times\mathbb{R}^m,\ \bar{y},\tilde{y}\in\mathbb{R}^n_+,\end{aligned} \tag{2.3}$$

where $\langle\cdot\rangle$ is the standard Euclidean inner product, $L_1\in(-\infty,\infty)$ and $L_2>0$. The first condition in (2.3) is called *one-sided* Lipschitz condition, and L_1 can be positive or negative. The following lemma (see Appendix A for the proof) about the combinatorial identities is useful for the convergence analysis for method (1.11).

LEMMA 2.2. *Let $r>0$ and $\psi(r,J,k)=\sum_{j_1=1}^{J}\sum_{j_2=1}^{j_1}\cdots\sum_{j_k=1}^{j_{k-1}}r^{J-j_k}$. Then, we have*

$$\psi(r,J,k)=\begin{cases}\frac{1}{(1-r)^k}-\sum_{l=1}^{k}\binom{J+k-l-1}{k-l}\frac{r^J}{(1-r)^l}, & r\ne 1,\\ \binom{J+k-1}{k}, & r=1.\end{cases} \tag{2.4}$$

THEOREM 2.3. *For problem* (1.1), *suppose the functions G and F satisfy* (2.1), (2.2), *and* (2.3). *Then, for method* (1.11) *the error $e_j^k=x_j-x_j^k$ satisfies*

$$\max_{0\le j\le J}\|e_j^k\|_2\le\begin{cases}(h\widetilde{\eta})^k\psi(1,J,k)\max_{0\le j\le J}\|e_j^0\|_2 & \text{if } L_1=0,\\ \psi((1-hL_1)^{-1},J,k)\left(\frac{h\widetilde{\eta}}{1-hL_1}\right)^k\max_{1\le j\le J}\|e_j^0\|_2 & \text{if } L_1\ne 0,\end{cases} \tag{2.5}$$

provided $hL_1<1$, where $\widetilde{\eta}=L_2\eta$ and x_j is the converged solution (i.e., the solution of the fully nonlinear one-step problem (1.4)*).*

Proof. By using Lemma 2.1, we represent y_j^{k+1} as $\mathcal{Y}(x_j^k)$ and the converged solution y_j in (1.4) as $\mathcal{Y}(x_j)$. Therefore, we can rewrite (1.4) and (1.11) as

$$x_j = x_{j-1} + hF(t_j, x_j, \mathcal{Y}(x_j)), \quad x_j^{k+1} = x_{j-1}^{k+1} + hF(t_j, x_j^{k+1}, \mathcal{Y}(x_j^k)).$$

For the simplicity of notation, we let $e_j^k = x_j - x_j^k$, $q = \mathcal{Y}(x_j)$, and $q^k = \mathcal{Y}(x_j^k)$. Then,

$$e_j^{k+1} = e_{j-1}^{k+1} + h\left[F(t_j, x_j, q) - F(t_j, x_j^{k+1}, q^k)\right].$$

To get an estimate of $\|e_j^{k+1}\|_2$, we consider the following inner product:

$$\begin{aligned}\langle e_j^{k+1}, e_j^{k+1}\rangle = \langle e_j^{k+1}, e_{j-1}^{k+1}\rangle &+ h\langle e_j^{k+1}, F(t_j, x_j, q) - F(t_j, x_j^{k+1}, q)\rangle \\ &+ h\langle e_j^{k+1}, F(t_j, x_j^{k+1}, q) - F(t_j, x_j^{k+1}, q^k)\rangle.\end{aligned}$$

By using the Cauchy–Schwarz inequality and the two Lipschitz conditions in (2.3), we get

$$\|e_j^{k+1}\|_2^2 \le \|e_j^{k+1}\|_2\|e_{j-1}^{k+1}\|_2 + hL_1\|e_j^{k+1}\|_2^2 + hL_2\|e_j^{k+1}\|_2\|q - q^k\|_2. \tag{2.6}$$

For the last term in (2.6), by using Lemma 2.1 it holds that

$$\|q - q^k\|_2 = \|\mathcal{Y}(x_j) - \mathcal{Y}(x_j^k)\|_2 \le \eta\|e_j^k\|_2.$$

Substituting this into (2.6) gives $\|e_j^{k+1}\|_2 \le \|e_{j-1}^{k+1}\|_2 + hL_1\|e_j^{k+1}\|_2 + h\widetilde{\eta}\|e_j^k\|_2$, with $\widetilde{\eta} = L_2\eta$. The assumption $hL_1 < 1$ implies $\|e_j^{k+1}\|_2 \le \frac{1}{1-hL_1}\|e_{j-1}^{k+1}\|_2 + \frac{h\widetilde{\eta}}{1-hL_1}\|e_j^k\|_2$.

Let $\{\epsilon_j^k\}_{j=0}^J$ be the sequence defined by

$$\epsilon_j^{k+1} = \frac{1}{1-hL_1}\epsilon_{j-1}^{k+1} + \frac{h\widetilde{\eta}}{1-hL_1}\epsilon_j^k, \text{ with } \{\epsilon_j^0 = \|e_j^0\|_2\}_{j=0}^J \text{ and } \{\epsilon_0^k = 0\}_{k\ge 0}.$$

It is clear $\|e_j^k\|_2 \le \epsilon_j^k$. So, it suffices to establish a suitable inequality between $\max_{1\le i\le J}\epsilon_j^k$ and $\max_{1\le i\le J}\epsilon_j^0$. We have

$$\epsilon_j^{k+1} = \frac{h\widetilde{\eta}}{1-hL_1}\sum\nolimits_{l=1}^{j}(1-hL_1)^{-(j-l)}\epsilon_l^k = h\widetilde{\eta}\sum\nolimits_{l=1}^{j}(1-hL_1)^{l-j-1}\epsilon_l^k, \tag{2.7}$$

and a successive application of this relation yields

$$\epsilon_j^k = (h\widetilde{\eta})^k\sum\nolimits_{j_1=1}^{j}\sum\nolimits_{j_2=1}^{j_1}\cdots\sum\nolimits_{j_k=1}^{j_{k-1}}(1-hL_1)^{j_k-j-k}\epsilon_{j_k}^0,$$

which gives $\epsilon_j^k \le \left[(h\widetilde{\eta})^k\sum_{j_1=1}^{j}\sum_{j_2=1}^{j_1}\cdots\sum_{j_k=1}^{j_{k-1}}(1-hL_1)^{j_k-j-k}\right]\max_{0\le l\le j}\epsilon_l^0$. This relation holds for all $j \in \{0, 1, \ldots, J\}$, and we therefore get

$$\max_{0\le j\le J}\epsilon_j^k \le \left(\frac{h\widetilde{\eta}}{1-hL_1}\right)^k\left[\sum\nolimits_{j_1=1}^{J}\sum\nolimits_{j_2=1}^{j_1}\cdots\sum\nolimits_{j_k=1}^{j_{k-1}}(1-hL_1)^{j_k-J}\right]\max_{0\le j\le J}\epsilon_j^0. \tag{2.8}$$

It remains to estimate the nested summation in (2.8). We consider the following two cases. Let $r = \frac{1}{1-hL_1}$. Then, it holds that $\sum_{j_1=1}^{J}\sum_{j_2=1}^{j_1}\cdots\sum_{j_k=1}^{j_{k-1}}(1-hL_1)^{j_k-J} = \psi(r, J, k)$, where ψ is the function given by Lemma 2.2. Applying Lemma 2.2 gives the desired result (2.5). ∎

Based on Theorem 2.3, we next study the asymptotic convergence rate of method (1.11) in the case that the step-size h approaches zero. Such an asymptotic convergence analysis gives more convenient estimate of the convergence rate. We will distinguish two situations:

1. method (1.11) is used with a fixed number of time points, i.e., J is fixed;
2. method (1.11) is used with a fixed length of time interval, i.e., $T = hJ$ is fixed.

2.1. Asymptotic convergence rate when J is fixed. We first consider the case that the iterative method (1.11) is applied to a fixed number of time steps, i.e., J is fixed. Based on Theorem 2.3 we have the following result.

THEOREM 2.4. *Under the assumptions of Theorem* 2.3*, it holds for* $k \gg 1$ *that*

$$\max_{0\le j\le J} \|e_j^k\|_2 \le \rho^k \max_{0\le j\le J} \|e_j^0\|_2, \tag{2.9}$$

where $\rho = \frac{h\widetilde{\eta}}{1-hL_1} = \mathcal{O}(h)$ *and* $\widetilde{\eta} = \eta L_2$*, provided* $h(L_1+\widetilde{\eta}) < 1$ *and* J *is fixed.*

Proof. To see this, we need to use

$$\binom{J+k-1}{k} \le (1+k)^{J-1},$$

which can be verified directly. For the case $L_1 = 0$, from Theorem 2.3 we have

$$\lim_{k\to\infty} \sqrt[k]{\binom{J+k-1}{k}} \le \lim_{k\to\infty} \left(e^{(J-1)\frac{\log(1+k)}{k}}\right) = 1. \tag{2.10}$$

This proves (2.9) for the case $L_1 = 0$ in Theorem 2.3. We next consider the case $L_1 \neq 0$. First, for $L_1 > 0$ it holds that $1 - hL_1 \in (0,1)$ and then by using Lemma 2.2 we have

$$\psi((1-hL_1)^{-1}, J, k) \le \frac{1}{(1-hL_1)^J}\psi(1,J,k) = \frac{1}{(1-hL_1)^J}\binom{J+k-1}{k}.$$

Then, by using (2.10) we have $\lim_{k\to\infty} \sqrt[k]{\psi((1-hL_1)^{-1},J,k)} \le 1$ since

$$\lim_{k} \sqrt[k]{(1-hL_1)^{-J}} = 1.$$

Second, for $L_1 < 0$ we have $(1-hL_1)^{-1} < 1$, and therefore,

$$\psi((1-hL_1)^{-1}, J, k) \le \psi(1,J,k) = \binom{J+k-1}{k}.$$

Again, by using (2.10) we have $\lim_{k\to\infty} \sqrt[k]{\psi((1-hL_1)^{-1},J,k)} \le 1$. ∎

2.2. Asymptotic convergence rate when $T = hJ$ is fixed. Theorem 2.4 implies that if the number of time points is fixed, a smaller step-size h results in faster convergence for method (1.11). For a fixed length of time interval, say $t \in [0,T]$ with T being a fixed quantity, this conclusion does not hold, because in this case J increases linearly as h reduces, and therefore, the limit in (2.10) is not correct. When hJ is fixed and $h \to 0$, method (1.11) reverts to the following continuous analogue:

$$0 \le y^{k+1}(t) \perp G(t, x^k(t), y^{k+1}(t)) \ge 0,\ \dot{x}^{k+1}(t) = F(t, x^{k+1}(t), y^{k+1}(t)),\ t \in (0,T). \tag{2.11}$$

THEOREM 2.5. *Under the assumption of Theorem* 2.3*, it holds that*

$$\|x^k(t) - x(t)\|_2 \le \frac{\max\{1, e^{L_1 t}\}(t\widetilde{\eta})^k}{k!} \sup_{t\in[0,T]} \|x^0(t) - x(t)\|_2,\ t \in (0,T), \tag{2.12}$$

where $\widetilde{\eta} = L_2\eta$ *and* $x(t)$ *is the converged solution of* (2.11).

Proof. In (2.11), according to Lemma 2.1 we represent $y^{k+1}(t)$ as $\mathcal{Y}(x^k(t))$ and then we rewrite (2.11) as

$$\dot{x}^{k+1}(t) = F(t, x^{k+1}(t), \mathcal{Y}(x^k(t))),\ t \in (0,T), \tag{2.13}$$

where $x^k(0) = x_0$ for all $k \geq 0$. Similarly, we can rewrite (1.1) as

$$\dot{x}(t) = F(t, x(t), \mathcal{Y}(x(t))), \quad t \in (0,T), \tag{2.14}$$

where $x(0) = x_0$. Let $e^k(t) = x^k(t) - x(t)$. Then, from (2.13) and (2.14) we have

$$\dot{e}^{k+1}(t) = F(t, x^{k+1}(t), \mathcal{Y}(x^k(t))) - F(t, x(t), \mathcal{Y}(x(t))),\ t \in (0,T), \tag{2.15}$$

where $e^k(0) = 0$ for all $k \geq 0$.

For the Euclidean inner product, it holds for any differentiable function $e(t) \neq 0$ that

$$\begin{cases} \frac{d\|e(t)\|_2^2}{dt} = 2\langle \dot{e}(t), e(t)\rangle, \\ \frac{d\|e(t)\|_2^2}{dt} = 2\|e(t)\|_2 \frac{d\|e(t)\|_2}{dt} \end{cases} \Rightarrow \|e(t)\|_2 \frac{d\|e(t)\|_2}{dt} = \langle \dot{e}(t), e(t)\rangle. \tag{2.16}$$

Applying this to differential equation (2.15) gives

$$\begin{aligned} &\|e^k(t)\|_2 \frac{d\|e^k(t)\|_2}{dt} = \langle \dot{e}^k(t), e^k(t)\rangle = \langle F(t, x^k, \mathcal{Y}(x^{k-1})) - F(t, x, \mathcal{Y}(x)), x^k - x\rangle \\ &= \langle F(t, x^k, \mathcal{Y}(x^{k-1})) - F(t, x, \mathcal{Y}(x^{k-1})), x^k - x\rangle \\ &+ \langle F(t, x, \mathcal{Y}(x^{k-1})) - F(t, x, \mathcal{Y}(x)), x^k - x\rangle \\ &\leq L_1 \|e^k(t)\|_2^2 + L_2 \eta \|e^{k-1}(t)\|_2 \|e^k(t)\|_2, \end{aligned}$$

where for the "$\leq$" we used (2.3) and Lemma 2.1. We have

$$\frac{d\|e^k(t)\|_2}{dt}(t) \leq L_1 \|e^k(t)\|_2 + \widetilde{\eta} \|e^{k-1}(t)\|_2, \quad t \in (0,T),$$

where $\widetilde{\eta} = L_2 \eta$ and $\zeta^k(0) = 0$ for all $k \geq 0$, i.e.,

$$\|e^k(t)\|_2 \leq \widetilde{\eta} \int_0^t e^{L_1(t-s)} \|e^{k-1}(s)\|_2 ds, \quad t \in (0,T). \tag{2.17}$$

Using (2.17) recursively gives

$$\begin{aligned} &\|e^k(t)\|_2 \\ &\leq \widetilde{\eta}^k \left(\int_0^t e^{L_1(t-s_1)} \int_0^{s_1} e^{L_1(s_1-s_2)} \cdots \int_0^{s_{k-1}} e^{L_1(s_{k-1}-s_k)} ds_k \cdots ds_1 \right) \sup_{t\in[0,T]} \|e^0(t)\|_2. \end{aligned}$$

The k-fold integral can be estimated as follows:

$$\begin{aligned} &\int_0^t e^{L_1(t-s_1)} \int_0^{s_1} e^{L_1(s_1-s_2)} \cdots \int_0^{s_{k-1}} e^{L_1(s_{k-1}-s_k)} ds_k \cdots ds_1 \\ &= e^{L_1 t} \int_0^t \int_0^{s_1} \cdots \int_0^{s_{k-1}} e^{-L_1 s_k} ds_k \cdots dt s_1 \\ &\leq e^{L_1 t} \max\{1, e^{-L_1 t}\} \int_0^t \int_0^{s_1} \cdots \int_0^{s_{k-1}} 1\, ds_k \cdots ds_1 = \max\{1, e^{L_1 t}\} \frac{t^k}{k!}. \end{aligned}$$

From this, we get $\zeta^k(t) \leq \frac{\max\{1, e^{L_1 t}\}(t\widetilde{\eta})^k}{k!} \sup_{t\in[0,T]} \zeta^0(t)$. ∎

Remark 2.2 (about the convergence rate). From Theorems 2.4 and 2.5, we see that method (1.11) has two different convergence rates. If J is fixed, the method converges with a rate $\rho = \mathcal{O}(h)$ and this implies that the method converges if a smaller step-size h is used. In particular, for $J = 1$ we know that the iterative algorithm (1.8) used for each single time point converges with a rate $\mathcal{O}(h)$, since in this case the multipoint algorithm (1.11) reduces to (1.8). If the length of time interval is fixed, i.e., the quantity T is fixed and J increases as h decreases, the factorial term $k!$ in (2.12) implies that the method converges superlinearly with a rate independent of h.

Remark 2.3. From Theorems 2.3 and 2.5 we see that a negative Lipschitz constant L_1 results in faster convergence, compared to a positive L_1. In the linear case, i.e., for DLCP (1.3), L_1 is negative when the matrix A is similar to its Jordan canonical form via an orthogonal transformation and the real parts of the eigenvalues are negative. This is often the case when the differential system arises from semidiscretizing a partial differential equation, e.g., the parabolic Signorini problems [14].

2.3. The case of nonlinear $\mathbf{P_0}$-function. The requirement that G is a uniform P-function can be slightly relaxed, namely G is a P_0-function of y, i.e., G still satisfies (2.1) but $L_0 = 0$. In this case, we can use the idea of *regularization* [10] to deal with DNCP (1.1). Precisely, with a small quantity $\varepsilon > 0$, we rewrite (1.1) as

(2.18)
$$\dot{x}(t) = F(t, x(t), y(t)),\ 0 \le y(t) \perp \widehat{G}(t, x(t), y(t)) \ge 0,\ \widehat{G}(t, x, y) = G(t, x, y) + \varepsilon y.$$

Now, it is clear that $\widehat{G}$ is a *uniform* P-function of y. It was proved in [10] that the solution $y_\varepsilon(t)$ of the regularized complementarity system in (2.18) approaches $y(t)$—a solution of the original complementarity system in (1.1) if (1.1) has a solution, when $\varepsilon \to 0^+$. Hence, with a suitable regularization parameter $\varepsilon > 0$, which is comparable with the temporal discretization error $\mathcal{O}(h)$, all the results obtained in this section are directly applicable to (2.18).

3. Convergence analysis in the case of linear complementarity. We now consider the case that the function $G(t, x, y)$ is not a P-function with respect to y, in the following DNCP form:

$$\dot{x}(t) = F(t, x(t), y(t)), \quad 0 \le y(t) \perp My(t) + \widetilde{G}(t, x(t)) \ge 0, \tag{3.1}$$

where $x(0) = x_0$ and $\widetilde{G} : \mathbb{R}_+ \times \mathbb{R}^m \to \mathbb{R}^n$ can be a nonlinear function of x. In particular, we consider the case that M is a positive semidefinite matrix. (The case that M is a Z-matrix can be treated similarly; see Remark 3.1.) DNCPs (3.1) have wide applications; see, e.g., [13, 26, 30]. For positive semidefinite matrix M, the LCS in (3.1) has a unique *least-norm* solution if the feasible set $FEA(M, \widetilde{G}(t, x)) := \{y | y \ge 0, \widetilde{G}(t, x) + My \ge 0\}$ is nonempty. Choosing the least-norm solution from the solution set leads to

$$\dot{x}(t) = F(t, x(t), y(t)), \quad y(t) = \operatorname{argmin}\{\|v\|_2 : 0 \le v \perp Mv + \widetilde{G}(t, x(t)) \ge 0\}. \tag{3.2}$$

Then, similar to (1.11) we define the following iterative method:

$$\begin{cases} y_j^{k+1} = \operatorname{argmin}\{\|v\|_2 : 0 \le v \perp Mv + \widetilde{G}(t_j, x_j^k) \ge 0\}, & j = 1, 2, \ldots, J, \\ x_j^{k+1} = x_{j-1}^{k+1} + hF(t_j, x_j^{k+1}, y_j^{k+1}), & j = 1, 2, \ldots, J. \end{cases} \tag{3.3}$$

The assumption that M is a positive semidefinite matrix, together with FEA(M, $\widetilde{G}(t, x)) \ne \emptyset$, guarantees the unique existence of the least-norm solution of the LCS,

but this cannot be used as a criterion in practice, because the solution $x(t)$ is not known a priori, and therefore, it is difficult to justify whether the feasible set is empty or not. Our task in this section is to answer the following two questions:

1. Under what conditions does DNCP (3.2) have a unique solution $(x(t), y(t))$?
2. Under what conditions does the iterative method (3.3) converge?

The following lemma plays a central role for our analysis.

LEMMA 3.1 (Theorem 2.3 in [9]). *Let $M \in \mathbb{R}^{n\times n}$ be a positive semidefinite matrix and $q_1, q_2 \in \mathbb{R}^n$ such that $FEA(M, q_1) \neq \emptyset$ and $FEA(M, q_2) \neq \emptyset$. Then, we have $\|\mathcal{Y}(q_1) - \mathcal{Y}(q_2)\|_2 \leq \eta_0 \|q_1 - q_2\|_2$, where $\mathcal{Y}(q)$ denotes the least-norm solution of $0 \leq y \perp q + My \geq 0$ and $\eta_0 > 0$ is a constant.*

We assume that the nonlinear function $\widetilde{G}(t, x)$ satisfies the following Lipschitz condition:

$$\|\widetilde{G}(t, x_1) - \widetilde{G}(t, x_2)\|_2 \leq \eta_1 \|x_1 - x_2\|_2 \ \forall t \in \mathbb{R}_+,\ x_1, x_2 \in \mathbb{R}^m. \tag{3.4}$$

THEOREM 3.2 (unique existence of the solution of (3.2)). *For DNCP (3.2) with M being a positive semidefinite matrix, suppose $F(t, x, y)$ satisfies the Lipschitz condition (2.3) with $L_1 \in \mathbb{R}$ and $L_2 > 0$. Assume that the nonlinear function $\widetilde{G}$ satisfies the Lipschitz condition (3.4) and that there exist $T > 0$ and $\beta > 0$ such that*

$$FEA(M, \widetilde{G}(t, v)) \neq \emptyset \text{ for } t \in [0, T] \quad \text{and} \quad v \in \mathcal{B}(x_0, \beta) := \{v : \|v - x_0\|_2 \leq \beta\}. \tag{3.5a}$$

Then, the DNCP (3.2) has a unique least-norm solution $(x(t), y(t)) \in C^1(0, t) \times C(0, t)$ in the interval $t \in (0, T^*)$, with T^* being defined by

$$T^* = \begin{cases} T & \text{if } L_1 \leq -\frac{\widetilde{\eta}\beta + C_0}{\beta}, \\ \min\left\{T, \frac{1}{L_1} \log\left(1 + \frac{L_1 \beta}{\widetilde{\eta}\beta + C_0}\right)\right\} & \text{if } L_1 > -\frac{\widetilde{\eta}\beta + C_0}{\beta}, \end{cases} \tag{3.5b}$$

where $\widetilde{\eta} = L_2 \eta_0 \eta_1$, $C_0 = \max_{t\in[0,T]} \|F(x_0, \mathcal{Y}(\widetilde{G}(t, x_0))\|_2$ and $\mathcal{Y}(\widetilde{G}(t, x_0))$ denotes the least-norm solution of the complementarity system $0 \leq y \perp \widetilde{G}(t, x_0) + My \geq 0$.

The condition (3.5a) shall be used throughout this section. It is essentially the same assumption that the authors used in [9] to study the semismooth Newton method. With the quantity T^* defined by (3.5b) it holds that

$$\frac{e^{L_1 t} - 1}{L_1} \leq \frac{\beta}{\widetilde{\eta}\beta + C_0} \text{ for } t \leq T^*, \tag{3.6}$$

where for $L_1 = 0$ the left quantity is defined in the limit, i.e., $\lim_{L_1 \to 0} \frac{e^{L_1 t} - 1}{L_1} = t$.

Proof. Let $\{x^k(t), y^k(t)\}_{k\geq 0}$ be the functional sequence generated by

$$\begin{aligned} y^{k+1}(t) &= \operatorname{argmin}\{\|v\|_2 : 0 \leq v \perp Mv + \widetilde{G}(t, x^k(t)) \geq 0\}, \\ \dot{x}^{k+1}(t) &= F(t, x^{k+1}(t), y^{k+1}(t)), \end{aligned}$$

where $x^k(0) = x_0$ for $k \geq 0$. For the initial iterate, i.e., $k = 0$, we choose $x^0(t) = x_0$. We claim that starting from this initial guess every iterate $x^k(t)$ still lies in the ball $\mathcal{B}(x_0, \beta)$ under condition (3.5b).

For $k = 0$, since $x^0(t)$ lies in the ball $\mathcal{B}(x_0, \beta)$ we know that $y^1(t)$ is uniquely existent for $t \in [0, T]$. Hence, $\dot{x}^1(t) = F(t, x^1(t), \mathcal{Y}(\widetilde{G}(t, x^0(t))))$. We have

$$\begin{aligned} \frac{d[x^1(t) - x_0(t)]}{dt} = &\Big[F(t, x^1(t), \mathcal{Y}(\widetilde{G}(t, x^0(t)))) - F(t, x^0(t), \mathcal{Y}(\widetilde{G}(t, x^0(t))))\Big] \\ &+ F(t, x^0(t), \mathcal{Y}(\widetilde{G}(t, x^0(t)))). \end{aligned}$$

By using (2.16) and the Lipschitz condition for F (cf. (2.3)), we have

$$\frac{d\|x^1(t)-x_0(t)\|_2}{dt} \le L_1\|x^1(t)-x_0(t)\|_2 + C_0 \Rightarrow \|x^1(t)-x_0(t)\|_2 \le C_0\frac{e^{L_1 t}-1}{L_1}.$$

This, together with (3.6), implies that under condition (3.5b) the solution $x^1(t)$ lies in the ball $\mathcal{B}(x_0,\beta)$ for $t\in[0,T^*]$.

Suppose $x^k(t)\in\mathcal{B}(x_0,\beta)$ for $t\in[0,T^*]$. Then, we have

$$\begin{aligned}\frac{d[x^{k+1}(t)-x_0(t)]}{dt} &= \Big[F(t,x^{k+1}(t),\mathcal{Y}(\widetilde{G}(t,x^k(t)))) - F(t,x^0(t),\mathcal{Y}(\widetilde{G}(t,x^k(t))))\Big]\\ &\quad+\Big[F(t,x^0(t),\mathcal{Y}(\widetilde{G}(t,x^k(t)))) - F(t,x^0(t),\mathcal{Y}(\widetilde{G}(t,x^0(t))))\Big]\\ &\quad+F(t,x^0(t),\mathcal{Y}(\widetilde{G}(t,x^0(t)))).\end{aligned}$$

Similar to the deduction for $\|x^1(t)-x_0(t)\|_2$, by using Lemma 3.1 and the Lipschitz condition (3.4) it holds that

$$\frac{d\|x^{k+1}(t)-x_0(t)\|_2}{dt} \le L_1\|x^{k+1}(t)-x_0(t)\|_2 + L_2\eta_0\eta_1\|x^k(t)-x^0(t)\|_2 + C_0.$$

This implies

$$\begin{aligned}\|x^{k+1}(t)-x_0(t)\|_2 &\le \int_0^t e^{L_1(t-s)}[\widetilde{\eta}\|x^k(s)-x_0(s)\|_2 + C_0]ds\\ &\le \int_0^t e^{L_1(t-s)}(\widetilde{\eta}\beta + C_0)ds = (\widetilde{\eta}\beta+C_0)\frac{e^{L_1 t}-1}{L_1}.\end{aligned}$$

By using (3.6), it holds that $\|x^{k+1}(t)-x_0(t)\|_2\le\beta$ for $t\le T^*$, which implies $x^{k+1}(t)\in\mathcal{B}(x_0,\beta)$.

Now, for any $k\ge 0$ we have $x^k(t),x^{k+1}(t)\in\mathcal{B}(x_0,\beta)$ for $t\in[0,T^*]$, and therefore,

(3.7a)
$$\begin{aligned}\frac{d[x^{k+1}(t)-x^k(t)]}{dt} &= F(t,x^{k+1}(t),\mathcal{Y}(\widetilde{G}(t,x^k(t)))) - F(t,x^k(t),\mathcal{Y}(\widetilde{G}(t,x^k(t))))\\ &\quad+F(t,x^k(t),\mathcal{Y}(\widetilde{G}(t,x^k(t)))) - F(t,x^k(t),\mathcal{Y}(\widetilde{G}(t,x^{k-1}(t)))).\end{aligned}$$

By using (2.16) and the Lipschitz condition for F (cf. (2.3)), we get

$$\frac{d\|x^{k+1}(t)-x^k(t)\|_2}{dt} \le L_1\|x^{k+1}(t)-x^k(t)\|_2 + \widetilde{\eta}\|x^k(t)-x^{k-1}(t)\|_2, \tag{3.7b}$$

which, after an integration, gives $\|x^{k+1}(t)-x^k(t)\|_2\le\widetilde{\eta}\int_0^t e^{L_1(t-s)}\|x^k(s)-x^{k-1}(s)\|_2 ds$. Then, similar to the proof of Theorem 2.5 we have

$$\|x^{k+1}(t)-x^k(t)\|_2 \le \left(\widetilde{\eta}\max\{1,e^{L_1 t}\}\right)^k\frac{t^k}{k!}\max_{s\in[0,t]}\|x^1(s)-x^0(s)\|_2.$$

Clearly, for any $\epsilon>0$ there exists some integer $K>\widetilde{\eta}\max\{1,e^{L_1T^*}\}T^*$ such that

$$\frac{1}{1-\frac{\widetilde{\eta}\max\{1,e^{L_1T^*}\}T^*}{K+1}}\frac{\left(\widetilde{\eta}\max\{1,e^{L_1T^*}\}T^*\right)^K}{K!}\max_{t\in[0,T^*]}\|x^1(t)-x^0(t)\|_2<\epsilon. \tag{3.8}$$

Then, for any integers K_1 and K_2 satisfying $K_2 > K_1 \geq K$, it holds that

$$\begin{aligned}\|x^{K_2}(t) - x^{K_1}(t)\|_2 &\leq \sum\nolimits_{k=0}^{K_2-K_1-1} \|x^{K_1+k}(t) - x^{K_1+k+1}(t)\|_2 \\ &\leq \left(\sum\nolimits_{k=0}^{K_2-K_1-1} \frac{\left(\widetilde{\eta}\max\{1, e^{L_1T^*}\}T^*\right)^{K_1+k}}{(K_1+k)!}\right) \max_{t\in[0,T^*]} \|x^1(t) - x^0(t)\|_2 \\ &\leq \frac{1}{1 - \frac{\widetilde{\eta}\max\{1,e^{L_1T^*}\}T^*}{K+1}} \frac{\left(\widetilde{\eta}\max\{1, e^{L_1T^*}\}T^*\right)^{K_1}}{K_1!} \max_{t\in[0,T^*]} \|x^1(t) - x^0(t)\|_2.\end{aligned}$$

This, together with (3.8), implies $\|x^{K_2}(t) - x^{K_1}(t)\|_2 < \epsilon$. Therefore, $\{x^k(t)\}_{k\geq 0}$ is a Cauchy sequence for $t \in [0, T^*]$. Let $\lim_{k\to\infty} x^k(t) = x^*(t)$. Then, it is easy to know that $x^*(t)$ is the solution of DNCP (3.2) for $t \in [0, T^*]$.

To prove uniqueness of $x^*(t)$, we suppose $x_1^*(t)$ and $x_2^*(t)$ are solutions of (3.2). Then, similar to (3.7a)–(3.7b) we have $\frac{d\|x_1^*(t)-x_2^*(t)\|_2}{dt} \leq (L_1 + \widetilde{\eta})\|x_1^*(t) - x_2^*(t)\|_2$. This gives $\frac{d\left(e^{-(L_1+\widetilde{\eta})t}\|x_1^*(t)-x_2^*(t)\|_2\right)}{dt} \leq 0$. Since $x_1^*(0) = x_2^*(0)$, it holds that $\|x_1^*(t) - x_2^*(t)\|_2 \leq 0$. Hence, $x_1^*(t) \equiv x_2^*(t)$ for $t \in [0, T^*]$. ∎

We next analyze the convergence of the iterative method (3.3). For the uniform P-function case, the Lipschitz continuity of $\mathcal{Y}(q)$ plays a central role for proving the convergence of method (1.11). Comparing Lemma 3.1 to Lemma 2.1, we see that, to ensure such a Lipschitz continuity in the LCS case with positive semidefinite coefficient matrix M, we need $\mathrm{SOL}(M, q) \neq \emptyset$ as an additional condition. Therefore, it is clear that for the least-norm iterative method (3.3) if $\mathrm{FEA}(M, \widetilde{G}(t, x_j^k)) \neq \emptyset$ $(\forall j = 1, \ldots, J)$ for each iteration, the results given by Theorems 2.3–2.5 still hold.

THEOREM 3.3. *For DNCP* (3.2) *with M being a positive semidefinite matrix, let $F(t, x, y)$ satisfy the Lipschitz condition* (2.3) *and $\widetilde{G}$ satisfy* (3.4). *Suppose* (3.5a) *holds for some constants $T > 0$ and $\beta > 0$. Then, the least-norm iterative method* (3.3) *for the time points $\{t_j\}_{j=0}^{J^*}$ is well-defined, i.e.,*

$$\mathrm{FEA}(M, \widetilde{G}(t, x_j^k)) \neq \emptyset \quad \forall j \in \{0, 1, \ldots, J^*\},\ \forall k \geq 1, \tag{3.9a}$$

provided $\{x_j^0 = x_0\}_{j=1}^{J}$, $T^* \leq T$, and the following conditions are satisfied:

$$hL_1 < 1, \quad \begin{cases} J^* = J & \text{if } L_1 \leq -\frac{\widetilde{\eta}\beta + C_0}{\beta}, \\ J^* = \min\left\{J, \left[\frac{\log\left(1+\frac{\beta L_1}{\widetilde{\eta}\beta+C_0}\right)}{\log((1-hL_1)^{-1})}\right]\right\} & \text{if } L_1 > -\frac{\widetilde{\eta}\beta + C_0}{\beta}, \end{cases} \tag{3.9b}$$

where $J = \frac{T}{h}$, $C_0 = \max_{j=0}^{J} \|F(t, x_0, \mathcal{Y}(\widetilde{G}(t_j, x_0)))\|_2$, $\widetilde{\eta} = L_2\eta_0\eta_1$ (with η_0 being the constant given by Lemma 3.1 *and η_1 being given by* (3.4)*), and $[v]$ denotes the integer part of $v \in \mathbb{R}$.*

Similar to (3.6), under condition (3.9b) it is easy to verify that

$$\frac{C_0 + \widetilde{\eta}\beta}{L_1}\left[\frac{1}{(1-hL_1)^j} - 1\right] \leq \beta \ \ \forall j \in \{0, 1, \ldots, J^*\}. \tag{3.10}$$

Proof. The proof given below can be regarded as a discrete version of the proof of Theorem 3.2. Since $\{x_j^0\}_{j=0}^{J^*} \subseteq \mathcal{B}(x_0, \beta)$, from Lemma 3.1 we know that $\{\mathcal{Y}(\widetilde{G}(t_j, x_j^0))\}_{j=0}^{J^*}$ are uniquely existent. Hence, by using the first Lipschitz condition in (2.3) we have

$$\|x_j^1 - x_j^0\|_2 \leq \|x_{j-1}^1 - x_{j-1}^0\|_2 + hL_1\|x_j^1 - x_j^0\|_2 + hC_0,$$

i.e., $\|x_j^1 - x_j^0\|_2 \le \frac{1}{1-hL_1}\|x_{j-1}^1 - x_{j-1}^0\|_2 + \frac{h}{1-hL_1}C_0$. Since $x_0^1 = x_0^0 = x_0$, it holds that

$$\|x_j^1 - x_j^0\|_2 \le \frac{hC_0}{1-hL_1}\sum\nolimits_{l=0}^{j-1}\frac{1}{(1-hL_1)^l} = \frac{C_0}{L_1}\left(\frac{1}{(1-hL_1)^j} - 1\right). \tag{3.11}$$

Hence, from (3.10) we get $\|x_j^1 - x_j^0\|_2 \le \beta$ for $j \in \{0, 1, \ldots, J^*\}$ and thus $x_j^1 \in \mathcal{B}(x_0, \beta)$.

To perform an induction proof, we assume $x_j^k \in \mathcal{B}(x_0, \beta)$ for $j = 0, 1, \ldots, J^*$. Then, by using Lemma 3.1 the sequence $\{\mathcal{Y}(\widetilde{G}(t_j, x_j^k))\}_{j=0}^{J^*}$ is uniquely existent. Therefore,

$$\|x_j^{k+1} - x_j^0\|_2 \le \|x_{j-1}^{k+1} - x_{j-1}^0\|_2 + hL_1\|x_j^{k+1} - x_j^0\|_2 + h\widetilde{\eta}\|x_j^k - x_j^0\|_2 + hC_0.$$

Similar to (3.11), this gives

$$\|x_j^{k+1} - x_j^0\|_2 \le \frac{h(C_0 + \widetilde{\eta}\beta)}{1-hL_1}\sum\nolimits_{l=0}^{j-1}\frac{1}{(1-hL_1)^l} = \frac{C_0 + \widetilde{\eta}\beta}{L_1}\left(\frac{1}{(1-hL_1)^j} - 1\right).$$

Now, by using (3.10) again we have $\|x_j^{k+1} - x_0\|_2 \le \beta$, i.e., $x_j^{k+1} \in \mathcal{B}(x_0, \beta)$ for $j \in \{0, 1, \ldots, J^*\}$, and this completes the proof of (3.9a). ∎

Remark 3.1 (the case M is a Z-matrix). At the end of this section, we consider the case that the matrix M in (3.1) is a Z-matrix, which is another representative case in the field of complementarity problems [13, 16]. In this case, there exists a unique least-element solution[5] for the LCS in (3.1) if the feasible set $\mathrm{FEA}(M, \widetilde{G}(t,x)) := \{y | y \ge 0, \widetilde{G}(t,x) + My \ge 0\}$ is nonempty; see [9]. This leads to the following least-element DNCP:

$$\dot{x}(t) = F(t, x(t), y(t)), \quad y(t) = \mathrm{argmin}\{\|v\|_1 : v \ge 0, \widetilde{G}(t, x(t)) + Mv \ge 0\}. \tag{3.12}$$

For (3.12), similar to (3.3) we define the iterative method as

$$\begin{cases} y_j^{k+1} = \mathrm{argmin}\{\|v\|_1 : v \ge 0, Mv + \widetilde{G}(t_j, x_j^k) \ge 0\}, & j = 1, 2, \ldots, J, \\ x_j^{k+1} = x_{j-1}^{k+1} + hF(t_j, x_j^{k+1}, y_j^{k+1}), & j = 1, 2, \ldots, J, \end{cases} \tag{3.13}$$

where $x_0^k = x_0$ for all $k \ge 0$. According to Theorem 2.3 in [9], the Lipschitz continuity of the least-element solution of the LCS $0 \le y \perp q + My \ge 0$ holds as well; see Lemma 3.1. Hence, the results obtained in this section are directly applicable to (3.1) if M is a Z-matrix.

4. Applications and numerical results. In this section, we show applications of the proposed iterative method for the nonsmooth circuit systems and the projected dynamic systems. For each application, we provide numerical results to validate the convergence properties of the proposed iterative method (1.11). The iteration stops when

$$\max\nolimits_{0 \le j \le J} \|x_j^k - x_j\|_2 \le 10^{-8}, \tag{4.1}$$

where $\{x_j\}_{j=1}^J$ denotes the converged solution.

[5]The least-element solution $y_{\min}$ is a solution of $0 \le y \perp My + \widetilde{G}(t,x) \ge 0$ satisfying $y_{\min} \le y$ for all $y \in \mathrm{FEA}(M, \widetilde{G}(t,x))$.

4.1. The 4-diode bridge wave rectifier. The 4-diode bridge wave rectifier shown in Figure 4.1 is a widely studied circuit model; see, e.g., [1, 4, 10, 16, 18, 20]. The circuit consists of a capacitor $C > 0$ with randomly perturbed value and a nonlinear resistor R, which lie inside the bridge formed by four ideal diodes. Let x be the voltage for the capacitor. Let $V_{DF1,DF2}$ and $I_{DF1,DF2}$ be the voltages and currents for the diodes, respectively. Then, by applying Kirchhoff voltage/current laws, we get the modified nodal analysis state equation described by a DNCP as (see [1, Chapter 2] for more details)

$$\begin{aligned} &C(\xi)\dot{x}(t,\xi) = -R(x(t,\xi)) + By(t,\xi) + I_s(t),\\ &0 \le y(t,\xi) \perp Nx(t,\xi) + My(t,\xi) \ge 0, \end{aligned} \tag{4.2a}$$

where $x(0,\xi) = 0$, $C(\xi)$ denotes the capacitor with value perturbed by random variable $\xi = (\xi_1, \xi_2, \ldots, \xi_d)^\top$, $I_s(t)$ is the current source (i.e., $e(t)$ shown in Figure 4.1), and

$$B = [1,0,1,0], \quad N = \begin{pmatrix} 1\\0\\1\\0 \end{pmatrix}, \quad M = \begin{pmatrix} 0 & -1 & 0 & 0\\ 1 & 0 & 1 & -1\\ 0 & -1 & 0 & 0\\ 0 & 1 & 0 & 0 \end{pmatrix}. \tag{4.2b}$$

For the nonlinear resistor, the current source, and the random capacitor we use the data

$$\begin{aligned} &R(v) = e^{v/50} - 1 \text{ (Schottky resistor)}, \ I_s(t) = 10\sin(18\pi t + 2) - 0.5,\\ &C(\xi) = 1.8 + \frac{\xi_1}{1+\xi_2^2} \text{ with } \begin{cases} \xi_1 \in [-1,1], & \text{uniform distribution},\\ \xi_2 \in (-\infty,\infty), & \text{Gaussian distribution}. \end{cases} \end{aligned} \tag{4.2c}$$

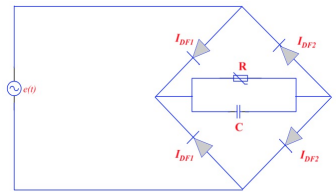

FIG. 4.1. *The circuit of a 4-diode bridge wave rectifier.*

For circuit (4.2a)–(4.2c), the random source of the capacitor can be caused by many factors, e.g., external environmental fluctuations such as temperature variation. Such an uncertainty may lead to remarkable performance variations at both circuit and system levels, and it cannot be ignored if we want to make a correct prediction of the behavior of the circuit. Here, we are interested in the basic stochastic information of the circuit, i.e., the mean values of $x(t,\xi)$ and $y(t,\xi)$ and the standard deviations. To this end, we use the technique of generalized polynomial chaos (gPC) expansion [34] together with the so-called *stochastic testing strategy* [35] to treat the random space. The gPC expansion technique has gained increasing interest in recent years thanks to its high order accuracy.

Briefly speaking, the gPC expansion technique transforms the random circuit system to a deterministic system with larger size as follows:

$$\widehat{C}X(t) = -R(X(t)) + \widehat{B}Y(t) + \widehat{I}_s(t), \quad 0 \le Y(t) \perp \widehat{N}X(t) + \widehat{M}Y(t) \ge 0. \tag{4.3}$$

where $X(t) \in \mathbb{R}^q$, $Y(t) \in \mathbb{R}^{4q}$, $\widehat{B} = B \otimes I_q$, $\widehat{C} = \mathrm{diag}(C_1, C_2, \ldots, C_q)$, $\widehat{M} = M \otimes I_q$, $\widehat{N} = N \otimes I_q$, $\widehat{I}_s(t) = I_q \otimes I_s(t)$, $I_q \in \mathbb{R}^{q\times q}$ is an identity matrix, and q denotes the number of the gPC basis functions (for the numerical results given below $q = 6$). The details for deriving the deterministic system (4.3) is given in Appendix B. Since M is a positive semidefinite matrix, the matrix $\widehat{M}$ in (4.3) is a positive semidefinite matrix as well. Thus, we choose for $Y(t)$ the least-norm solution from the solution set of the LCS in (4.3):

$$\begin{aligned} &\widehat{C}\dot{X}(t) = -R(X(t)) + \widehat{B}Y(t) + \widehat{I}_s(t), \\ &Y(t) = \mathrm{argmin}\left\{\|v\|_2 : 0 \le v \perp \widehat{M}v + \widehat{N}X(t) \ge 0\right\}. \end{aligned} \tag{4.4}$$

We now apply method (3.3) to (4.4) in the case $J = N_t$, i.e., we do Gauss–Seidel iterations for the whole time interval instead at each single time point (cf. Remark 2.2). In each iteration the nonlinear equations at $t = t_j$ arising in the discrete ODE system are solved by the `fsolve` command in MATLAB and the LCS is solved by using the `quadprog` command in MATLAB according to [8]. In Figure 4.2, we show the stochastic information of the circuit, i.e., the mean values and the standard derivations of $x(t,\xi)$ and $y(t,\xi)$.[6] For the complementarity variable $y(t,\xi)$, it holds that $\mathbb{E}(y_{2,4}(t,\xi)) \equiv 0$ and that $\mathbb{E}(y_1(t,\xi)) = \mathbb{E}(y_3(t,\xi))$, so we show the stochastic information for $y_1(t,\xi)$ and $y_3(t,\xi)$ together in the left column of Figure 4.2.

In Figure 4.3 we show the measured convergence rates of method (3.3) in two situations: in the left subfigure we consider the case that the length of time interval (i.e., T) is fixed and the step-size h varies; in the right subfigure we consider the case that the number of discrete time points (i.e., $J = N_t$) is fixed and h varies. We see that the method converges superlinearly with a robust convergence rate with respect to h, if T is fixed. In the case that J is fixed, the method converges faster when the step-size h becomes smaller. All these numerical results confirm Remark 2.2 very well.

As we mentioned in section 1, in each iteration of the iterative method (3.3) the computation of the complementarity system at all the discrete time points is parallel. We now show numerical results to illustrate that such an advantage can dramatically reduce the computation time. To carry out the parallel experiments, we use the following software and hardware configurations:

- CPU: Intel Core i7-3770K 3.5 GHz and 32 GB RAM using gcc 4.8.1. A single CPU was used for the sequential implementation of the proposed iterative methods. The codes were tested with gcc's fast math option (`ffast_math`).
- GPU: NVIDIA GeForce GTX 660 installed in a system with the above described CPU. The GPU operates at 1.10 GHz clock speed and consists of five multiprocessors (each contains 192 CUDA cores). We compiled the code using CUDA version 5.5 in combination with the gcc 4.8.1 compiler with fast math option (`use_fast_math`).

In Figure 4.4, we show the computation time (measured in seconds) for the semismooth Newton method and the iterative method (3.3). The semismooth Newton method is implemented in a sequential pattern by using a single CPU. For the new iterative method (3.3), the complementarity system is solved in parallel by GPU. Here, similar to Figure 4.3 we also consider two cases: the length of the time interval is fixed (i.e., $t \in [0,T]$ with $T = 0.2$), and the number of time points is fixed (i.e.,

[6] Such stochastic information can be obtained from the solutions $X(t)$ and $Y(t)$ due to the theory of the gPC expansion technique; see explanation in Appendix B.

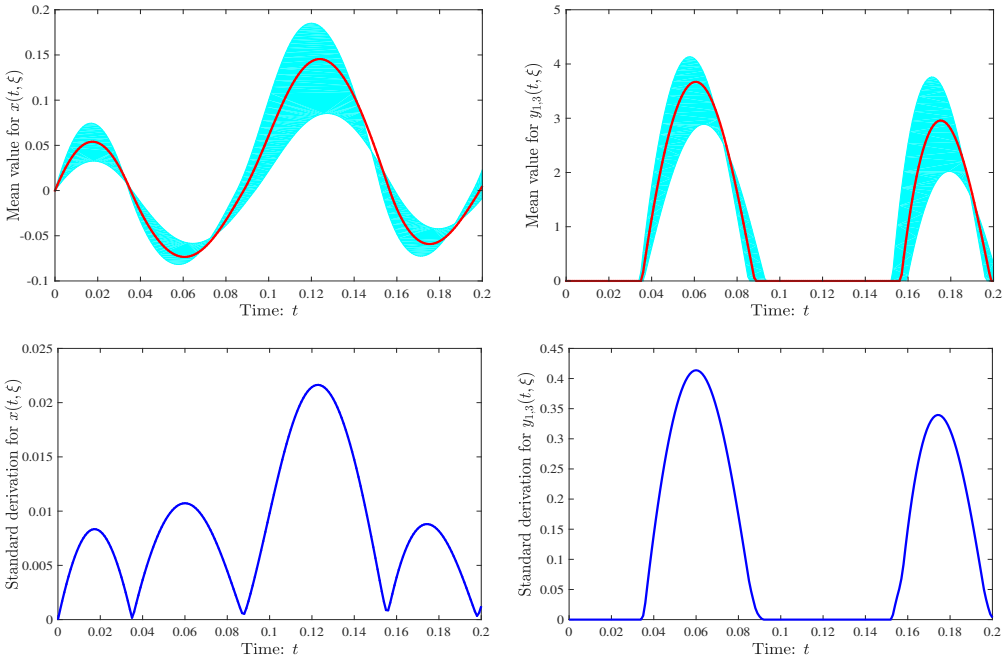

FIG. 4.2. *Stochastic information of the* 4*-diode bridge wave rectifier. Top row: the mean values of* $x(t,\xi)$ *(left) and* $y_{1,3}(t,\xi)$ *(right). The shading region is filled by the solutions obtained by Monte Carlo simulations with* $10,000$ *samples of* ξ*. Bottom row: standard derivations for* $x(t,\xi)$ *(left) and* $y_{1,3}(t,\xi)$ *(right).*

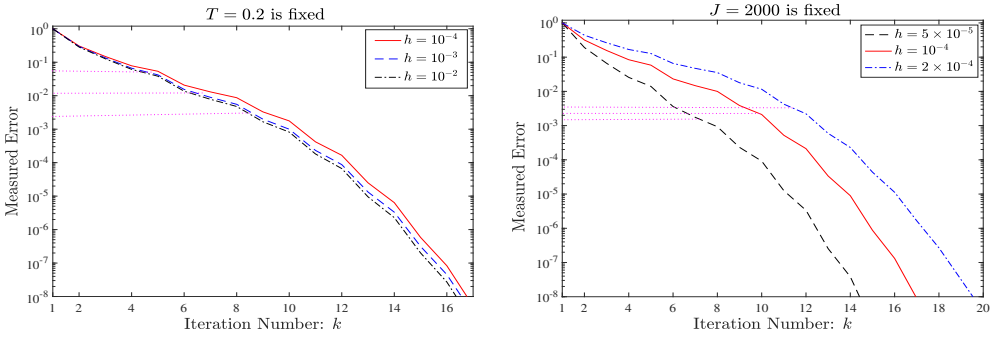

FIG. 4.3. *Measured convergence rates of the least-norm iterative method* (3.3) *applied to* (4.4)*. Left:* $T = 0.2$ *is fixed and* h *varies. Right:* $J = 2000$ *is fixed and* h *varies* ($J = N_t = T/h$ *denotes the number of discrete time points). The (dotted) horizontal line indicates where the method should stop in practice.*

$J = N_t = 2000$). From the results shown in Figure 4.4, we see that in the parallel circumstance the new iterative method needs much less computation time compared to the semismooth Newton method. For example, for the cases $T = 0.2$ and $h = 2^{-13}$ (left subfigure) the computation time for the semismooth Newton method is around 9587 seconds ($\approx$ 158 minutes), while by parallel computation the time for the new

iterative method is around 2023 seconds (≈ 34 minutes).

For the case $J = N_t = 2000$ (right subfigure), it is interesting to see that the computation time for the new iterative method (3.3) decreases as h becomes smaller; see Figure 4.4 on the right. This can be explained by using the convergence rate $\rho = \mathcal{O}(h)$ in the case of crossing a fixed number of time points, namely the new iterative method converges faster as h becomes smaller; see Remark 2.2.

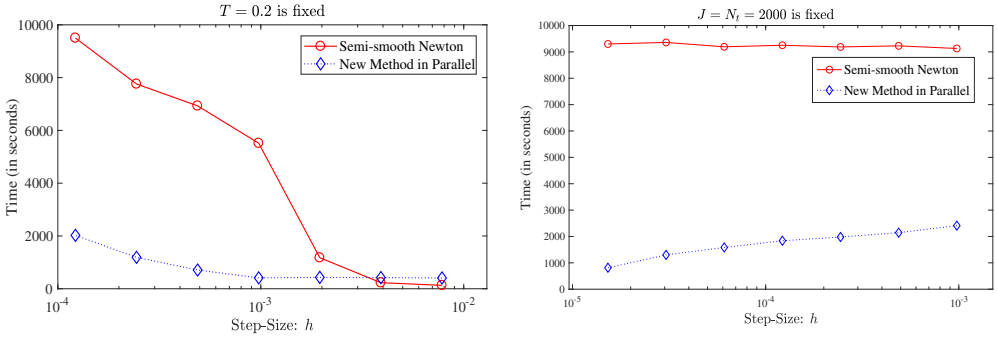

FIG. 4.4. *For DNCP* (4.4)*, comparison with respect to computation time between the semi-smooth Newton method and the new iterative method* (3.3) *with* $J = N_t$*. Left: the case* $T = 0.2$ *is fixed and* h *varies from* 2^{-7} *to* 2^{-13}*. Right: the case* $J = N_t = 2000$ *is fixed and* h *varies from* 2^{-10} *to* 2^{-16} *(in this case* T *decreases as* h *decreases).*

4.2. A projected dynamic system: The spatial price equilibrium. Let $\Omega \subset \mathbb{R}^m$ be a convex set. The projected dynamic system is described by the equation $\dot{x}(t) = F(t, x(t))$ on the interior of Ω, but on the boundary a modification is applied to prevent the solution from leaving the constraint set Ω. To be more specific, let $\mathcal{P}_\Omega$ be the projection operator that assigns to each vector $x \in \mathbb{R}^m$ the vector in Ω that is closest to x, i.e., $\mathcal{P}_\Omega(x) = \operatorname{argmin}_{v\in\Omega}\|x - v\|_2$. Then the projected dynamical system is defined by

$$\dot{x}(t) = \Pi_\Omega(x(t); F(t, x(t))) \text{ with } \Pi_\Omega(x, v) = \lim_{\delta\to 0^+} \frac{\mathcal{P}_\Omega(x + \delta v) - x}{\delta}. \tag{4.5}$$

In [23], Nagurney and Zhang mentioned several interesting applications of the projected dynamic systems, including the oligopolistic markets, the traffic networks, and the spatial price equilibrium.

In this section, we consider the spatial price equilibrium problem, which can be described as follows (see [23, Chapter 6] for more details). Suppose we have m supply markets and n demand markets involved in the production and consumption of a homogeneous commodity under perfect competition. Denote a typical supply market by i and a typical demand market by j. Let θ_j denote the supply and γ_i the supply price of the commodity at supply market i. Let d_j denote the demand, and let α_j be the demand price at demand market j. Let x_{ij} denote the nonnegative commodity shipment between the supply and demand market pair (i, j), and let c_{ij} denote the unit transaction cost associated with trading the commodity between i and j. The supply price at any supply market depends on the supply of the commodity at every supply market, that is, $\gamma = \gamma(\theta)$, where $\gamma = (\gamma_1, \ldots, \gamma_m)$, $\theta = (\theta_1, \ldots, \theta_m)^\top$. Similarly, the demand price at any demand market depends on the demand of the commodity at every demand market, i.e., $\alpha = \alpha(d)$, where $\alpha = (\alpha_1, \ldots, \alpha_n)^\top$ and $d = (d_1, \ldots, d_n)^\top$. The

unit transaction cost between a pair of supply and demand markets may depend upon the shipments of the commodity between every pair of markets, that is, $c = c(x)$, where $c = (c_{11}, \ldots, c_{1n}, \ldots, c_{m1}, \ldots, c_{mn})^\top$ and $x = (x_{11}, \ldots, x_{1n}, \ldots, x_{m1}, \ldots, x_{mn})^\top$. The supplies, demands, and shipments of the commodity must satisfy the following feasibility conditions:

$$\begin{aligned} \theta_i &= \sum\nolimits_{j=1}^{n} x_{ij}, \quad i = 1, 2, \ldots, m, \\ d_j &= \sum\nolimits_{i=1}^{m} x_{ij}, \quad j = 1, 2, \ldots, n. \end{aligned} \tag{4.6}$$

Define the *excess price* F_{ij}, between a pair of markets (i, j), as

$$F_{ij}(x) := \alpha_j(d) - \gamma_i(\theta) - c_{ij}(x). \tag{4.7}$$

In practice, we need to dynamically adjust the commodity shipments between the supply and demand markets. Under perfect competition, the rate of change of the commodity shipment between the supply and demand market pair (i, j) is in proportion to the excess price F_{ij}. This results in the following projected dynamic system:

$$\dot{x}(t) = \Pi_\Omega(x(t), F(x(t))), \tag{4.8}$$

where $F(x) = (F_{11}(x), \ldots, F_{1n}(x), \ldots, F_{m1}(x), \ldots, F_{mn}(x))^\top$. We are interested in the case that the commodity shipment x_{ij} is nonnegative and is not larger than a prescribed upper bound $x_{\max}$. In this case, the domain Ω can be described as

$$\Omega = \{x \in \mathbb{R}^{mn} | \phi_{ij}(x) \geq 0, i = 1, 2, \ldots, m, j = 1, 2, \ldots, n\}, \quad \phi_{ij}(x) := \begin{bmatrix} x_{ij} \\ x_{\max} - x_{ij} \end{bmatrix}. \tag{4.9}$$

According to [3, 19], the projected dynamic system (4.8) can be reformulated as a DNCP:

$$\dot{x}(t) = F(x(t)) + \Phi^\top(x(t))y(t), \quad 0 \leq \phi(x(t)) \perp y(t) \geq 0, \tag{4.10}$$

where Φ is the Jacobian matrix of ϕ, i.e., $\Phi = \frac{\partial \phi}{\partial_x}$. For the case (4.9), we have

$$\Phi = \begin{bmatrix} 1 & & & & \\ -1 & & & & \\ & 1 & & & \\ & -1 & & & \\ & & & \vdots & \\ & & & & 1 \\ & & & & -1 \end{bmatrix}_{2mn \times mn}, \quad \phi(x) = \Phi x + b, \quad b = \begin{bmatrix} 0 \\ x_{\max} \\ 0 \\ x_{\max} \\ \vdots \\ 0 \\ x_{\max} \end{bmatrix}_{2mn \times 1}. \tag{4.11}$$

Then, we can rewrite (4.10) as

$$\dot{x}(t) = F(x(t)) + \Phi^\top y(t), \quad 0 \leq \Phi x(t) + b \perp y(t) \geq 0. \tag{4.12}$$

Applying the implicit Euler method with a step-size h to (4.12) gives

$$x_{n+1} = x_n + hF(x_{n+1}) + h\Phi^\top y_{n+1}, \quad 0 \leq \Phi x_{n+1} + b \perp y_{n+1} \geq 0, \tag{4.13}$$

where $n = 0, 1, \ldots, N_t$. A direct application of the proposed iterative method to (4.13) leads to

$$0 \le \Phi x_{n+1}^k + b \perp y_{n+1}^{k+1} \ge 0, \quad x_{n+1}^{k+1} = x_n + hF(x_{n+1}^{k+1}) + h\Phi^\top y_{n+1}^{k+1},$$

where $k \ge 0$ is the iteration index and for $k = 0$ the quantity x_{n+1}^0 is an initial guess. However, this iteration process does not converge because the solution of the LCS is always zero, i.e., $y_{n+1}^{k+1} \equiv 0$ for $k \ge 0$. To apply the method correctly, we first obtain Φx_{n+1} from the discretized ODEs in (4.13) as

$$\Phi x_{n+1} = \Phi x_n + h\Phi F(x_{n+1}) + h\Phi\Phi^\top y_{n+1},$$

and then by substituting Φx_{n+1} into the LCS we get

$$0 \le \Phi x_n + h\Phi F(x_{n+1}) + h\Phi\Phi^\top y_{n+1} + b \perp y_{n+1} \ge 0.$$

We now define the following iterations:

$$\begin{aligned} &0 \le y_{n+1}^{k+1} \perp \left(\Phi x_n + h\Phi F(x_{n+1}^k) + b\right) + h\Phi\Phi^\top y_{n+1}^{k+1} \ge 0, \\ &x_{n+1}^{k+1} = x_n + hF(x_{n+1}^{k+1}) + h\Phi^\top y_{n+1}^{k+1}. \end{aligned} \tag{4.14}$$

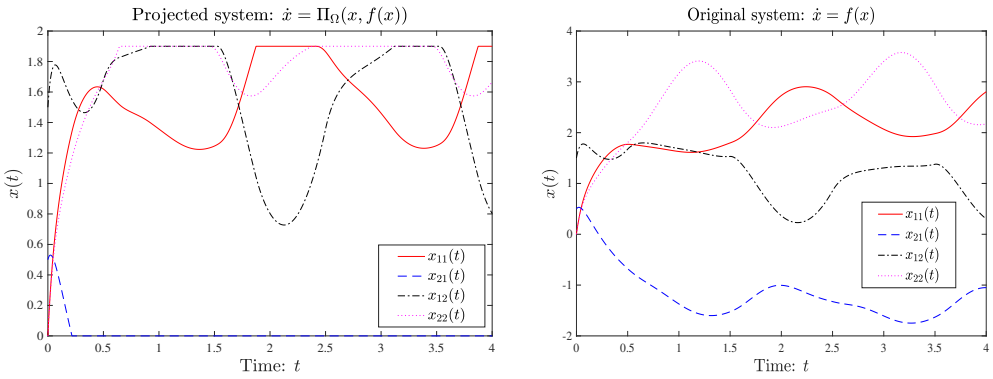

FIG. 4.5. *Evolution of the solutions of the projected system (left) and the original system (without projection on the domain Ω). Here, a step-size $h = 2^{-9}$ is used.*

We consider a spatial market problem consisting of two supply markets and two demand markers with the following data:

$$\begin{aligned} &\gamma_1(\theta) = 5\theta_1 + \theta_2 + 2, \quad \gamma_2(\theta) = 2\theta_2 + 1.5\theta_1 + 1.5, \\ &\alpha_1(d) = -2(1 + 0.25\sin(2\pi t))d_1 - 1.5d_2 + 28.75, \\ &\alpha_2(d) = -4(1 + 0.47\cos(\pi t))d_2 - d_1 + 41, \\ &c_{11}(x) = 0.01x_{11}^2 + 0.5x_{11}, \quad c_{12}(x) = 0.02x_{12}^2 + 2x_{12} + 7|\cos(\pi t)|, \\ &c_{21}(x) = 0.03x_{21}^2 + 3x_{21} + 16.25, \quad c_{22}(x) = 0.02x_{12}^2 + 2x_{12} + 11.5. \end{aligned} \tag{4.15}$$

We note that the quantities $\theta_1, \theta_2, d_1, d_2$ also depend on x_{ij} because of (4.6). With the initial condition $x = (0, 0.5, 1.5, 0)^\top$ and $x_{\max} = 1.9$, the numerical solutions of the projected dynamic system (4.8) are shown in Figure 4.5 on the left. For comparison, the solution of the original system $\dot{x} = F(x)$ is also shown in Figure 4.5 on the right.

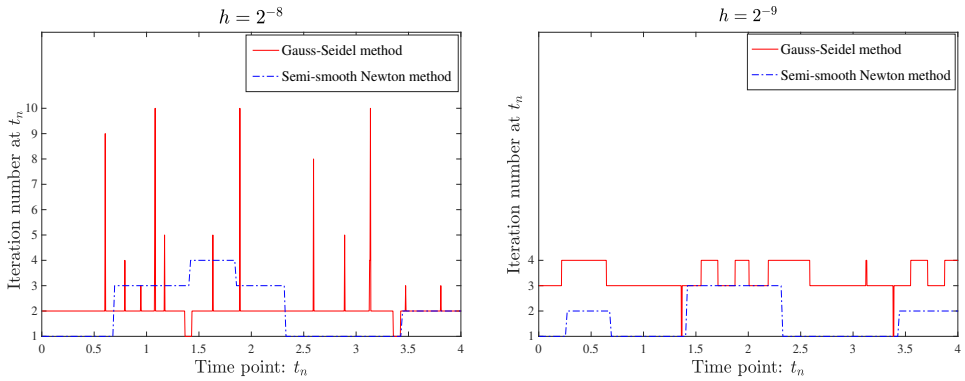

FIG. 4.6. *The iteration number (at each time point t_n) of the iterative method* (4.14) *and the semismooth Newton method mentioned in section* 1. *Left: $h = 2^{-8}$. Right: $h = 2^{-9}$.*

In Figure 4.6, we show the iteration number at each time point for the iterative method (4.14) and the semismooth Newton method mentioned in section 1. We see that when $h = 2^{-8}$ the method needs more iterations, while if we reduce the step-size to $h = 2^{-9}$, it converges with a rate similar to the semismooth Newton method. For the proposed iterative method (4.14), the reduction of the iteration number confirms our theoretical analysis very well (cf. Remark 2.2).

5. Conclusions. We proposed an iterative method for DNCP and made a convergence analysis based on the one-sided Lipschitz condition for the ODE system and the classical Lipschitz condition for the complementarity system. We proved that the new iterative method has two different convergence properties. In the case when the method is applied to a fixed number of time points, i.e., J is fixed, it converges with a rate $\rho = O(h)$, and therefore, a smaller h results in a better convergence rate. In the case when the method is applied to a fixed length of time interval, i.e., T is fixed, the method converges superlinearly with a rate independent of h. The general idea behind the proposed iterative method is that we solve the complementarity system and the differential system separately via an iteration of Gauss–Seidel style and thus many existing numerical methods for each of these two systems can be used without changes. In particular, in some cases we can solve the complementarity system efficiently via optimization solvers. For DLCPs, the new iterative method avoids solving a lot of linear systems that need to form the matrix M_h for the direct method (cf. (1.6a)–(1.6b)) and the Clarke Jacobian matrix V_j^k for the semismooth Newton method (cf. (1.7)). For large-scale problems, this is an important advantage for saving memory storage and computation time. The time-integrator used in this paper is the Backward-Euler method, by which the precision of the numerical solution is only of oder $\mathcal{O}(h)$. To improve the precision we can solve the differential system by higher-oder implicit Runge–Kutta methods and then similar to (1.8) (or (1.11)) we solve the discrete differential and complementarity systems separately via iterations. The details on this aspect, especially the convergence analysis of the corresponding iterative algorithms and the computation of the discrete differential system, will be addressed in future work.

Appendix A. The proof of Lemma 2.2. The result for $\psi(1, J, k)$ (i.e., $r = 1$) is well-known and can be verified by routine calculation. For $r \neq 1$, if $k = 1$ we have $\sum_{j_1=1}^{J} r^{-j_1} = \frac{r^{-1}(1-r^{-J})}{1-r^{-1}} = \frac{r^{-J}}{1-r} - \frac{1}{1-r}$ and this implies that (2.4) holds for $k = 1$.

Suppose (2.4) holds some $k > 1$. Then, we have

$$\begin{aligned}
&\sum_{j_1=1}^{J}\sum_{j_2=1}^{j_1}\cdots\sum_{j_k=1}^{j_{k-1}}\sum_{j_{k+1}=1}^{j_k} r^{-j_{k+1}} \\
&= \sum_{j_1=1}^{J}\left[\sum_{j_2=1}^{j_1}\cdots\sum_{j_k=1}^{j_{k-1}}\sum_{j_{k+1}=1}^{j_k} r^{-j_{k+1}}\right] \\
&= \sum_{j_1=1}^{J}\left[\frac{r^{-j_1}}{(1-r)^k} - \sum_{l=1}^{k}\binom{j_1+k-l-1}{k-l}\frac{1}{(1-r)^l}\right] \\
&= \frac{\frac{r^{-J}}{1-r}-\frac{1}{1-r}}{(1-r)^k} - \sum_{l=1}^{k}\frac{1}{(1-r)^l}\left[\sum_{j_1=1}^{J}\binom{j_1+k-l-1}{k-l}\right] \\
&= \frac{r^{-J}}{(1-r)^{k+1}} - \frac{1}{(1-r)^{k+1}} - \sum_{l=1}^{k}\frac{1}{(1-r)^l}\binom{J+k-l}{k-l+1} \\
&= \frac{r^{-J}}{(1-r)^{k+1}} - \sum_{l=1}^{k+1}\frac{1}{(1-r)^l}\binom{J+k-l}{k-l+1},
\end{aligned}$$

where for the fourth equality we used a well-known identity about the binomial coefficients, namely

$$\sum_{l=1}^{p}\binom{l+\tilde{k}-1}{\tilde{k}} = \binom{p+\tilde{k}}{\tilde{k}+1}$$

for any $\tilde{k} \geq 0$. Thus, (2.4) also holds for $k+1$.

Appendix B. Details of gPC expansion technique. In this appendix, the show details about how to get the deterministic system (4.3) via the gPC expansion technique. The gPC seeks to obtain a global polynomial approximation for a parametric (random) function. Let $\phi_{q_l}(\xi_l)$ be the univariate orthogonal polynomial of degree q_l (the special form of $\phi_{q_l}(\xi_l)$ depends on the density function of the input random parameter ξ_l). Then, the high dimensional polynomial bases $\Phi_q(\xi)$ are constructed by tensorizing the one dimensional bases

$$\Phi_q(\xi) = \prod_{l=1}^{d}\phi_{q_l}(\xi_l). \tag{B.1}$$

Moreover, the bases are chosen as $\int \rho(\xi)\Phi_{p_1}(\xi)\Phi_{p_2}(\xi) = \delta_{p_1p_2}$, where $\rho(\xi)$ is the joint density function of the random parameters.

With the above gPC basis functions, we approximate the ξ-dependent functions $x(t,\xi)$ and $y(t,\xi)$ via the following truncated expansions:

$$x(t,\xi) \approx \sum_{l=1}^{q}\tilde{x}_l(t)\Phi_l(\xi), \quad y(t,\xi) \approx \sum_{l=1}^{q}\tilde{y}_l(t)\Phi_l(\xi), \tag{B.2}$$

where $\{\tilde{x}_q(t), \tilde{y}_q(t)\}_{l=1}^{q}$ are coefficients that we need to compute. If the maximal polynomial degree is p, the number of the basis gPC functions is $q = \frac{(p+d)!}{p!d!}$ (see e.g., [34]). If we order the gPC basis functions $\{\Phi_q(\xi)\}_{l=1}^{q}$ by the polynomial degree in an ascending order, e.g., in lexicographical order $deg(\Phi_1(\xi)) < deg(\Phi_2(\xi)) < \cdots < deg(\Phi_q(\xi))$, the mean values and the standard derivations can be calculated as

$$\begin{cases}
\mathbb{E}(x(t,\xi)) = \tilde{x}_1(t),\ \mathbb{E}(y(t,\xi)) = \tilde{y}_1(t), & \text{mean value,} \\
\sigma(x(t,\xi)) = \sqrt{\sum_{l\geq 2}^{q}|\tilde{x}_l(t)|^2},\ \sigma(y(t,\xi)) = \sqrt{\sum_{l\geq 2}^{q}|\tilde{y}_l(t)|^2}, & \text{standard derivation.}
\end{cases} \tag{B.3}$$

Next we introduce the basic idea of the stochastic testing (collocation) method in [35]. Let $\{\xi^1, \xi^1, \ldots, \xi^q\}$ be q collocation points. Then, by substituting gPC expansions (B.2) into the circuit equation (4.2a) and by imposing the collocation condition on the sample points, we get the following large-scale deterministic DNCP:

$$\begin{aligned} &C_i \sum\nolimits_{l=1}^{q} \dot{\tilde{x}}_l(t)\Phi_{i,l} = -R\left(\sum\nolimits_{l=1}^{q} \tilde{x}_l(t)\Phi_{i,l}\right) + B\sum\nolimits_{l=1}^{q} \tilde{y}_l(t)\Phi_{i,l} + I_s(t), \\ &0 \le \sum\nolimits_{l=1}^{q} \tilde{y}_l(t)\Phi_{i,l} \perp N\sum\nolimits_{l=1}^{q} \tilde{x}_l(t)\Phi_{i,l} + M\sum\nolimits_{l=1}^{q} \tilde{y}_l(t)\Phi_{i,l} \ge 0, \end{aligned} \tag{B.4}$$

where $i = 1, 2, \ldots, q$, $C_i = C(\xi^i)$, and $\Phi_{i,q} = \Phi_q(\xi^i)$. Notice that choosing a good collocation set to guarantee the well-posedness of the above scheme is not a trivial work. Here, we adopt the strategy in [35], by which the collocation points are chosen as a subset (with the largest possible contribution) of the tensor grid of Gaussian quadrature points. Other types of collocation methods such as sparse grid, least-squares, and compressed sampling can also be used, and one can refer to [21, 22] for more details.

By defining the symbols

$$\widehat{\Phi} = \begin{pmatrix} \Phi_{1,1} & \cdots & \Phi_{1,q} \\ \vdots & \ddots & \vdots \\ \Phi_{q,1} & \cdots & \Phi_{q,q} \end{pmatrix}, \quad X(t) = (\Phi \otimes I_m)\begin{pmatrix} \tilde{x}_1(t) \\ \vdots \\ \tilde{x}_q(t) \end{pmatrix}, \quad Y(t) = (\Phi \otimes I_n)\begin{pmatrix} \tilde{y}_1(t) \\ \vdots \\ \tilde{y}_q(t) \end{pmatrix},$$

$$\widehat{B} = B \otimes I_q, \quad \widehat{C} = \operatorname{diag}(C_1, C_2, \ldots, C_q) \otimes I_m,$$

$$\widehat{M} = M \otimes I_q, \quad \widehat{N} = N \otimes I_q, \quad \widehat{I}_s(t) = I_q \otimes I_s(t),$$

we can rewrite (B.4) as (4.3). When $\{X(t), Y(t)\}$ are ready, we can compute $\{\tilde{x}_q(t), \tilde{y}_q(t)\}_{l=1}^{q}$ as follows:

$$\begin{pmatrix} \tilde{x}_1(t) \\ \vdots \\ \tilde{x}_q(t) \end{pmatrix} = (\widehat{\Phi}^{-1} \otimes I_m)X(t), \quad \begin{pmatrix} \tilde{y}_1(t) \\ \vdots \\ \tilde{y}_q(t) \end{pmatrix} = (\widehat{\Phi}^{-1} \otimes I_m)Y(t).$$

Then, we can explore the stochastic information of the circuit according to (B.3).

For the numerical experiments in section 4.1, we set $p = 2$ (the maximal polynomial degree of the basis gPC functions), and therefore, we have $q = 6$ gPC functions in total. For the reader's convenience, we list them as follows:

$$\Phi_1(\xi) = 1, \quad \Phi_2(\xi) = \xi_1, \quad \Phi_3(\xi) = \frac{1}{2}(3\xi_1^2 - 1),$$
$$\Phi_4(\xi) = \xi_2, \quad \Phi_5(\xi) = \xi_1\xi_2, \quad \Phi_6(\xi) = \xi_2^2 - 1.$$

Acknowledgment. The authors are very grateful to the anonymous referees for their careful reading of a preliminary version of the manuscript and for their valuable suggestions, which greatly improved the quality of this paper.

REFERENCES

[1] V. Acary, O. Bonnefon, and B. Brogliato, *Nonsmooth Modeling and Simulation for Switched Circuits*, Lect. Notes Electr. Eng. 69, Springer, Dordrecht, 2011.

[2] S. Billups and M. C. Ferris, *Solutions to affine generalized equations using proximal mappings*, Math. Oper. Res., 24 (1999), pp. 219–236.

[3] B. Brogliato, A. Daniilidis, C. Lemaréchal, and V. Acary, *On the equivalence between complementarity systems, projected systems and unilateral differential inclusions*, Systems Control Lett., 55 (2006), pp. 45–51.

[4] M. K. Camlibel, *Complementarity Methods in the Analysis of Piecewise Linear Dynamical Systems*, Ph.D. thesis, Center for Economic Research, Tilburg University, Tilburg, The Netherlands, 2001.
[5] M. K. Camlibel, W. P. M. H. Heemels, and J. M. Schumacher, *Consistency of a time stepping method for a class of piecewise-linear networks*, IEEE Trans. Circuits Systems I. Fund. Theory Appl., 49 (2002), pp. 349–357.
[6] M. K. Camlibel, J.-S. Pang, and J. Shen, *Lyapunov stability of complementarity and extended systems*, SIAM J. Optim., 17 (2006), pp. 1056–1101, https://doi.org/10.1137/050629185.
[7] X. Chen, Z. Nashed, and L. Qi, *Smoothing methods and semismooth methods for nondifferentiable operator equations*, SIAM J. Numer. Anal., 38 (2000), pp. 1200–1216, https://doi.org/10.1137/S0036142999356719.
[8] X. Chen and S. Xiang, *Sparse solutions of linear complementarity problems*, Math. Program. Ser. A, 159 (2016), pp. 539–556.
[9] X. Chen and S. Xiang, *Newton iterations in implicit time-stepping scheme for differential linear complementarity systems*, Math. Program. Ser. A, 138 (2013), pp. 579–606.
[10] X. Chen and Z. Wang, *Convergence of regularized time-stepping methods for differential variational inequalities*, SIAM J. Optim., 23 (2013), pp. 1647–1671, https://doi.org/10.1137/120875223.
[11] X. Chen and Z. Wang, *Differential variational inequality approach to dynamic games with shared constraints*, Math. Program. Ser. A, 146 (2014), pp. 379–408.
[12] C. Christof, *Sensitivity analysis and optimal control of obstacle-type evolution variational inequalities*, SIAM J. Control Optim., 57 (2019), pp. 192–218, https://doi.org/10.1137/18M1183662.
[13] R. W. Cottle, J.-S. Pang, and R. E. Stone, *The Linear Complementarity Problem*, Academic Press, Boston, 1992.
[14] D. Danielli, N. Garofalo, A. Petrosyan, and T. To, *Optimal regularity and the free boundary in the parabolic Signorini problem*, Mem. Amer. Math. Soc., 249 (2017), 1181.
[15] P. Deuflhard, *Newton Methods for Nonlinear Problems: Affine Invariance and Adaptive Algorithms*, Springer Ser. Comput. Math. 35, Springer, Heidelberg, 2011.
[16] F. Facchinei and J.-S. Pang, *Finite-dimensional Variational Inequalities and Complementarity Problems* Volume II, Springer-Verlag, New York, 2003.
[17] M. C. Ferris and T. S. Munson, *Interfaces to PATH* 3.0*: Design, implementation and usage*, Comput. Optim. Appl., 12 (1999), pp. 207–227.
[18] S. Greenhalgh, V. Acary, and B. Brogliato, *On preserving dissipativity properties of linear complementarity dynamical systems with the θ-method*, Numer. Math., 125 (2013), pp. 601–637.
[19] W. P. M. H. Heemels, J. M. Schumacher, and S. Weiland, *Projected dynamical systems in a complementarity formalism*, Oper. Res. Lett., 27 (2000), pp. 83–91.
[20] L. Han, A. Tiwari, M. K. Camlibel, and J.-S. Pang, *Convergence of time-stepping schemes for passive and extended linear complementarity systems*, SIAM J. Numer. Anal., 47 (2009), pp. 3768–3796, https://doi.org/10.1137/080725258.
[21] F. Nobile, R. Tempone, and C. G. Webster, *An anisotropic sparse grid stochastic collocation method for partial differential equations with random input data*, SIAM J. Numer. Anal., 46 (2008), pp. 2411–2442, https://doi.org/10.1137/070680540.
[22] A. Narayan, J. D. Jakeman, and T. Zhou, *A Christoffel function weighted least squares algorithm for collocation approximations*, Math. Comp., 86 (2017), pp. 1913–1947.
[23] A. Nagurney and D. Zhang, *Projected Dynamical Systems and Variational Inequalities with Applications*, Kluwer, Dordrecht, 1996.
[24] J. Pang and J. Shen, *Strongly regular differential variational systems*, IEEE Trans. Automat. Control, 52 (2007), pp. 242–255.
[25] F. A. Potra, M. Anitescu, B. Gavrea, and J. Trinkle, *A linearly implicit trapezoidal method for integrating stiff multibody dynamics with contact, joints, and friction*, Internat. J. Numer. Methods Engrg., 66 (2006), pp. 1079–1124.
[26] J.-S. Pang and D. Stewart, *Differential variational inequalities*, Math. Program. Ser. A, 113 (2008), pp. 345–424.
[27] J. S. Pang and D. Chan, *Iterative methods for variational and complementarity problems*, Math. Program. Ser. A., 24 (1982), pp. 284–313.
[28] J. Shen, *Robust Non-zenoness of piecewise affine systems with applications to linear complementarity systems*, SIAM J. Optim., 24 (2014), pp. 2023–2056, https://doi.org/10.1137/130937068.

[29] J. M. Schumacher, *Complementarity systems in optimization*, Math. Program. Ser. B, 101 (2004), pp. 263–296.
[30] J. Shen and J.-S. Pang, *Linear complementarity systems: Zeno states*, SIAM J. Control Optim., 44 (2005), pp. 1040–1066, https://doi.org/10.1137/040612270.
[31] A. Tanwani, B. Brogliato, and C. Prieur, *Well-posedness and output regulation for implicit time-varying evolution variational inequalities*, SIAM J. Control Optim., 56 (2018), pp. 751–781, https://doi.org/10.1137/16M1083657.
[32] S. L. Wu and X. Chen, *A parallel iterative algorithm for differential linear complementarity problems*, SIAM J. Sci. Comput., 39 (2017), pp. A3040–A3066, https://doi.org/10.1137/16M1103749.
[33] Z. Wang and X. Chen, *An exponential integrator-based discontinuous Galerkin method for linear complementarity systems*, IMA J. Numer. Anal., 38 (2017), pp. 2145–2165.
[34] D. Xiu, *Numerical Methods for Stochastic Computations: A Spectral Method Approach*, Princeton University Press, Princeton, NJ, 2010.
[35] Z. Zhang, *Uncertainty Quantification for Integrated Circuits and Microelectromechanical Systems*, Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, MA, 2015.