SIAM J. OPTIM.
Vol. 23, No. 2, pp. 1167–1188

A GLOBALLY CONVERGENT PROBABILITY-ONE HOMOTOPY FOR LINEAR PROGRAMS WITH LINEAR COMPLEMENTARITY CONSTRAINTS*

LAYNE T. WATSON†, STEPHEN C. BILLUPS‡, JOHN E. MITCHELL§, AND DAVID R. EASTERLING¶

Abstract. A solution of the standard formulation of a linear program with linear complementarity constraints (LPCC) does not satisfy a constraint qualification. A family of relaxations of an LPCC, associated with a probability-one homotopy map, proposed here is shown to have several desirable properties. The homotopy map is nonlinear, replacing all the constraints with nonlinear relaxations of NCP functions. Under mild existence and rank assumptions, (1) the LPCC relaxations RLPCC(λ) have a solution for $0 \le \lambda \le 1$; (2) RLPCC(1) is equivalent to LPCC; (3) the Kuhn–Tucker constraint qualification is satisfied at every local or global solution of RLPCC(λ) for almost all $0 \le \lambda < 1$; (4) a point is a local solution of RLPCC(1) (and LPCC) if and only if it is a Kuhn–Tucker point for RLPCC(1); and (5) a homotopy algorithm can find a Kuhn–Tucker point for RLPCC(1). Since the homotopy map is a globally convergent probability-one homotopy, robust and efficient numerical algorithms exist to find solutions of RLPCC(1). Numerical results are included for some small problems.

Key words. complementarity, constraint qualification, globally convergent, homotopy algorithm, linear program, probability-one homotopy

AMS subject classifications. 65F10, 65F50, 65H10, 65K10, 90C33

DOI. 10.1137/11082868X

1. Introduction.

Problems in diverse areas can be formulated as mathematical programs with complementarity constraints (MPCCs). The recent paper by Pang [39] describes applications in deregulated electricity markets, mechanical systems with frictional contacts, genetic regulatory networks in cell biology, control theory, and bilevel optimization. The complementarity constraints result in disjunctive mathematical programs, with feasible regions that may consist of many disjoint pieces. Linear programs with complementarity constraints (LPCCs) play an analogous role in disjunctive programming to that of linear programs in nonlinear programming. The LPCC has many applications of its own, as surveyed by Hu, Mitchell, and Pang [23].

Complementarity constraints arise naturally in bilevel optimization, through the use of the Kuhn–Tucker conditions to express the requirement that a feasible point must solve the lower level problem. If the upper level problem is linear and if the lower level problem is either a linear program or a convex quadratic program, then

*Received by the editors March 28, 2011; accepted for publication (in revised form) March 7, 2013; published electronically June 10, 2013.
http://www.siam.org/journals/siopt/23-2/82868.html
†Departments of Computer Science and Mathematics, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061-0106 (lwatson@computer.org). This author's work was partially supported by National Science Foundation grant CCF-0726763, Air Force Research Laboratory grant FA8650-09-2-3938, and Air Force Office of Scientific Research grant FA9550-09-1-0153.
‡Department of Mathematics, University of Colorado at Denver, Denver, CO 80217-3364 (stephen.Billups@ucdenver.edu).
§Department of Mathematical Sciences, Rensselaer Polytechnic Institute, Troy, NY 12180-1590 (mitchj@rpi.edu). This author's research was partially supported by Air Force Office of Scientific Research grants FA9550-08-1-0081 and FA9550-11-1-0260.
¶Department of Computer Science, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061-0106 (dreast@mohawk.cs.vt.edu).

the bilevel optimization problem is equivalent to an LPCC [12]. Surveys of bilevel optimization include [10] and [11], and a natural application is in the solution of Stackelberg games. Other related work on bilevel optimization includes [8], [21], and [38]. Inverse convex quadratic optimization problems can be cast as bilevel programs, when it is desired to choose the linear part of the objective and the right-hand side of the quadratic program so that they and the solution to the quadratic program are close to target values. If proximity is measured using a linear norm, then the problem is equivalent to an LPCC. Algorithms for inverse quadratic programs are given in [62] and [65]. Parameter estimation problems also lead to LPCCs in certain situations, such as in cross-validated support vector regression problems [28]. Other applications of LPCCs include quantile minimization such as chance constrained programming [43] and minimization of value-at-risk [41] and nonconvex quadratic programming [7, 22].

MPCCs and even LPCCs are nonlinear programs that do not satisfy most of the standard constraint qualifications [16, 32, 40, 45], which poses challenges in the development of algorithms and in proofs of their convergence. Various concepts of different types of stationary points have been developed and optimality conditions have been derived [32, 63, 64]. Typically the aim is to show that an algorithm converges to a stationary point of a certain type under certain conditions. Standard nonlinear programming algorithms can be applied, provided the complementarity condition is either penalized or initially relaxed and then gradually tightened. Regularization [46] and penalty [25] methods have been applied, and Anitescu, Tseng, and Wright [3] showed that an elastic mode approach converges to a strongly stationary point under boundedness assumptions and a constraint qualification. Fukushima and Tseng [19] showed that an ϵ-active set method converges to a strongly stationary point under a constraint qualification. It was shown computationally [17] that an SQP method can be very effective at solving MPCCs, leading to theoretical analysis of SQP approaches. Local convergence to a strongly stationary point is proved in [18], and Jiang and Ralph [26] show that a smoothed SQP method converges globally under certain assumptions (including strict complementarity for one variant) and investigate convergence of a quasi-Newton variant of an SQP approach [27]. An interior point method using shifted barriers was shown to have fast local convergence under certain assumptions in [37]. The analysis of interior point approaches to MPCCs was refined by Leyffer, Lopéz-Calva, and Nocedal [30], who showed global convergence to a strongly stationary point under standard assumptions. Other related work on MPCCs includes [14] and [31].

The homotopy method presented here uses nonlinear complementarity problem (NCP) functions. Leyffer [29] has proposed the use of NCP functions to solve MPCCs and has shown that the set of stationary points is unchanged with the use of certain NCP functions. He also proves local convergence to strongly stationary points and uses a filter SQP method to achieve global convergence in practice. Fang, Leyffer, and Munson [15] have recently proposed a pivoting method for LPCCs, generalizing the simplex method for LP. Hu et al. [24] have described a logical Benders decomposition method for determining a global optimal solution to an LPCC and for verifying infeasibility or unboundedness if appropriate.

The application of simplicial methods, a discrete version of continuous homotopy methods [2], to complementarity problems dates to the 1970s [13, 36, 44]. The first application of the modern probability-one homotopy theory [9, 49] to complementarity problems was in 1979 [50] and to optimization in general in 1980 [51]. Since then the development of homotopy methods (and interior point methods, which can be viewed as a variant of homotopy methods) in optimization has blossomed—[20, 42, 47, 48, 53, 54, 55, 58], just to mention a few homotopy references. The application of classical continuation, homotopy algorithms, and probability-one homotopy algorithms (see [20] or [55] for a discussion of the distinction between these three) to linear complementarity problems was thoroughly explored in [57], based on the theory in [52, 56].

The probability-one homotopy theory and algorithms in [61] have recently been successfully applied to general mixed complementarity problems [1, 4, 5, 6]. The present work is an outgrowth of that mixed complementarity work and the LPCC work in [24].

After defining the notation and terminology in section 2, the proposed relaxation and homotopy map are presented in section 3, followed by the convergence theory in section 4. Section 5 gives a few numerical results on very small problems, just as a sanity check. The conclusion addresses an alternative formulation, some issues not considered, and future directions.

2. Notation.

Let E^n denote real n-dimensional Euclidean space and $E^{m\times n}$ the set of real $m \times n$ matrices. For subsets $I \subset \{1,\ldots, m\}$, $J \subset \{1, \ldots, n\}$, a matrix $A \in E^{m\times n}$ and vector $x \in E^n$, $A_{I\cdot}$ denotes the rows of A indexed by I, $A_{\cdot J}$ is the columns of A indexed by J, A_{IJ} is the submatrix of A with row indices in I and column indices in J, and x_J is the subvector of x indexed by J. x_i denotes components of $x \in E^n$, $x > 0$ means all $x_i > 0$, $x \geqq 0$ means all $x_i \geqq 0$, and $x \geq 0$ means $x \geqq 0$ and $x \neq 0$. For $x, y \in E^n$, $x \perp y$ means the inner product $x^t y = \sum_{i=1}^n x_i y_i = 0$.

2.1. Stationarity conditions for LPCCs.

Let $c \in E^n$, $d \in E^m$, $A \in E^{k\times n}$, $B \in E^{k\times m}$, $f \in E^k$, $q \in E^m$, $N \in E^{m\times n}$, $M \in E^{m\times m}$. The problem under consideration is to find $x \in E^n$, $y \in E^m$ that solve

$$\text{(LPCC)} \qquad \begin{aligned} \min \quad & c^t x + d^t y \\ \text{subject to} \quad & Ax + By \geqq f, \\ & 0 \leqq y \perp (q + Nx + My) \geqq 0. \end{aligned}$$

In general, (LPCC) can be (i) infeasible, (ii) feasible with $c^t x + d^t y$ unbounded below, or (iii) feasible with $c^t x + d^t y$ bounded below.

Many different stationarity concepts have been proposed for general MPCCs. The two that are most relevant for LPCCs are *strong stationarity* and *Bouligand stationarity* [15]. Let $w := q + Nx + My$, so the complementarity condition can be represented as $y_i w_i = 0$ for $i = 1,\ldots, m$. Given a feasible point (x, y, w), say that the ith complementarity relationship is satisfied strictly if $y_i + w_i > 0$; otherwise, it is degenerate. Let $D(x, y, w)$ denote the set of degenerate indices.

Definition 1. *A feasible point* (x, y, w) *is* strongly stationary *if there exist dual multipliers* $\pi \in E^k$, $\rho \in E^m$, *and* $\mu \in E^m$ *satisfying*

$$\begin{aligned} c - A^T\pi - N^T\mu &= 0, \\ d - B^T\pi - M^T\mu - \rho &= 0, \\ 0 \leqq Ax + By - f \perp \pi &\geqq 0, \\ w_i > 0 \Rightarrow \mu_i &= 0, \quad i = 1,\ldots, m, \\ y_i > 0 \Rightarrow \rho_i &= 0, \quad i = 1,\ldots, m, \\ \mu_i \geqq 0, \quad \rho_i &\geqq 0, \quad i \in D(x, y, w). \end{aligned}$$

It should be noted that the dual multipliers for the complementarity terms are restricted to be nonnegative for only the degenerate indices. Any local minimizer of (LPCC) is a strongly stationary point, but the converse is not necessarily true. The set of local minimizers is equivalent to the set of Bouligand stationary points, defined next. The conditions for Bouligand stationarity involve looking at all the pieces of (LPCC) corresponding to a feasible point (x, y, w), where a piece is given by imposing either $y_i = 0$ or $w_i = 0$ for $i = 1, \ldots, m$. Each piece is a linear program.

Given a feasible point (x, y, w), let $P \subseteq D(x, y, w)$. In the piece defined by P, let $y_i = 0$ for $i \in P$ and $w_i = 0$ for $i \in D(x, y, w) \setminus P$. For $i \notin D(x, y, w)$, the complementarity restriction is chosen to agree with (x, y, w). The point is optimal in the piece defined by P if there exist dual multipliers $\pi \in E^k$, $\rho \in E^m$, and $\mu \in E^m$ satisfying

$$\begin{aligned}
c - A^T\pi - N^T\mu &= 0,\\
d - B^T\pi - M^T\mu - \rho &= 0,\\
0 \leqq Ax + By - f \perp \pi &\geqq 0,\\
w_i > 0 \Rightarrow \mu_i &= 0, \quad i = 1, \ldots, m, \quad i \notin D(x, y, w),\\
y_i > 0 \Rightarrow \rho_i &= 0, \quad i = 1, \ldots, m, \quad i \notin D(x, y, w),\\
\rho_i &\geqq 0, \quad i \in P,\\
\mu_i &\geqq 0, \quad i \in D(x, y, w) \setminus P.
\end{aligned}$$

Definition 2. *The point* (x, y, w) *is* Bouligand stationary *if it is optimal in each piece corresponding to some* $P \subseteq D(x, y, w)$.

Strongly stationary points are Bouligand stationary points. Further, the two concepts coincide for nondegenerate points. A degenerate Bouligand stationary point may not be strongly stationary. Kuhn–Tucker points can be defined for the nonlinear programming formulation of (LPCC) given by imposing the complementarity relationship through the constraints $y_i w_i = 0$ for $i = 1, \ldots, m$; the Kuhn–Tucker points are then the strongly stationary points. The set of nondegenerate Kuhn–Tucker points is equivalent to the set of nondegenerate Bouligand stationary points and the set of nondegenerate local minimizers of (LPCC).

The complementarity relationship can be represented using a positively oriented NCP function.

Definition 3. *A continuous function* $\hat{\psi} : E \times E \to E$ *is called an* NCP function *if* $\hat{\psi}(a, b) = 0 \iff 0 \leqq a \perp b \geqq 0$*; it is* positively oriented *if* $\hat{\psi}(a, b) \geqq 0 \iff a \geqq 0$ *and* $b \geqq 0$.

In [50], a family of smooth positively oriented NCP functions based on [34] is defined by

$$(\hat{\psi}^{(k)}) \qquad \hat{\psi}^{(k)}(a, b) := -|a - b|^k + a^k + b^k, \quad k > 0, \quad k \text{ odd}.$$

Observe that this function is C^{k-1}; moreover, for $b > 0$, $\hat{\psi}^{(k)}(\cdot, b)$ is monotone strictly increasing and is onto E for odd $k \geqq 3$.

For any positively oriented NCP function $\hat{\psi}$, problem (LPCC) is equivalent to the problem

$$(\hat{\psi}\text{LPCC}) \qquad \begin{aligned}
\min \quad & c^t x + d^t y\\
\text{subject to} \quad & Ax + By \geqq f,\\
& w - (q + Nx + My) = 0,\\
& y \geqq 0,\\
& w \geqq 0,\\
& \hat{\psi}(y_i, w_i) \leqq 0, \quad i = 1, \ldots, m.
\end{aligned}$$

It was shown by Leyffer [29] that the set of Kuhn–Tucker points for formulations of MPCCs using positively oriented NCP functions coincide with the set of strongly stationary points provided the gradients of the NCP functions satisfy three criteria:

(ψ1) $\frac{\partial\hat{\psi}}{\partial a} = \frac{\partial\hat{\psi}}{\partial b} = 0$ if $a = b = 0$;

(ψ2) $\frac{\partial\hat{\psi}}{\partial a} > 0$ and $\frac{\partial\hat{\psi}}{\partial b} = 0$ if $a = 0$ and $b > 0$;

(ψ3) $\frac{\partial\hat{\psi}}{\partial a} = 0$ and $\frac{\partial\hat{\psi}}{\partial b} > 0$ if $a > 0$ and $b = 0$.

It is easily verified that $\hat{\psi}^{(k)}$ satisfies these criteria for odd $k \geqq 3$. Thus, the sets of strongly stationary points for (LPCC) and Kuhn–Tucker points for ($\hat{\psi}$LPCC) coincide. The homotopy algorithm proposed here, under mild existence and rank assumptions, obtains a nondegenerate Kuhn–Tucker point for ($\hat{\psi}$LPCC), so this is a strongly stationary point and a local minimizer for (LPCC).

2.2. Probability-one homotopy theory. Let $F : E^n \to E^n$ be a C^2 function. The main idea of a probability-one homotopy method for solving $F(x) = 0$ is to embed F in a homotopy map $\rho_a : [0, 1) \times E^n \to E^n$ such that $\rho_a(1, x) = F(x)$ and $\rho_a(0, \cdot)$ has a known zero x^0. The algorithm works by tracking a zero curve γ of ρ_a emanating from $(0, x^0)$ until reaching an accumulation point $(1, \bar{x})$ of γ, which then yields a solution $\bar{x}$ to $F(x) = 0$. The theoretical foundation of all probability-one homotopy methods is the following differential geometry definition and theorem.

DEFINITION 4. *Let $U \subset E^m$ and $V \subset E^n$ be nonempty open sets, and let $\rho : U \times [0, 1) \times V \to E^n$ be a C^2 map. ρ is said to be* transversal to zero *if the Jacobian matrix $\nabla\rho$ has full rank on $\rho^{-1}(0)$.*

PARAMETRIZED SARD'S THEOREM (see [9]). *If $\rho(a, \lambda, x)$ is transversal to zero, then for almost all $a \in U$, the map*

$$\rho_a(\lambda, x) = \rho(a, \lambda, x)$$

is also transversal to zero; i.e., with probability one, the Jacobian matrix $\nabla\rho_a(\lambda, x)$ has full rank on $\rho_a^{-1}(0)$.

The importance of this theorem is that for almost all $a \in U$, the zero set $\rho_a^{-1}(0)$ consists of smooth, nonintersecting curves in $[0, 1) \times V$. These curves either are closed loops or have endpoints in $\{0\} \times V$ or $\{1\} \times V$ or go to infinity. Another important consequence is that these curves have finite arc length on any compact subset of $[0, 1) \times V$.

As a rule, ρ_a should be constructed such that there is a unique zero $(0, x^0)$ at $\lambda = 0$ and such that rank $\nabla_x\rho_a(0, x^0) = n$. This latter requirement ensures that the zero curve γ emanating from $(0, x^0)$ is not a closed loop. Thus, the curve must either approach $\lambda = 1$ or wander off to infinity.

The following section defines a homotopy map for solving the Kuhn–Tucker conditions of a reformulation of (LPCC).

3. Probability-one homotopy map. A family of relaxations of (LPCC) is defined for $0 \leqq \lambda \leqq 1$ by

$$\text{(RLPCC)} \qquad \begin{aligned} &\min && c^t x + d^t y \\ &\text{subject to} && Ax + By - f + (1 - \lambda)f^0 \geqq 0, \\ & && \psi\big(\lambda, x, y, q^0, g^0\big) = 0, \end{aligned}$$

where given an initial guess x^0 and $y^0 > 0$, $f^0 > 0$ is chosen so $Ax^0 + By^0 - f + f^0 > 0$, and

$$\psi_i\big(\lambda, x, y, q^0, g^0\big) = \hat{\psi}^{(5)}\left(\left((1 - \lambda)q^0 + q + Nx + My\right)_i, y_i\right) - (1 - \lambda)g_i^0, \;\; i = 1, \dots, m,$$

in which $g^0 > 0$, and q^0 is chosen such that $\psi\big(0, x^0, y^0, q^0, g^0\big) = 0$. Note that q^0 is uniquely determined; this follows because $y_i^0 > 0$, so $\hat{\psi}^{(5)}(\cdot, y_i^0)$ is monotone strictly increasing and onto E for each i. With f^0, g^0, and q^0 so chosen, $\big(x^0, y^0\big)$ is a strictly feasible (the feasible set has a nonempty relative interior) point for the problem (RLPCC) with $\lambda = 0$. Moreover, since $\hat{\psi}^{(5)}$ is positively oriented and $g^0 > 0$, it follows that for any feasible point (x, y) of (RLPCC) with $0 \leqq \lambda < 1$, $y > 0$ and $\big((1-\lambda)q^0 + q + Nx + My\big) > 0$. Observe that the problem (RLPCC) with $\lambda = 1$ is equivalent to (LPCC).

For $0 \leqq \lambda \leqq 1$ define

$$\Omega_\lambda = \big\{(x, y) \mid Ax + By - f + (1-\lambda)f^0 \geqq 0\big\}.$$

Assume (LPCC) has a solution and the set $\Omega_1 \neq \emptyset$ is bounded. (The boundedness assumption is a sufficient, but not necessary, condition for the results proved in this paper. When x and y are physical quantities, it is reasonable to assume they are bounded, so explicit bound constraints can be added if necessary. Unbounded models can occur during the modeling process and can often indicate that aspects of the system have been overlooked in the model. Proper selection of the bounds may require an iterative approach.) Note that

$$\Omega_0 \supset \Omega_{\lambda_1} \supset \Omega_{\lambda_2} \supset \Omega_1 \neq \emptyset \text{ for } 0 < \lambda_1 < \lambda_2 < 1,$$

all the sets Ω_λ are bounded, and the interior int $(\Omega_\lambda) \neq \emptyset$ for $0 \leqq \lambda < 1$.

Looking ahead, it will be shown that

(C1) (RLPCC) has a solution for $0 \leqq \lambda < 1$;

(C2) for almost all $f^0 > 0$ and $g^0 > 0$ chosen as above, the Kuhn–Tucker constraint qualification is satisfied at a solution of (RLPCC) for almost all $0 \leqq \lambda < 1$;

(C3) a nondegenerate local solution $(\bar{x}, \bar{y})$ of (RLPCC) at $\lambda = 1$ together with some $\bar{u} \geqq 0$ and $\bar{v}$ satisfies the Kuhn–Tucker conditions;

(C4) a globally convergent, probability-one homotopy algorithm can find a Kuhn–Tucker point $(\bar{x}, \bar{y}, \bar{u}, \bar{v})$ of (RLPCC) at $\lambda = 1$, which is a Kuhn–Tucker point for a reformulation of (LPCC).

The Kuhn–Tucker conditions for the constraint $Ax + By - f + (1-\lambda)f^0 \geqq 0$ become

$$\begin{gathered} Ax + By - f + (1-\lambda)f^0 \geqq 0, \\ u \geqq 0, \\ u^t\big(Ax + By - f + (1-\lambda)f^0\big) = 0, \end{gathered}$$

which are replaced by $\phi\big(\lambda, x, y, u, f^0, h^0\big) = 0$, where $E^k \ni h^0 > 0$ and

$$\phi_i\big(\lambda, x, y, u, f^0, h^0\big) = \hat{\psi}^{(3)}\left(\big(Ax + By - f + (1-\lambda)f^0\big)_i, u_i\right) - (1-\lambda)h_i^0,$$

$i = 1, \ldots, k$. Note that the definitions of ψ and ϕ involve different members of the $\hat{\psi}^{(k)}$ family of NCP functions. The choice of k in each case is dictated by the need to make the homotopy map C^2, which is required to apply the parametrized Sard's theorem. Choosing $k = 3$ suffices for ϕ, but $k = 5$ is needed for ψ since $\nabla\psi$ is used in the definition of the homotopy map.

Let $a = \big(x^0, y^0, f^0, g^0, h^0\big)$, from which q^0 is uniquely determined as noted previously. Similarly $\phi_i\big(0, x^0, y^0, u, f^0, h^0\big) = 0$ uniquely determines $u_i = u_i^0 > 0$.

The proposed probability-one homotopy map is

$$\rho_a(\lambda,x,y,u,v)=\rho(a,\lambda,x,y,u,v)$$
$$=\begin{bmatrix}\lambda\left[\begin{pmatrix}c\\d\end{pmatrix}-\begin{pmatrix}A^t\\B^t\end{pmatrix}u+\left(\nabla_{(x,y)}\psi\right)^t v\right]+(1-\lambda)\begin{pmatrix}x-x^0\\y-y^0\end{pmatrix}\\ \phi\left(\lambda,x,y,u,f^0,h^0\right)\\ \psi\left(\lambda,x,y,q^0,g^0\right)+\left(1-\tanh\left(\frac{60\lambda}{1-\lambda}\right)\right)\left(v-v^0\right)\end{bmatrix}.$$

Choose v^0 in the above homotopy map as the minimum norm least squares solution to

$$\begin{pmatrix}c\\d\end{pmatrix}-\begin{pmatrix}A^t\\B^t\end{pmatrix}u^0+\left(\nabla_{(x,y)}\psi\left(0,x^0,y^0,q^0,g^0\right)\right)^t v=0.$$

Some observations on this homotopy map ρ_a follow. Let E^n_{++} denote the positive orthant in n-dimensional Euclidean space E^n. For brevity, write $w=(x,y,u,v)$.

(O1) $\rho(a,\lambda,x,y,u,v)$ is C^2 and transversal to zero. To see this, observe that the columns of $\nabla\rho$ corresponding to x^0, y^0, h^0, and g^0 form an $(n+2m+k)\times(n+2m+k)$ diagonal matrix with $\lambda-1$ on the diagonal. Thus, $\nabla\rho$ has full row rank for $\lambda\in[0,1)$, so ρ is transversal to zero. Therefore, by the parametrized Sard's theorem, for almost all

$$a=\left(x^0,y^0,f^0,g^0,h^0\right)\in E^n\times E^m_{++}\times E^k_{++}\times E^m_{++}\times E^k_{++}$$

the map $\rho_a:[0,1)\times E^{n+2m+k}\to E^{n+2m+k}$ is also transversal to zero.

(O2) The term $(v-v^0)$ is necessary to force a unique solution to $\rho_a(\lambda,x,y,u,v)=0$ at $\lambda=0$. $\psi\left(0,x^0,y^0,q^0,g^0\right)=0$ by construction. From a theoretical perspective, any strictly decreasing C^2 function $\zeta:[0,1]\to E$ satisfying $\zeta(0)=1$ and $\zeta(\lambda)=o(1-\lambda)$ as $\lambda\to1$ can be used in place of the function $\left(1-\tanh\left(\frac{60\lambda}{1-\lambda}\right)\right)$. The choice given here has proved to be very effective in practice since it is approximately zero for $\lambda\geqq0.1$. This forces the last component of ρ_a to essentially be ψ for $\lambda\geqq0.1$, modeling the complementarity constraint better for intermediate λ.

(O3) By construction $\rho_a(0,x,y,u,v)=0$ has the unique solution $w^0=\left(x^0,\,y^0,\,u^0,\,v^0\right)$, and the (square) Jacobian matrix $D_w\rho_a\left(0,x^0,y^0,u^0,v^0\right)$ is invertible. (To see this, observe that $\frac{\partial}{\partial u_i}\phi_i(0,x^0,y^0,u^0,v^0)>0$, since $u^0>0$. Thus, $D_w\rho_a\left(0,x^0,y^0,u^0,v^0\right)$ is a lower triangular matrix with nonzero diagonal elements.) The invertibility of $D_w\rho_a\left(0,x^0,y^0,u^0,v^0\right)$ ensures that the zero set of ρ_a intersects $\lambda=0$ transversally, i.e., the zero set is not tangent to the hyperplane $\lambda=0$.

(O4) From (O1) and (O3), the probability-one homotopy theory (derived from the parametrized Sard's theorem) says that for almost all vectors a described in (O1), there exists a C^1 zero curve γ of ρ_a, emanating from $\left(0,x^0,y^0,u^0,v^0\right)$, along which the Jacobian matrix $D\rho_a$ has full rank. γ does not return to $\lambda=0$, does not intersect itself, is disjoint from any other zeros of ρ_a, cannot just stop at some point with $\lambda<1$, and has finite arc length in every compact subset of $[0,1)\times E^{n+2m+k}$. γ either wanders off to infinity or has an accumulation point $(1,\bar{x},\,\bar{y},\,\bar{u},\,\bar{v})$ at $\lambda=1$, which is then a Kuhn–Tucker point for (RLPCC) with $\lambda=1$.

(O5) If γ is bounded, then a globally convergent homotopy algorithm consists of tracking γ from $\left(0, x^0, y^0, u^0, v^0\right)$ at $\lambda = 0$ to $(1, \bar{x}, \bar{y}, \bar{u}, \bar{v})$ at $\lambda = 1$, and this is guaranteed to work almost surely (for almost all choices of a). Note further than λ need not increase monotonically along γ; thus tracking γ is not simple continuation in λ from 0 to 1.

(O6) If γ is not bounded, it may still lead to a Kuhn–Tucker point of (RLPCC) at $\lambda = 1$. First observe that (λ, x, y) along γ is bounded, since $0 \leqq \lambda \leqq 1$ and $\phi = 0$ along γ implies that $(x, y) \in \Omega_\lambda \subset \Omega_0$, which is bounded. The shifted NCP map ϕ used here for linear constraints is exactly the same as that used in [54], where under the same assumption that Ω_λ is nonempty and bounded for all $0 \leqq \lambda \leqq 1$, but without the v and ψ terms in ρ_a, it is proven that either x, y, and u remain bounded along γ, or (λ, x, y) along γ accumulate at $(1, \bar{x}, \bar{y})$ and there exists $\bar{u}$ such that $(1, \bar{x}, \bar{y}, \bar{u})$ is a zero of the homotopy map and $(\bar{x}, \bar{y})$ solves the nonlinear program. In this sense, γ is said to *reach the point* $(1, \bar{x}, \bar{y}, \bar{u})$. In the following section, similar results are established under the assumption that either u or v is bounded along γ.

4. Convergence theorems. Homotopy convergence theorems simultaneously prove the existence of a solution; hence it is not necessary to assume (LPCC) has a solution. The standing assumption that Ω_1 is nonempty and bounded, however, cannot be easily dispensed with. By (O5) and (O6), the goal is to show that either γ is bounded or there exists a point $(\bar{x}, \bar{y}, \bar{u}, \bar{v})$ such that $\rho_a(1, \bar{x}, \bar{y}, \bar{u}, \bar{v}) = 0$.

LEMMA 0. *Assume $\Omega_1 \neq \emptyset$ is bounded. For any $\tilde{\lambda} \in [0, 1)$, γ is bounded for $0 \leqq \lambda \leqq \tilde{\lambda}$.*

Proof. Suppose v is not bounded along γ for $0 \leqq \lambda \leqq \tilde{\lambda}$. Then there exists a sequence $\left\{\left(\lambda_k, x^k, y^k, u^k, v^k\right)\right\} \subset \gamma$ for which $\{v^k\}$ has no accumulation point and $\lambda_k \leqq \tilde{\lambda}$ for all k. By (O6) λ, x, and y are bounded on γ, so it may be assumed, after passing to a subsequence, that $\left\{\left(\lambda_k, x^k, y^k\right)\right\}$ converges to an accumulation point $(\bar{\lambda}, \bar{x}, \bar{y})$. Note that $\bar{\lambda} \leqq \tilde{\lambda} < 1$, so by the third component of ρ_a,

$$\lim_{k\to\infty} v^k = v^0 - \psi\left(\bar{\lambda}, \bar{x}, \bar{y}, q^0, g^0\right) \;/\; \left(1 - \tanh(60\bar{\lambda}/(1-\bar{\lambda}))\right),$$

contradicting the unboundedness of $\{v^k\}$. Therefore, v must be bounded along γ.

Since λ, x, y, and v are bounded along γ for $0 \leqq \lambda \leqq \tilde{\lambda}$, every term in the first component of ρ_a is bounded except possibly the term containing u. Now the argument (passing to an accumulation point of (λ, x, y, v) along γ, observe that u unbounded implies the component ϕ of $\rho_a = 0$ on γ is also unbounded) used in the proof of Theorem 5.1 of [54] for $\bar{\lambda} < 1$ to prove u is also bounded applies in this context (the component of ρ_a involving ψ plays no role in this proof). Therefore u along γ is also bounded for $0 \leqq \lambda \leqq \tilde{\lambda}$. $\square$

LEMMA 1. *Assume $\Omega_1 \neq \emptyset$ is bounded and u is bounded along γ. v is bounded along γ if $\nabla_{(x,y)}\psi$ always has full row rank along γ for $0 \leqq \lambda \leqq 1$.*

Proof. Suppose v is not bounded along γ. Then there exists a sequence $\left\{\left(\lambda_k, x^k, y^k, u^k, v^k\right)\right\} \subset \gamma$ for which $\{v^k\}$ has no accumulation point. Since λ, x, y, and u are bounded on γ, the sequence $\left\{\left(\lambda_k, x^k, y^k, u^k\right)\right\}$ has an accumulation point, which by Lemma 0 must be at $\lambda = 1$. By passing to a subsequence, it may be assumed that $\left\{\left(\lambda_k, x^k, y^k, u^k\right)\right\}$ converges to $(1, \bar{x}, \bar{y}, \bar{u})$. By assumption, $\left(\nabla_{(x,y)}\psi\right)^t$ always has full column rank, so v^k is the unique solution to

$$M_k v = -\frac{1-\lambda_k}{\lambda_k}\begin{pmatrix} x^k - x^0 \\ y^k - y^0 \end{pmatrix} - \begin{pmatrix} c \\ d \end{pmatrix} + \begin{pmatrix} A^t \\ B^t \end{pmatrix} u^k, \tag{L1}$$

where $M_k = \left(\nabla_{(x,y)}\psi\left(\lambda_k, x^k, y^k, q^0, g^0\right)\right)^t$. The right-hand side is bounded (since $\lambda_k \to 1$).

Let σ_k denote the smallest singular value of M_k. Since $\{v^k\}$ is unbounded, $\{\sigma_k\}$ must converge to zero. $\{M_k\}$ converges to $\overline{M} = \left(\nabla_{(x,y)}\psi\left(1,\ \bar{x},\ \bar{y},\ q^0,\ g^0\right)\right)^t$, which therefore has zero as a singular value, contradicting the assumption that $(\nabla_{(x,y)}\psi)^t$ always has full column rank. Hence v must be bounded along γ. $\square$

To understand the full rank assumption for $\nabla_{(x,y)}\psi$ along γ, $\nabla_{(x,y)}\psi$ must be studied in detail. By Lemma 0 and (O4), regardless of rank $\nabla_{(x,y)}\psi$, for any $\tilde{\lambda} < 1$, γ must eventually leave $[0, \tilde{\lambda}] \times E^{n+2m+k}$ and reach points with $\lambda > \tilde{\lambda}$. The term $v - v^0$ in ρ_a is both a blessing and a curse. $v - v^0$, or something similar, is required to uniquely determine v at $\lambda = 0$ and also to prevent γ from asymptotically approaching $\lambda = 0$ with $\|v^k\| \to \infty$. The curse is that for points $\left(\lambda_k,\ x^k,\ y^k,\ u^k,\ v^k\right)$ along γ,

$$\lim_{\lambda_k \to 1} \left(1 - \tanh\left(\frac{60\lambda_k}{1-\lambda_k}\right)\right)\left(v^k - v^0\right) = -\psi\left(1, \bar{x}, \bar{y}, q^0, g^0\right) \neq 0$$

is possible when $\|v^k\| \to \infty$, so the point $(\bar{x}, \bar{y})$ does not satisfy complementarity conditions, and $\nabla_{(x,y)}\psi\left(1,\ \bar{x},\ \bar{y},\ q^0,\ g^0\right)$ could be rank deficient. Ideally, as $\lambda_k \to 1$,

$$\psi\left(\lambda_k, x^k, y^k, q^0, g^0\right) = o(1 - \lambda_k),$$

and it is worthwhile to analyze the situation when $\psi = 0$ for $0 \ll \tilde{\lambda} \leqq \lambda < 1$, where both $(1-\lambda)q^0 + q + Nx + My > 0$ and $y > 0$ along γ. Examining all the cases for $\operatorname{sgn}\left(\left((1-\lambda)q^0 + q + Nx + My\right)_i - y_i\right)$ yields that

$$\nabla_{(x,y)}\psi\left(\lambda, x, y, q^0, g^0\right) = \left(\Sigma\, N, \quad \Sigma\, M + \Delta\right),$$

where $\Sigma = \Sigma\left(\lambda, x, y, q^0\right)$, $\Delta = \Delta\left(\lambda, x, y, q^0\right) \in E^{m\times m}$ are diagonal matrices with positive diagonal elements for (λ, x, y) along γ $(\lambda < 1)$. Explicit expressions for Σ and Δ are given later in the proof of Lemma 4. Now at $\lambda = 1$, assume also that the complementarity solution (x, y) is nondegenerate: $q + Nx + My + y > 0$. In this case

$$\begin{aligned} y_i = 0 &\Longrightarrow \Sigma_{ii} = 0, \quad \Delta_{ii} > 0; \\ y_i > 0 &\Longrightarrow \Sigma_{ii} > 0, \quad \Delta_{ii} = 0. \end{aligned}$$

Observe that if the complementarity solution is degenerate, say, $(q + Nx + My)_i + y_i = 0$, then the ith row of $\nabla_{(x,y)}\psi$ is zero.

Summarizing the preceding discussion, if $\psi\left(\lambda, x, y, q^0, g^0\right) = o(1-\lambda)$ along γ as $\lambda \to 1$ and the limit point $(1, \bar{x}, \bar{y})$ is a nondegenerate complementarity solution, then a sufficient condition that $\nabla_{(x,y)}\psi$ have full rank along γ for $0 \ll \tilde{\lambda} < \lambda \leqq 1$ is that

$$\text{(A1)} \qquad \operatorname{rank}\left(\Sigma\, N, \quad \Sigma\, M + \Delta\right) = m$$

for all nonnegative diagonal matrices Σ, $\Delta \in E^{m\times m}$ satisfying $\Sigma_{ii} + \Delta_{ii} > 0$ for $i = 1, \ldots, m$. There are numerous ways this condition could be satisfied, e.g.,

(RA1) rank $N = m$,
(RA2) M is a P-matrix (all principal minors are positive),
(RA3) M is positive definite,
(RA4) M is strictly diagonally dominant with positive diagonal elements.

Conversely, if v is bounded along γ, then

$$\psi\big(\lambda_k, x^k, y^k, q^0, g^0\big) = -\left(1 - \tanh\left(\frac{60\lambda_k}{1-\lambda_k}\right)\right)\big(v^k - v^0\big) = o(1-\lambda_k)$$

as $\lambda_k \to 1$ along γ. There is no obvious simple weak assumption to guarantee that v remains bounded along γ.

LEMMA 2. *Assume $\Omega_1 \neq \emptyset$ is bounded and u is bounded along γ. v is bounded along γ if*

$$\text{rank } \nabla_{(x,y)}\psi\big(1, x, y, q^0, g^0\big) = m$$

over Ω_1.

Proof. Suppose v is unbounded along γ . Using arguments similar to the proof of Lemma 1, there exists a sequence $\big\{\big(\lambda_k,\, x^k,\, y^k,\, u^k,\, v^k\big)\big\} \subset \gamma$ for which $\big\{\big(\lambda_k,\, x^k,\, y^k,\, u^k\big)\big\}$ converges to $(1, \bar{x}, \bar{y}, \bar{u})$ and $\{v^k\}$ is unbounded. Observe that $(\bar{x}, \bar{y}) \in \Omega_1$ since $\phi\big(\lambda_k,\, x^k,\, y^k,\, u^k,\, f^0,\, h^0\big) \to \phi\big(1,\, \bar{x},\, \bar{y},\, \bar{u},\, f^0,\, h^0\big) = 0$. Let $M_k = \big(\nabla_{(x,y)}\psi\big(\lambda_k,\, x^k,\, y^k,\, q^0,\, g^0\big)\big)^t$ and note that $\{M_k\}$ converges to $\overline{M} = \big(\nabla_{(x,y)}\psi\big(1,\, \bar{x},\, \bar{y},\, q^0,\, g^0\big)\big)^t$, which has rank m since $(\bar{x}, \bar{y}) \in \Omega_1$. By continuity, M_k has full rank for k sufficiently large; so for large k, by the first component of ρ_a, v^k is the unique solution to (L1). Since the right-hand side of (L1) is bounded and $\{v^k\}$ is unbounded, the smallest singular values of $\{M_k\}$ must converge to zero, contradicting the assumption that $\overline{M}$ has full rank. ∎

The preceding observations (O1)–(O6) and lemmas prove the following theorem.

THEOREM 1. *Let the set $\Omega_1 = \{(x, y) \mid Ax + By \geqq f\}$ be nonempty and bounded, and let*

$$\text{rank } \nabla_{(x,y)}\psi\big(1, x, y, q^0, g^0\big) = m$$

over the set Ω_1. Then for almost all

$$a = \big(x^0, y^0, f^0, g^0, h^0\big) \in E^n \times E^m_{++} \times E^k_{++} \times E^m_{++} \times E^k_{++}$$

there exists a C^1 zero curve γ of ρ_a, emanating from $\big(0, x^0, y^0, u^0, v^0\big)$, along which the Jacobian matrix $D\rho_a$ has full rank. γ does not return to $\lambda = 0$, does not intersect itself, is disjoint from any other zeros of ρ_a, and has finite arc length in every compact subset of $[0, 1) \times E^{n+2m+k}$. If either u or v is bounded along γ, then γ reaches a point $(1,\, \bar{x},\, \bar{y},\, \bar{u},\, \bar{v})$ (in the sense of (O6)*) at $\lambda = 1$, and $(\bar{x},\, \bar{y},\, \bar{u},\, \bar{v})$ is a Kuhn–Tucker point for* (RLPCC) *with $\lambda = 1$.*

Proof. By (O1), ρ is transversal to zero, so by the parametrized Sard's theorem, for almost all specified a, ρ_a is also transversal to zero. That is, $\nabla\rho_a$ has full rank on the zero set $\rho_a^{-1}(0)$. Therefore, the zero set consists of smooth, nonintersecting curves in $[0, 1) \times E^n \times E^m \times E^k \times E^m$. Let γ be the component of $\rho_a^{-1}(0)$ containing $(0,\, x^0,\, y^0,\, u^0,\, v^0)$. By (O4), γ either wanders off to infinity or has an accumulation point at $\lambda = 1$. If u is bounded along γ, then by Lemma 2, γ is bounded, so it reaches a point $(1,\, \bar{x},\, \bar{y},\, \bar{u},\, \bar{v})$, which is a Kuhn–Tucker point for (RLPCC) with $\lambda = 1$.

Suppose u is unbounded and v is bounded along γ. Then there exists a sequence $\big\{\big(\lambda_k,\, x^k,\, y^k,\, u^k,\, v^k\big)\big\} \subset \gamma$ for which $\big\{\big(\lambda_k,\, x^k,\, y^k,\, v^k\big)\big\}$ converges to $(1, \bar{x}, \bar{y}, \bar{v})$ and $\{u^k\}$ is unbounded. Then

$$\begin{pmatrix} c \\ d \end{pmatrix} - \begin{pmatrix} A^t \\ B^t \end{pmatrix} u^k + \big(\nabla_{(x,y)}\psi\big(1, \bar{x}, \bar{y}, q^0, g^0\big)\big)^t \bar{v} \to 0$$

and $\bar{u}$ can be constructed from u^k such that $\rho_a(1,\, \bar{x},\, \bar{y},\, \bar{u},\, \bar{v}) = 0$. ∎

The conclusion of Theorem 1 includes the conjecture (C4) mentioned earlier. Conjectures (C1)–(C3) concern the relationship between solutions of (LPCC), (RLPCC), and the Kuhn–Tucker point of Theorem 1. A proof of conjecture (C1) will be taken up next, using the machinery and results developed so far.

Due to technical requirements on the homotopy map ρ_a involving transversality and a unique solution at $\lambda = 0$, ρ_a could not simply have ψ, like ϕ, as a component. The import of this is that ψ is not zero everywhere along γ (unlike ϕ), and thus points along γ are not generally feasible points for the relaxed problems (RLPCC). Let (RLPCC(λ)) denote the problem (RLPCC) for a particular fixed λ. Fix $\tilde{\lambda}$, $0 \leqq \tilde{\lambda} < 1$, and let (MRLPCC($\lambda$)) denote the family of problems obtained from (RLPCC) by replacing $(1-\lambda)$ by $(1-\lambda)+\lambda(1-\tilde{\lambda})$ or, equivalently, λ by $\lambda\tilde{\lambda}$. Then (MRLPCC(1)) is the same as (RLPCC($\tilde{\lambda}$)), and the feasible set Ω_λ for (MRLPCC(λ)) matches up with $\Omega_{\lambda\tilde{\lambda}}$ for (RLPCC($\lambda\tilde{\lambda}$)). Define $\tilde{\psi}(\lambda,\cdot) = \psi(\lambda\tilde{\lambda},\cdot)$, $\tilde{\phi}(\lambda,\cdot) = \phi(\lambda\tilde{\lambda},\cdot)$, and define $\tilde{\rho_a}$ the same as ρ_a, except with $\tilde{\psi}$ and $\tilde{\phi}$ in lieu of ψ and ϕ. Let $\gamma(\tilde{\lambda})$ denote the zero curve of $\tilde{\rho_a}$ emanating from $(0,\, x^0,\, y^0,\, u^0,\, v^0)$.

THEOREM C1. *Assume* $\Omega_1 \neq \emptyset$ *is bounded. If rank* $\nabla_{(x,y)}\psi(\lambda,\cdot) = m$ *over* Ω_λ *for* $0 \leqq \lambda < 1$, *and either* u *or* v *is bounded along* $\gamma(\tilde{\lambda})$ *for* $0 \leqq \tilde{\lambda} < 1$, *then the problem* (RLPCC(λ)) *has a solution for* $0 \leqq \lambda < 1$.

Proof. Fix $\tilde{\lambda}$, $0 \leqq \tilde{\lambda} < 1$. Since (MRLPCC($\lambda$)) has exactly the same structure as (RLPCC(λ)), all the earlier discussion and results apply to the homotopy map $\tilde{\rho_a}$, which has the same (unique) zero at $\lambda = 0$ as ρ_a. So assuming

$$\operatorname{rank} \nabla_{(x,y)}\tilde{\psi}(1,\cdot) = \operatorname{rank} \nabla_{(x,y)}\psi(\tilde{\lambda},\cdot) = m$$

over $\Omega_{\tilde{\lambda}}$ (with respect to (RLPCC)), and either u or v is bounded along $\gamma(\tilde{\lambda})$ for $0 \leqq \tilde{\lambda} < 1$, Theorem 1 applies to $\tilde{\rho_a}$. Then $\tilde{\rho_a}(1,\, \bar{x},\, \bar{y},\, \bar{u},\, \bar{v}) = 0$ means

$$\begin{bmatrix} \begin{pmatrix} c \\ d \end{pmatrix} - \begin{pmatrix} A^t \\ B^t \end{pmatrix} \bar{u} + \Big(\nabla_{(x,y)}\tilde{\psi}\big(1,\bar{x},\bar{y},q^0,g^0\big)\Big)^t \bar{v} \\ \tilde{\phi}\big(1,\bar{x},\bar{y},\bar{u},f^0,h^0\big) \\ \tilde{\psi}\big(1,\bar{x},\bar{y},q^0,g^0\big) \end{bmatrix} = 0,$$

which implies $A\bar{x} + B\bar{y} - f + (1-\tilde{\lambda})f^0 \geqq 0$ and $\psi\big(\tilde{\lambda},\bar{x},\bar{y},q^0,g^0\big) = 0$, so $(\bar{x},\bar{y})$ is a feasible point for (RLPCC($\tilde{\lambda}$)). Therefore (RLPCC($\tilde{\lambda}$)) has a solution, since its feasible set is nonempty and compact. ∎

Observe that $\tilde{\rho_a}(1,\, \tilde{x},\, \tilde{y},\, \tilde{u},\, \tilde{v}) = 0$ implies $\tilde{\psi}(1,\, \tilde{x},\, \tilde{y},\, \cdot) = \psi(\tilde{\lambda},\, \tilde{x},\, \tilde{y},\, \cdot) = 0$ and $\tilde{\phi}(1,\, \tilde{x},\, \tilde{y},\, \cdot) = \phi(\tilde{\lambda},\, \tilde{x},\, \tilde{y},\, \cdot) = 0$, and the latter implies (because ϕ involves a positively oriented NCP function) $A\tilde{x} + B\tilde{y} - f + (1-\tilde{\lambda})f^0 > 0$. Hence the proof of Theorem C1 gives the following corollary.

COROLLARY C1. *Under the assumptions for Theorem* C1, *there exists a point* $(\tilde{x},\tilde{y}) \in \operatorname{int} \Omega_{\tilde{\lambda}}$ *satisfying* $\psi(\tilde{\lambda},\, \tilde{x},\, \tilde{y},\, \cdot) = 0$ *for* $0 \leqq \tilde{\lambda} < 1$, *i.e.,* (RLPCC($\tilde{\lambda}$)) *has a strictly feasible point for* $0 \leqq \tilde{\lambda} < 1$.

THEOREM C2. *Assume* $\Omega_1 \neq \emptyset$ *is bounded. If rank* $\nabla_{(x,y)}\psi(\lambda,\cdot) = m$ *over* Ω_λ *for* $0 \leqq \lambda < 1$, *and either* u *or* v *is bounded along* $\gamma(\tilde{\lambda})$ *for* $0 \leqq \tilde{\lambda} < 1$, *then for almost all*

$$a = \big(x^0, y^0, f^0, g^0, h^0\big) \in E^n \times E^m_{++} \times E^k_{++} \times E^m_{++} \times E^k_{++}$$

satisfying $Ax^0 + By^0 - f + f^0 > 0$, *and for each* λ *in* $0 \leqq \lambda < 1$, *the Kuhn–Tucker constraint qualification* (7.3.3 *of* [33]) *is satisfied at some local solution* $(\bar{x},\bar{y})$ *of* (RLPCC(λ)).

Proof. Fix $\tilde{\lambda}$, $0 \leqq \tilde{\lambda} < 1$. By Theorem C1, the problem (RLPCC($\tilde{\lambda}$)) has a solution, and by Corollary C1 there is a strictly feasible point $(\tilde{x}, \tilde{y}) \in \text{int}\ \Omega_{\tilde{\lambda}}$. From the rank assumption on $\nabla_{(x,y)}\psi$, the set $\mathcal{M} = \{(x, y) \mid \psi(\tilde{\lambda}, x, y, \cdot) = 0\}$ is an n-dimensional manifold in a neighborhood of $\Omega_{\tilde{\lambda}}$. Let Γ be the (closed) connected component of $\mathcal{M} \cap \Omega_{\tilde{\lambda}}$ containing $(\tilde{x}, \tilde{y}) \in \text{int}\ \Omega_{\tilde{\lambda}}$, and let $(\bar{x}, \bar{y}) \in \Gamma$ be a local solution of (RLPCC($\tilde{\lambda}$)), which exists since Γ is compact. Write the linear constraints as

$$g(x, y) = f - (1 - \tilde{\lambda})f^0 - Ax - By \leqq 0$$

and let $I = \{i \mid g_i(\bar{x}, \bar{y}) = 0\}$. $\mathcal{M}$ has the n-dimensional tangent space $\mathcal{T} = \{z \mid \nabla_{(x,y)}\psi(\tilde{\lambda},\ \bar{x},\ \bar{y}, \cdot)z = 0\}$ at $(\bar{x}, \bar{y})$. With respect to $\Omega_{\tilde{\lambda}}$, the feasible directions at $(\bar{x}, \bar{y})$ are $\mathcal{F} = \left\{z \;\middle|\; \nabla g_I(\bar{x}, \bar{y})z \leqq 0\right\}$, and the directions of interest for the constraint qualification are $\mathcal{F} \cap \mathcal{T}$.

Let $0 \neq d \in \mathcal{F} \cap \mathcal{T}$ be a feasible direction. Since ψ is C^4, by the implicit function theorem, the connected set Γ is also path connected: there is a differentiable vector function $e(\tau)$ starting at $e(0) = (\bar{x}, \bar{y})$, lying in $\Gamma \subset \mathcal{M} \cap \Omega_{\tilde{\lambda}}$, and ending at $e(1) = (\tilde{x}, \tilde{y})$. Furthermore, since $\Omega_{\tilde{\lambda}}$ is convex, int $\Omega_{\tilde{\lambda}} \neq \emptyset$, and the curve $e([0, 1])$ lies entirely in $\Gamma \subset \mathcal{M} \cap \Omega_{\tilde{\lambda}}$, $e(\tau)$ can be chosen such that $e'(0) = \alpha d$ for some $\alpha > 0$, which is the Kuhn–Tucker constraint qualification. ∎

Theorem C2 provides sufficient conditions such that, for each λ, there exists a local solution of (RLPCC(λ)) that satisfies the Kuhn–Tucker constraint qualification, but it provides no guarantee that the constraint qualification is satisfied at a particular local solution. The following corollary provides sufficient conditions such that, for almost all λ, the constraint qualification will be satisfied at *any* local solution, provided that no more than n of the linear constraints are active at that solution.

COROLLARY C2. *Under the assumptions of Theorem* C2*, for almost all* $\lambda \in [0, 1)$*, any solution* $(\bar{x}, \bar{y})$ *of* (RLPCC(λ)) *will satisfy the Kuhn–Tucker constraint qualification, provided no more than n of the linear constraints* $f - (1-\lambda)f^0 - Ax - By \leqq 0$ *are active.*

Proof. Let λ be fixed, and let $(\bar{x}, \bar{y})$ be a local solution of (RLPCC(λ)) with no more than n of the linear constraints active. Let I denote the indices of the active linear constraints, that is, $I = \{i \mid (f - (1 - \lambda)f^0 - A\bar{x} - B\bar{y})_i = 0\}$. By assumption, $|I| \leqq n$.

Define

$$g\big(\lambda, x, y, f^0\big) = f - (1 - \lambda)f^0 - Ax - By, \quad F\big(\lambda, x, y, f^0, g^0\big) = \begin{pmatrix} g_I\big(\lambda, x, y, f^0\big) \\ \psi\big(\lambda, x, y, q^0, g^0\big) \end{pmatrix}.$$

Note that F is transversal to zero since the columns of $\nabla F\big(\lambda,\ \bar{x},\ \bar{y},\ f^0,\ g^0\big)$ corresponding to f_I^0 and g^0 form a multiple of the identity matrix. Therefore, by the parametrized Sard's theorem, for almost all $b = \big(\lambda,\ f^0,\ g^0\big)$, $F_b(x, y) = F\big(\lambda,\ x,\ y,\ f^0,\ g^0\big)$ is also transversal to zero. Thus

$$\text{rank}\ \nabla_{(x,y)}F_b(\bar{x}, \bar{y}) = \text{rank}\ \begin{pmatrix} \nabla_{(x,y)}g_I\big(\lambda, \bar{x}, \bar{y}, f^0\big) \\ \nabla_{(x,y)}\psi\big(\lambda, \bar{x}, \bar{y}, q^0, g^0\big) \end{pmatrix} = m + |I| \leqq m + n.$$

Since the columns of $\nabla F_b(\bar{x}, \bar{y})$ span $E^{m+|I|}$, there exists $z \in E^{m+n}$ such that

$$rl\nabla_{(x,y)}g_I\big(\lambda, \bar{x}, \bar{y}, f^0\big)z < 0,$$
$$\nabla_{(x,y)}\psi\big(\lambda, \bar{x}, \bar{y}, q^0, g^0\big)z = 0,$$

or equivalently, $z \in \mathcal{T} \cap \operatorname{int} \mathcal{F}$, where $\mathcal{T} = \{z \mid \nabla_{(x,y)}\psi(\lambda, \bar{x}, \bar{y}, \cdot)z = 0\}$ and $\mathcal{F} = \{z \mid \nabla g_I(\bar{x}, \bar{y})z \leqq 0\}$. The existence of such a (strictly) feasible direction z implies that $\mathcal{M} \cap \operatorname{int} \Omega_\lambda \cap B\big((\bar{x}, \bar{y}), \delta\big) \neq \emptyset$ for every open ball $B\big((\bar{x}, \bar{y})$ of radius $\delta > 0$ centered at $(\bar{x}, \bar{y})$, i.e., there exist strictly feasible points $(\tilde{x}, \tilde{y})$ arbitrarily close to $(\bar{x}, \bar{y})$. Thus for δ sufficiently small, $(\tilde{x}, \tilde{y})$ and $(\bar{x}, \bar{y})$ lie in the same (closed) connected component Γ of $\mathcal{M} \cap \Omega_\lambda$, and the proof of Theorem C2 applies for the satisfaction of the Kuhn–Tucker constraint qualification at $(\bar{x}, \bar{y})$, except now the conclusion is for almost all λ rather than each λ, $0 \leqq \lambda < 1$. ∎

The restriction on rank $\nabla_{(x,y)}\psi$ in Theorems 1, C1, and C2 is not so severe as it might seem, since the rank assumption holds generically (with probability one). To see this, fix q^0 defined by $\psi\big(0,\, x^0,\, y^0,\, q^0,\, g^0\big) = 0$, let $0 \leqq \lambda < 1$, $b = (\lambda, g^0)$, and define

$$\psi_b(x, y) = \psi\big(\lambda, x, y, q^0, g^0\big).$$

Observe that, with respect to the variables $\big(\lambda,\, x,\, y,\, g^0\big)$, ψ is transversal to zero. Therefore, by the parametrized Sard's theorem [9], for almost all $g^0 > 0$ (and q^0 determined by g^0) and almost all λ, $0 \leqq \lambda < 1$, $\psi_b : E^{n+m} \to E^m$ is also transversal to zero. This proves the following lemma.

LEMMA 3. *Let q^0 be defined by $\psi\big(0,\, x^0,\, y^0,\, q^0,\, g^0\big) = 0$. Then for almost all $g^0 > 0$ and almost all λ, $0 \leqq \lambda < 1$,*

$$\operatorname{rank} \nabla_{(x,y)}\psi_b(x, y) = \operatorname{rank} \nabla_{(x,y)}\psi\big(\lambda, x, y, q^0, g^0\big) = m$$

on the set $\psi_b^{-1}(0)$.

As $\lambda \to 1$ along the zero curve γ of ρ_a, $\psi_b = \psi \approx 0$, and Lemma 3 says that rank $\nabla_{(x,y)}\psi = m$ along γ as $\lambda \to 1$, almost surely. Thus the assumption of Lemma 2 (where Ω_1 could be replaced by the projection of γ onto Ω_1) is rather mild and is tantamount to assuming a nondegenerate complementarity solution $(\bar{x}, \bar{y})$ at $\lambda = 1$.

LEMMA 4. *Let $\big\{\big(\lambda_k, x^k, y^k, u^k, v^k\big)\big\}_{k=1}^{\infty} \subset \gamma$. Then*

$$\nabla_{(x,y)}\psi\big(\lambda_k, x^k, y^k, q^0, g^0\big) = \big(\Sigma^k N, \quad \Sigma^k M + \Delta^k\big),$$

where Σ^k and Δ^k are diagonal matrices with entries

$$\Sigma^k_{ii} = \begin{cases} -5(s_i^k - y_i^k)^4 + 5(s_i^k)^4 & \text{if } s_i^k \geqq y_i^k, \\ 5(s_i^k - y_i^k)^4 + 5(s_i^k)^4 & \text{if } s_i^k < y_i^k, \end{cases}$$

$$\Delta^k_{ii} = \begin{cases} 5(s_i^k - y_i^k)^4 + 5(y_i^k)^4 & \text{if } s_i^k \geqq y_i^k, \\ -5(s_i^k - y_i^k)^4 + 5(y_i^k)^4 & \text{if } s_i^k < y_i^k, \end{cases}$$

where $s_i^k = \big(Nx^k + My^k + q + (1-\lambda_k)q^0\big)_i$. Additionally, if $\psi(\lambda, x, y, q^0, g^0) = o(1-\lambda)$ along γ as $\lambda \to 1$, then for sufficiently large k, $s^k > 0$, $y^k > 0$, and the diagonal entries of Σ^k and Δ^k are positive.

Proof. The formula for $\nabla_{(x,y)}\psi\big(\lambda_k,\, x^k,\, y^k,\, q^0,\, g^0\big)$ follows from the chain rule and the definition of $\hat{\psi}^{(5)}$. If $\psi(\lambda,\, x,\, y,\, q^0,\, g^0) = o(1-\lambda)$ along γ as $\lambda \to 1$, then for k sufficiently large, $\hat{\psi}^{(5)}(s_i^k, y_i^k) = \psi_i\big(\lambda_k,\, x^k,\, y^k,\, q^0,\, g^0\big) + (1-\lambda_k)g_i^0 > \frac{(1-\lambda_k)}{2} g_i^0 > 0$ (since $\psi_i\big(\lambda_k,\, x^k,\, y^k,\, q^0,\, g^0\big) = o(1-\lambda_k)$). This implies that $s_i^k > 0$ and $y_i^k > 0$ for sufficiently large k, which ensures that $\Sigma^k_{ii} > 0$ and $\Delta^k_{ii} > 0$. ∎

LEMMA 5. *Suppose that* $\psi(\lambda, x, y, q^0, g^0) = o(1-\lambda)$ *along* γ *as* $\lambda \to 1$ *and the limit point* $(1, \bar{x}, \bar{y})$ *is a nondegenerate solution. Then* $\nabla_{(x,y)}\psi$ *has full rank along* γ *for* $0 \ll \tilde{\lambda} < \lambda \leqq 1$ *if*

$$\operatorname{rank}\begin{pmatrix} N_{J\cdot} & M_{J\cdot} \\ 0 & I_{K\cdot} \end{pmatrix} = m$$

for all partitions of $\{1, \ldots, m\}$ *into disjoint subsets* J, K.

Proof. Let $\{(\lambda_k, x^k, y^k, u^k, v^k)\}_{k=1}^{\infty} \subset \gamma$ and (λ_k, x^k, y^k) converge to $(1, \bar{x}, \bar{y})$. Define $s_i^k = \big(Nx^k + My^k + q + (1-\lambda_k)q^0\big)_i$. By continuity, $\{s_i^k\}$ converges to $\bar{s}_i = \big(N\bar{x} + M\bar{y} + q\big)_i$. By Lemma 4,

$$\nabla_{(x,y)}\psi\big(\lambda_k, x^k, y^k, q^0, g^0\big) = \big(\Sigma^k N, \quad \Sigma^k M + \Delta^k\big),$$

where Σ^k and Δ^k are the diagonal matrices defined in Lemma 4. Suppose that $\bar{y}_i = 0$. By nondegeneracy, for k sufficiently large, $s_i^k > \bar{s}_i/2 > y_i^k > 0$, where the last inequality is from Lemma 4. Thus, $\Sigma_{ii}^k = -5(s_i^k - y_i^k)^4 + 5(s_i^k)^4 \to 0$, and $\Delta_{ii}^k = 5(s_i^k - y_i^k)^4 + 5(y_i^k)^4 \to 5\bar{s}_i^4 > 0$.

Similarly, if $\bar{y}_i > 0$, then $\Sigma_{ii}^k = 5(s_i^k - y_i^k)^4 + 5(s_i^k)^4 \to 5\bar{y}_i^4 > 0$ and $\Delta_{ii}^k = -5(s_i^k - y_i^k)^4 + 5(y_i^k)^4 \to 0$. Let Υ^k be the diagonal matrix with entries

$$\Upsilon_{ii}^k = \begin{cases} 1/\big(5(s_i^k)^4\big) & \text{if } \bar{y}_i = 0, \\ 1/\big(5(y_i^k)^4\big) & \text{if } \bar{y}_i > 0. \end{cases}$$

Then

$$\lim_{k\to\infty} \Upsilon^k \big(\Sigma^k N, \quad \Sigma^k M + \Delta^k\big) = P\begin{pmatrix} N_{J\cdot} & M_{J\cdot} \\ 0 & I_{K\cdot} \end{pmatrix},$$

where P is a permutation matrix, $J := \{i \mid \bar{y}_i > 0\}$, $K := \{i \mid \bar{y}_i = 0\}$. It follows that for k sufficiently large,

$$\operatorname{rank} \nabla_{(x,y)}\psi\big(\lambda_k, x^k, y^k, q^0, g^0\big) = \operatorname{rank}\begin{pmatrix} N_{J\cdot} & M_{J\cdot} \\ 0 & I_{K\cdot} \end{pmatrix}.$$

The result follows immediately. ◻

LEMMA 6. *Suppose* (A1) *holds for all nonnegative diagonal matrices* Σ, $\Delta \in E^{m\times m}$ *satisfying* $\Sigma_{ii} + \Delta_{ii} > 0$ *for all* i *and that* $\psi(\lambda, x, y, q^0, g^0) = o(1-\lambda)$ *along* γ *as* $\lambda \to 1$. *Suppose that* $\{(\lambda_k, x^k, y^k, u^k, v^k)\} \subset \gamma$ *is such that* $(\lambda_k, x^k, y^k, u^k) \to (1, \bar{x}, \bar{y}, \bar{u})$. *Let* D *denote the degenerate indices of the complementarity constraints at* $(\bar{x}, \bar{y})$. *If* $v_D^k \leqq 0$ *for all* k *sufficiently large, then* $(\bar{x}, \bar{y})$ *is a local optimum point of* (LPCC).

Proof. Let J and K denote the nondegenerate indices with $\bar{y}_J > 0$ and $\bar{y}_K = 0$. By Lemma 4,

$$\nabla_{(x,y)}\psi\big(\lambda_k, x^k, y^k, q^0, g^0\big) = \big(\Sigma^k N, \quad \Sigma^k M + \Delta^k\big),$$

where Σ^k and Δ^k are diagonal matrices with positive diagonal entries for sufficiently large k. Let $\Upsilon^k = (\Sigma^k + \Delta^k)^{-1}$. Then $\Upsilon^k \nabla\psi(\lambda_k, x^k, y^k, q^0, g^0) = \Lambda^k(N, M) + (I - \Lambda^k)(0, I)$, where Λ^k is a diagonal matrix with diagonal entries between zero and one. Let

$$g^k = -\frac{1-\lambda_k}{\lambda_k}\begin{pmatrix} x^k - x^0 \\ y^k - y^0 \end{pmatrix} - \begin{pmatrix} c \\ d \end{pmatrix} + \begin{pmatrix} A^t \\ B^t \end{pmatrix} u^k,$$

and note that $\{g^k\}$ is bounded. By passing to a subsequence, it may be assumed that both $g^k \to \bar{g}$ and $\Lambda^k_{DD} \to \bar{\Lambda}$. By the definition of ρ_a,

$$g^k = \left(\nabla\psi\left(\lambda_k, x^k, y^k, q^0, g^0\right)\right)^t v^k = \begin{pmatrix} N^t\Lambda^k \\ M^t\Lambda^k + I - \Lambda^k \end{pmatrix} w^k = G^k w^k,$$

where $w^k = \left(\Upsilon^k\right)^{-1} v^k$. The matrix G^k converges to

$$\begin{pmatrix} (N_{J\cdot})^t & 0 & (N_{D\cdot})^t\bar{\Lambda} \\ (M_{J\cdot})^t & I_{\cdot K} & \left((M_{D\cdot})^t\bar{\Lambda} + I_{\cdot D} - I_{\cdot D}\bar{\Lambda}\right) \end{pmatrix} P,$$

where P is a permutation matrix and $\bar{\Lambda}$ is a $|D|\times|D|$ diagonal matrix with nonnegative entries less than or equal to one. By assumption (A1), this matrix has full column rank, so for k sufficiently large, the matrix G^k also has full column rank and the smallest singular value is bounded away from zero. Since $\{g^k\}$ is bounded, it follows that $\{w^k\}$ is bounded with an accumulation point $\bar{w}$. Let $\xi = \bar{\Lambda}\bar{w}_D$, and $\zeta = (I - \bar{\Lambda})\bar{w}_D$. Then equating the two expressions for $\bar{g}$ gives

$$\begin{pmatrix} c \\ d \end{pmatrix} - \begin{pmatrix} A^t \\ B^t \end{pmatrix} \bar{u} + \begin{pmatrix} N^t_{\cdot J} & 0 & N^t_{\cdot D} & 0 \\ M^t_{\cdot J} & I_{\cdot K} & M^t_{\cdot D} & I_{\cdot D} \end{pmatrix} \begin{pmatrix} \bar{w}_J \\ \bar{w}_K \\ \xi \\ \zeta \end{pmatrix} = 0.$$

Observe also that $\xi \leqq 0$, $\zeta \leqq 0$. Therefore, $(\bar{x}, \bar{y}, \bar{u}, \bar{w})$ is a Kuhn–Tucker point (so $(\bar{x}, \bar{y})$ is an optimal solution) for the following linear program:

$$\text{(RLP)} \qquad \begin{aligned} \min \quad & c^t x + d^t y \\ \text{subject to} \quad & Ax + By \geqq f, \\ & q_J + N_{J\cdot}x + M_{J\cdot}y = 0, \\ & y_K = 0, \\ & q_D + N_{D\cdot}x + M_{D\cdot}y \geqq 0, \\ & y_D \geqq 0. \end{aligned}$$

Since $\phi(1, \bar{x}, \bar{y}, \bar{u}, \cdot) = 0$ and $\psi(1, \bar{x}, \bar{y}, \cdot) = 0$, the point $(\bar{x}, \bar{y})$ is feasible for (LPCC); so, within a neighborhood of $(\bar{x}, \bar{y})$, the feasible set of (LPCC) is contained in the feasible region of (RLP). Thus, $(\bar{x}, \bar{y})$ is a locally optimal solution for (LPCC). ∎

THEOREM C3. *Suppose that* $(\bar{x}, \bar{y})$ *is feasible for* (LPCC) *and is nondegenerate with respect to the complementarity constraints. Then* $(\bar{x}, \bar{y})$ *is locally optimal for* (LPCC) *if and only if* $(\bar{x}, \bar{y}, \bar{u}, \bar{v})$*, for some* $\bar{u} \geqq 0$ *and* $\bar{v}$*, is a Kuhn–Tucker point for* (RLPCC) *at* $\lambda = 1$*.*

Proof. Let $J = \{i \mid \bar{y}_i > 0\}$ and $K = \{i \mid \bar{y}_i = 0\}$. Since $(\bar{x}, \bar{y})$ satisfies the complementarity constraints and is nondegenerate, $q_J + N_{J\cdot}\bar{x} + M_{J\cdot}\bar{y} = 0$ and $q_K + N_{K\cdot}\bar{x} + M_{K\cdot}\bar{y} > 0$. Thus, in a neighborhood of $(\bar{x}, \bar{y})$, the feasible region of (LPCC) coincides with the feasible region for the following linear program:

$$\begin{aligned} \min \quad & c^t x + d^t y \\ \text{subject to} \quad & Ax + By \geqq f, \\ & q_J + N_{J\cdot}x + M_{J\cdot}y = 0, \\ & y_K = 0. \end{aligned}$$

The Kuhn–Tucker conditions for this linear program are

$$\begin{gathered}
\begin{pmatrix} c \\ d \end{pmatrix} - \begin{pmatrix} A^t \\ B^t \end{pmatrix} u + \begin{pmatrix} N_{\cdot J}^t & 0 \\ M_{\cdot J}^t & I_{\cdot K} \end{pmatrix} w = 0, \\
Ax + By - f \geqq 0, \\
u \geqq 0, \\
u^t(Ax + By - f) = 0, \\
q_J + N_{J\cdot} x + M_{J\cdot} y = 0, \\
y_K = 0.
\end{gathered}$$

By arguments similar to those in the proof of Lemma 5,

$$\nabla_{(x,y)} \psi\big(1, \bar{x}, \bar{y}, q^0, g^0\big) = \big(\bar{\Sigma} N, \quad \bar{\Sigma} M + \bar{\Delta}\big),$$

where $\bar{\Sigma}$ and $\bar{\Delta}$ are diagonal matrices satisfying $\bar{\Sigma}_{ii} > 0$, $\bar{\Delta}_{ii} = 0$ for $i \in J$, and $\bar{\Sigma}_{ii} = 0$, $\bar{\Delta}_{ii} > 0$ for $i \in K$. Thus, for a permutation matrix P,

$$\nabla_{(x,y)} \psi\big(1, \bar{x}, \bar{y}, q^0, g^0\big)^t = \begin{pmatrix} N_{\cdot J}^t & 0 \\ M_{\cdot J}^t & I_{\cdot K} \end{pmatrix} \begin{pmatrix} \bar{\Sigma}_{JJ} & 0 \\ 0 & \bar{\Delta}_{KK} \end{pmatrix} P.$$

It follows that $(\bar{x}, \bar{y}, \bar{u}, \bar{w})$ satisfies the Kuhn–Tucker conditions for the linear program if and only if $(\bar{x}, \bar{y}, \bar{u}, \bar{v})$ satisfies the Kuhn–Tucker conditions for (RLPCC) at $\lambda = 1$ with $\bar{v} = P^t\big(\operatorname{diag}(\bar{\Sigma}_{JJ}, \bar{\Delta}_{KK})\big)^{-1} \bar{w}$. Therefore, $(\bar{x}, \bar{y})$ is a local optimum point of (LPCC) if and only if $(\bar{x}, \bar{y}, \bar{u}, \bar{v})$ is a Kuhn–Tucker point for (RLPCC) at $\lambda = 1$. ∎

5. Numerical results. Two small, difficult examples (one degenerate, the other a manifold of solutions) are used here to illustrate several aspects of homotopy algorithms. An important observation is that probability-one homotopy methods, similar to Monte Carlo methods, are rather insensitive to dimension but more sensitive to the nonlinearity of the problem. (This has been widely observed in structural mechanics [58], analogue circuit design [35], discretizations of boundary value problems [59, 60], and nonlinear complementarity problems [1], to mention just a few areas.) The point is that for problem classes of similar nonlinearity (like LPCCs), the nature of the homotopy curves (arc length, number of turning points, maximum curvature, sensitivity to starting point, etc.) will be qualitatively similar independent of the dimension. Thus nothing qualitatively new would be learned from solving a few large LPCCs beyond the small difficult ones here. The tanh term is used for technical reasons on large scale realistic problems and is not pertinent here, so $\tanh(60\lambda/(1-\lambda))$ is replaced by just λ. Just as the standard homotopy map

$$\lambda F(x) + (1-\lambda)(x-a),$$

while theoretically adequate for a large class of nonlinear systems $F(x) = 0$, is often not a good choice in practice [55], so the (theoretically correct) homotopy map ρ_a used here may not be a computationally efficient choice for LPCCs. A homotopy map ρ_a more intimately connected to the structure of an LPCC is desirable, but the development of such (as in [54] or [55], e.g.), and accompanying theory, is a topic for future work.

The subroutine FIXPNF in the package HOMPACK90 [61] is used here for the homotopy zero curve tracking. The first example is given by $k = 11$, $m = 3$, $n = 2$, $c = (1, 0)^t$, $d = (2, 0, -1)^t$,

$$A = \begin{pmatrix} 1 & 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}^t,$$
$$B = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 \end{pmatrix}^t,$$
$$f = (5, 0, 0, -7, -0.5, -0.5, -0.5, -0.5, 0, 0, 0)^t,$$
$$N = \begin{pmatrix} 0 & 1 \\ 1 & 0 \\ 1 & 1 \end{pmatrix}, \qquad M = \begin{pmatrix} 0 & 0 & -1 \\ 1 & 1 & 0 \\ 0 & -1 & 0 \end{pmatrix}, \qquad q = (1, 0, 2)^t.$$

This problem has a ray $\bar{x} = (x_1, 0)$, $\bar{y} = 0$ of nondegenerate solutions with $7 \geqq x_1 \geqq 5$, and the homotopy curve converges to some point along this ray. Note also that assumption (A1) is satisfied. The initial (arbitrarily chosen, infeasible) data is given by

$$x^0 = (1 \cdot 10^{-3}, 1 \cdot 10^{-3})^t, \qquad y^0 = (1 \cdot 10^{-3}, 1 \cdot 10^{-3}, 1 \cdot 10^{-3})^t,$$
$$f^0 = (1.000, 1.001, 1.001, 1.001, 1.001, 1.001, 1.001, 1.001, 1.001, 1.001, 1.001)^t,$$
$$g^0 = (0.3, 0.25, 6.25 \cdot 10^{-2})^t, \quad h^0 = (0.1, 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, 5.0, 5.5)^t,$$
$$u^0 = \left(5.77, 1.34 \cdot 10^{-2}, 0.80, 4.21 \cdot 10^{-2}, 0.5, 0.64, 0.79, 0.93, 1.54, 1.62, 1.69\right)^t,$$
$$v^0 = \left(-3.65 \cdot 10^{-3}, 3.36 \cdot 10^{-3}, 1.42 \cdot 10^{-2}\right)^t, \qquad q^0 = (1.78, -1.34, -3.76)^t,$$

where u^0, v^0, and q^0 are computed from the other initial data as described earlier.

Figure 5.1 shows the homotopy zero curve γ for this initial data. The dashed line in Figure 5.1 shows γ for the starting point $x^0 = (4.0,\ 1 \cdot 10^{-3})^t$, $y^0 = (1 \cdot 10^{-3},\ 1 \cdot 10^{-3},\ 1 \cdot 10^{-3})^t$, with f^0, g^0, h^0 unchanged, and u^0, v^0, q^0 computed accordingly. The shape of γ is very sensitive to the initial data, due to the nature of ρ_a, as shown in Figure 5.2, which corresponds to starting point $x^0 = (1 \cdot 10^{-3},\ 1 \cdot 10^{-3})^t$, $y^0 = (1 \cdot 10^{-3},\ 0.3,\ 1 \cdot 10^{-3})^t$, with everything else as before. γ remains bounded and will eventually reach a solution at $\lambda = 1$ but has a long arc length and many sharp turns. The functions $\hat{\psi}^{(k)}$, especially for $k = 5$, suffer from numerical truncation error for large arguments, making curve tracking difficult.

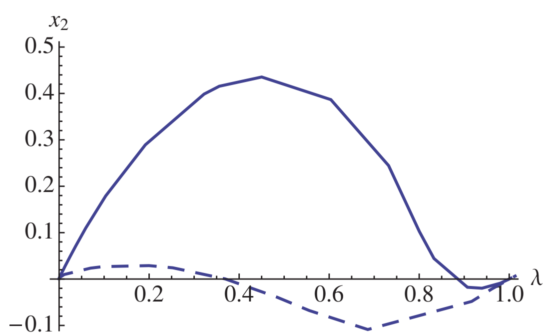

FIG. 5.1. *Plot of x_2 along the homotopy zero curve γ (solid line). The dashed line shows γ for a different starting point (x^0, y^0).*

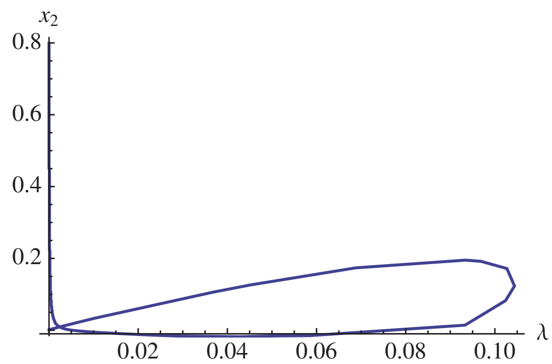

FIG. 5.2. *Plot of x_2 along the homotopy zero curve γ for a third starting point (x^0, y^0), showing the sensitivity of γ to the parameter vector a. γ eventually reaches the solution, after many sharp turns and a long arc length.*

The second example is given by $k = 6$, $m = 2$, $n = 2$, $c = (-2, 1)^t$, $d = (-1, 0)^t$,

$$A = \begin{pmatrix} -2 & -2 & -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}^t, \qquad B = \begin{pmatrix} 1 & -1 & -2 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}^t,$$

$$f = (-4, -6, -6, -2, 0, 0)^t,$$

$$N = \begin{pmatrix} 3 & 0 \\ 3 & 1 \end{pmatrix}, \qquad M = \begin{pmatrix} -2 & 0 \\ 6 & 0 \end{pmatrix}, \qquad q = (2, -14)^t.$$

This problem has the degenerate solution $\bar{x} = (2/3, 0)$, $\bar{y} = (2, 0)$ with value $-10/3$. Because of this degeneracy, regardless of the choice of a, the equation $\rho_a(1, \bar{x}, \bar{y}, u, v) = 0$ has no solution for u, v. Therefore no homotopy zero curve γ can reach this degenerate solution at $\lambda = 1$.

Consider the initial data given by

$$x^0 = (0.5, 1 \cdot 10^{-3})^t, \qquad y^0 = (1.5, 1 \cdot 10^{-3})^t,$$

$$f^0 = \left(1 \cdot 10^{-3}, 1 \cdot 10^{-3}, 1 \cdot 10^{-3}, 1 \cdot 10^{-3}, 1 \cdot 10^{-3}, 1 \cdot 10^{-3}\right)^t,$$

$$g^0 = (0.3, 0.25)^t, \quad h^0 = (0.1, 1.0, 1.5, 2.0, 2.5, 3.0)^t,$$

$$u^0 = \left(1.646 \cdot 10^{-3}, 2.741 \cdot 10^{-2}, 8.260 \cdot 10^{-2}, 1.357, 1.499, 22.362\right)^t,$$

$$v^0 = \left(3.538 \cdot 10^{-2}, 5.615 \cdot 10^{-3}\right)^t, \qquad q^0 = (-0.488, 6.159)^t,$$

where u^0, v^0, and q^0 are computed from the other initial data as described earlier.

For this data, Figure 5.3 (solid line) shows the homotopy zero curve γ, which approaches the hyperplane $\lambda = 1$ asymptotically without ever reaching a Kuhn–Tucker point. Along γ, $v_2 \to +\infty$ and (x, y) accumulates at the degenerate solution $(\bar{x}, \bar{y})$, but this happens very slowly. The figure shows y_2 accumulating at $\bar{y}_2 = 0$ as $\lambda \uparrow 1$. The dashed line in Figure 5.3 shows γ for the initial data $x^0 = (1 \cdot 10^{-3}, 1 \cdot 10^{-3})^t$, $y^0 = (1 \cdot 10^{-3}, 1 \cdot 10^{-3})^t$ reaching a nondegenerate local solution $\tilde{x} = (2, 8)^t$, $\tilde{y} = (0, 0.781)^t$ with value 4. As before, f^0, g^0, h^0 are unchanged, with u^0, v^0, q^0 computed accordingly. Figure 5.4 again shows the sensitivity of γ to the parameter vector a, with the curve shown corresponding to $x^0 = (1 \cdot 10^{-3}, 1 \cdot 10^{-3})^t$, $y^0 = (0.2, 1 \cdot 10^{-3})^t$.

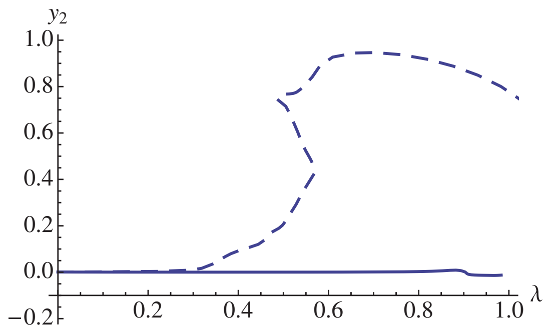

FIG. 5.3. *Plot of y_2 along the homotopy zero curve γ (solid line). The dashed line shows γ for a different starting point (x^0, y^0). Note that γ is C^2, but the plotted projections may have cusps.*

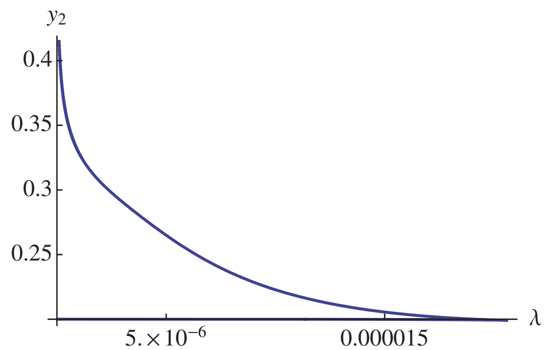

FIG. 5.4. *Plot of x_2 along the homotopy zero curve γ for a third starting point (x^0, y^0), showing the sensitivity of γ to the parameter vector a. γ eventually reaches a solution, after many sharp turns and a long arc length.*

6. Conclusions. A reformulation of (LPCC) with a simpler complementarity constraint is

$$\text{(LPCCw)} \qquad \begin{array}{rl} \min & c^t x + d^t y \\ \text{subject to} & Ax + By \geqq f, \\ & w - (q + Nx + My) = 0, \\ & 0 \leqq y \perp w \geqq 0. \end{array}$$

Using this simpler form (LPCCw) of an LPCC would have simplified all the proofs but at the expense of losing the explicit dependence of the results on the properties of the matrices M and N. Since the goal of applicable homotopy theory is to develop theory based on verifiable assumptions stated directly in terms of the quantities naturally arising in applications, a conscious decision was made to maintain the explicit dependence on M and N.

An important issue not explored here (beyond Lemma 6) is the extent to which the nondegeneracy assumptions can be relaxed. Many of the cited references address degeneracy, and it seems likely that the results here can be extended to cover some form of degeneracy. By construction, the probability-one homotopy algorithms involve nondegenerate and well-conditioned problems until $\lambda \approx 1$, and in practice the homotopy zero curve γ reaches degenerate solutions despite a lack of supporting theory. A careful examination of degeneracy should be a high priority for future work.

Based on the convergence theory for probability-one homotopy algorithms applied to nonlinear programs, the present results for LPCCs should be extensible to MPECs with pseudoconvex objective function and quasi-convex inequality constraints, as well as certain classes of general MPECs. Working out the details of these extensions is another fruitful avenue for future work.

Finally, as mentioned in the previous section, it is desirable that a homotopy map ρ_a, rather than being based on a direct relaxation of the Kuhn–Tucker conditions, be based on the structure of an LPCC; such a map ρ_a should have less meandering zero curves, if experience with many other applications [53] is any guide. The structure of such a "tailored" map would have the form $\lambda F(\lambda, x)+(1-\lambda)G(\lambda, x)$, where $F(1, x) = 0$ corresponds to the given LPCC, and $G(0, x) = 0$ is an LPCC similar to the given one, but with a unique, easily found solution. The technical challenge is to guarantee a solution for all $0 < \lambda < 1$ and to ensure the zero curves are bounded.

In summary, this paper showed that under existence and nondegeneracy assumptions, probability-one homotopies can be used to solve LPCCs and that the homotopy reformulation has certain advantages (e.g., satisfies a constraint qualification).

REFERENCES

[1] K. Ahuja, L. T. Watson, and S. C. Billups, *Probability-one homotopy maps for mixed complementarity problems*, Comput. Optim. Appl., 41 (2008), pp. 363–375.

[2] E. L. Allgower and K. Georg, *Numerical Continuation Methods*, Springer-Verlag, Berlin, 1990.

[3] M. Anitescu, P. Tseng, and S. J. Wright, *Elastic-mode algorithms for mathematical programs with equilibrium constraints: global convergence and stationarity properties*, Math. Program., 110 (2007), pp. 337–371.

[4] S. C. Billups, *A homotopy-based algorithm for mixed complementarity problems*, SIAM J. Optim., 12 (2002), pp. 583–605.

[5] S. C. Billups, A. L. Speight, and L. T. Watson, *Nonmonotone path following methods for nonsmooth equations and complementarity problems*, in Applications and Algorithms of Complementarity, M. C. Ferris, O. L. Mangasarian, and J.-S. Pang, eds., Kluwer, Norwell, MA, 2001, pp. 19–42.

[6] S. C. Billups and L. T. Watson, *A probability-one homotopy algorithm for nonsmooth equations and mixed complementarity problems*, SIAM J. Optim., 12 (2002), pp. 606–626.

[7] S. Burer and D. Vandenbussche, *Globally solving box-constrained nonconvex quadratic programs with semidefinite-based finite branch-and-bound*, Comput. Optim. Appl., 43 (2009), pp. 181–195.

[8] M. Campêlo and S. Scheimberg, *A note on a modified simplex approach for solving bilevel linear programming problems*, European J. Oper. Res., 126 (2000), pp. 454–458.

[9] S. N. Chow, J. Mallet-Paret, and J. A. Yorke, *Finding zeros of maps: Homotopy methods that are constructive with probability one*, Math. Comp., 32 (1978), pp. 887–899.

[10] B. Colson, P. Marcotte, and G. Savard, *An overview of bilevel optimization*, Ann. Oper. Res., 153 (2007), pp. 235–256.

[11] S. Dempe, *Foundations of Bilevel Programming*, Kluwer, Dordrecht, Netherlands, 2002.

[12] S. Dempe and J. Dutta, *Is bilevel programming a special case of a mathematical program with complementarity constraints?*, Math. Program., 131 (2012), pp. 37–48.

[13] B. C. Eaves and R. Saigal, *Homotopies for computation of fixed points on unbounded regions*, Math. Program., 3 (1972), pp. 225–237.

[14] F. Facchinei, H. Jiang, and L. Qi, *A smoothing method for mathematical programs with equilibrium constraints*, Math. Program., 85 (1999), pp. 107–134.

[15] H. Fang, S. Leyffer, and T. S. Munson, *A pivoting algorithm for linear programs with complementarity constraints*, Optim. Methods Softw., 27 (2012), pp. 89–114.

[16] M. L. Flegel and C. Kanzow, *On the Guignard constraint qualification for mathematical programs with equilibrium constraints*, Optimization, 54 (2005), pp. 517–534.

[17] R. Fletcher and S. Leyffer, *Solving mathematical programs with complementarity constraints as nonlinear programs*, Optim. Methods Softw., 19 (2004), pp. 15–40.

[18] R. Fletcher, S. Leyffer, D. Ralph, and S. Scholtes, *Local convergence of SQP methods for mathematical programs with equilibrium constraints*, SIAM J. Optim., 17 (2006), pp. 259–286.

[19] M. Fukushima and P. Tseng, *An implementable active-set algorithm for computing a B-stationary point for a mathematical program with linear complementarity constraints (with erratum)*, SIAM J. Optim., 12 (2002), pp. 724–739.

[20] Y. Ge, L. T. Watson, E. G. Collins, Jr., and D. S. Bernstein, *Probability-one homotopy algorithms for full and reduced order H^2/H^∞ controller synthesis*, Optimal Control Appl. Methods, 17 (1996), pp. 187–208.

[21] P. Hansen, B. Jaumard, and G. Savard, *New branch-and-bound rules for linear bilevel programming*, SIAM J. Sci. Statist. Comput., 13 (1992), pp. 1194–1217.

[22] J. Hu, J. E. Mitchell, and J.-S. Pang, *An LPCC approach to nonconvex quadratic programs*, Math. Program. 133(1–2), (2008), pp. 243–277.

[23] J. Hu, J. E. Mitchell, and J.-S. Pang, *On linear programs with linear complementarity constraints*, J. Global Optim., 53 (2012), pp. 29–51.

[24] J. Hu, J. E. Mitchell, J.-S. Pang, K. P. Bennett, and G. Kunapuli, *On the global solution of linear programs with linear complementarity constraints*, SIAM J. Optim., 19 (2008), pp. 445–471.

[25] X. M. Hu and D. Ralph, *Convergence of a penalty method for mathematical programming with complementarity constraints*, J. Optim. Theory Appl., 123 (2004), pp. 365–390.

[26] H. Jiang and D. Ralph, *Smooth SQP methods for mathematical programs with nonlinear complementarity constraints*, SIAM J. Optim., 10 (2000), pp. 779–808.

[27] H. Jiang and D. Ralph, *Extension of quasi-Newton methods to mathematical programs with complementarity constraints*, Comput. Optim. Appl., 123 (2004), pp. 365–390.

[28] G. Kunapuli, J. Hu, K. P. Bennett, and J.-S. Pang, *Classification model selection via bilevel programming*, Optim. Methods Softw., 23 (2008), pp. 475–489.

[29] S. Leyffer, *Complementarity constraints as nonlinear equations: Theory and numerical experience*, in Optimization with Multivalued Mappings, S. Dempe and V. Kalashnikov, eds., Springer, Berlin, 2006, pp. 169–208.

[30] S. Leyffer, G. Lopéz-Calva, and J. Nocedal, *Interior point methods for mathematical programs with complementarity constraints*, SIAM J. Optim., 17 (2006), pp. 52–77.

[31] S. Leyffer and T. S. Munson, *A Globally Convergent Filter Method for MPECs*, preprint ANL/MCS-P1457-0907 Argonne National Laboratory, Argonne, IL, 2009.

[32] Z.-Q. Luo, J.-S. Pang, and D. Ralph, *Mathematical Programs with Equilibrium Constraints*, Cambridge University Press, Cambridge, UK, 1996.

[33] O. L. Mangasarian, *Nonlinear Programming*, McGraw-Hill, New York, 1969.

[34] O. L. Mangasarian, *Equivalence of the complementarity problem to a system of nonlinear equations*, SIAM J. Appl. Math., 31 (1976), pp. 89–92.

[35] R. C. Melville, Lj. Trajković, S.-C. Fang, and L. T. Watson, *Artificial parameter homotopy methods for the DC operating point problem*, IEEE Trans. Computer-Aided Design, 12 (1993), pp. 861–877.

[36] O. H. Merrill, *Applications and Extensions of an Algorithm that Computes Fixed Points of Certain Non-Empty, Convex, Upper Semi-continuous Point to Set Mappings*, Tech. report 71-7, Department of Industrial Engineering, University of Michigan, 1971.

[37] A.-V. de Miguel, M. Friedlander, F. J. Nogales, and S. Scholtes, *A two-sided relaxation scheme for mathematical programs with equilibrium constraints*, SIAM J. Optim., 16 (2006), pp. 587–609

[38] H. Önal, *A modified simplex approach for solving bilevel linear programming problems*, European J. Oper. Res., 67 (1993), pp. 126–135.

[39] J.-S. Pang, *Three modeling paradigms in mathematical programming*, Math. Program., 125 (2010), pp. 297–323.

[40] J.-S. Pang and M. Fukushima, *Complementarity constraint qualifications and simplified B-stationarity conditions for mathematical programs with equilibrium constraints*, Comput. Optim. Appl., 13 (1999), pp. 111–136.

[41] J.-S. Pang and S. Leyffer, *On the global minimization of the value-at-risk*, Optim. Methods Softw., 19 (2004), pp. 611–631.

[42] A. B. Poore and Q. Al-Hassan, *The expanded Lagrangian system for constrained optimization problems*, SIAM J. Control Optim., 26 (1988), pp. 417–427.

[43] A. Prékopa, *Probabilistic Programming*, in Stochastic Programming, A. Ruszczynski and A. Shapiro, eds., Handb. Operations Research and Management Science 10, Elsevier, Amsterdam, 2003.

[44] R. Saigal, *On the convergence rate of algorithms for solving equations that are based on methods of complementary pivoting*, Math. Oper. Res., 22 (1977), pp. 108–124.

[45] H. Scheel and S. Scholtes, *Mathematical programs with complementarity constraints: Stationarity, optimality and sensitivity*, Math. Oper. Res., 25 (2000), pp. 1–22.
[46] S. Scholtes, *Convergence properties of a regularization scheme for mathematical programs with complementarity constraints*, SIAM J. Optim., 11 (2001), pp. 918–936.
[47] H. Sellami, *A homotopy continuation method for solving normal equations*, Math. Program., 82 (1998), pp. 317–337.
[48] H. Sellami and S. M. Robinson, *Implementation of a continuation method for normal maps*, Math. Program., 76 (1997), pp. 563–578.
[49] L. T. Watson, *A globally convergent algorithm for computing fixed points of C^2 maps*, Appl. Math. Comput., 5 (1979), pp. 297–311.
[50] L. T. Watson, *Solving the nonlinear complementarity problem by a homotopy method*, SIAM J. Control Optim., 17 (1979), pp. 36–46.
[51] L. T. Watson, *Computational experience with the Chow-Yorke algorithm*, Math. Program., 19 (1980), pp. 92–101.
[52] L. T. Watson, *Numerical linear algebra aspects of globally convergent homotopy methods*, SIAM Rev., 28 (1986), pp. 529–545.
[53] L. T. Watson, *A survey of probability-one homotopy methods for engineering optimization*, Arabian J. Sci. Eng., 16 (1991), pp. 297–323.
[54] L. T. Watson, *Theory of globally convergent probability-one homotopies for nonlinear programming*, SIAM J. Optim., 11 (2000), pp. 761–780.
[55] L. T. Watson, *Probability-one homotopies in computational science* J. Comput. Appl. Math., 140 (2002), pp. 785–807.
[56] L. T. Watson, S. C. Billups, and A. P. Morgan, HOMPACK: *A suite of codes for globally convergent homotopy algorithms*, ACM Trans. Math. Software, 13 (1987), pp. 281–310.
[57] L. T. Watson, J. P. Bixler, and A. B. Poore, *Continuous homotopies for the linear complementarity problem*, SIAM J. Matrix Anal. Appl., 10 (1989), pp. 259–277.
[58] L. T. Watson and R. T. Haftka, *Modern homotopy methods in optimization*, Comput. Methods Appl. Mech. Engrg., 74 (1989), pp. 289–305.
[59] L. T. Watson and M. R. Scott, *Solving spline collocation approximations to nonlinear two-point boundary value problems by a homotopy method*, Appl. Math. Comput., 24 (1987), pp. 333–357.
[60] L. T. Watson and L. R. Scott, *Solving Galerkin approximations to nonlinear two-point boundary value problems by a globally convergent homotopy method*, SIAM J. Sci. Statist. Comput., 8 (1987), pp. 768–789.
[61] L. T. Watson, M. Sosonkina, R. C. Melville, A. P. Morgan, and H. F. Walker, *Algorithm 777:* HOMPACK90: *A suite of FORTRAN* 90 *codes for globally convergent homotopy algorithms*, ACM Trans. Math. Software, 23 (1997), pp. 514–549.
[62] X. Xiao, L. Zhang, and J. Zhang, *A smoothing Newton method for a type of inverse semidefinite quadratic programming problem*, J. Comput. Appl. Math., 223 (2009), pp. 485–498.
[63] J. J. Ye, *Optimality conditions for optimization problems with complementarity constraints*, SIAM J. Optim., 9 (1999), pp. 374–387.
[64] J. J. Ye, *Necessary and sufficient optimality conditions for mathematical programs with equilibrium constraints*, J. Math. Anal. Appl., 307 (2005), pp. 350–369.
[65] J. Zhang, and L. Zhang, *An augmented Lagrangian method for a type of inverse quadratic programming problem*, Appl. Math. Optim., 61 (2010), pp. 57–83.