

RESEARCH

Open Access



# Globally convergent method for designing twice spline contractual function

Zhichuan Zhu<sup>1</sup>, Yonghong Yao<sup>2</sup> and Mihai Postolache<sup>3,4,5\*</sup>

\*Correspondence:

[mihai@mathem.pub.ro](mailto:mihai@mathem.pub.ro)

<sup>3</sup>Center for General Education, China Medical University, Taichung, Taiwan

<sup>4</sup>Gh. Mihoc-C. Iacob Institute of Mathematical Statistics and Applied Mathematics, Romanian Academy, Bucharest, Romania

Full list of author information is available at the end of the article

## Abstract

To design a quadratic spline contractual function in the case of discretely unknown nodes, a modified constraint shifting homotopy algorithm for solving principal–agent problems is constructed in the paper. Then the existence of globally convergent solution to KKT systems for the principal–agent problem with spline contractual function is proved under a much weaker condition. The proposed algorithm only requires that any initial point is in the shifted feasible set but not necessarily in the original feasible set.

**MSC:** 90C30; 65H10

**Keywords:** Principal–agent problem; Homotopy method; Normal cone condition; Constraint shifting

## 1 Problem description

In the paper, we will consider the following principal–agent bilevel programming problem, in which the principal can neither observe nor verify the agent’s action:

$$\begin{aligned} & \max_{s(x), a \in A} U(s(x), a), \\ & \text{s.t. } V(s(x), a) \geq V_0 \\ & a \in \arg \max V(s(x), a), \end{aligned} \tag{1}$$

where  $U(s(x), a)$  and  $V(s(x), a)$  denote expected utility of the principal and agent, respectively,  $s(x)$  is the contract when the agent’s action is  $a$  and a realized output  $x \in X = [x, \bar{x}] \subset \mathbb{R}_{++}$  which denotes the set of possible outcomes. The agent chooses an action  $a \in [a, \bar{a}]$  on the basis of the agreed payment schedule  $s(x)$  and has a separable von Neumann–Morgenstern utility  $v(s(x), a) = v(s(x)) - c(a)$ , which is concave. Let  $F(x, a)$  and  $f(x, a)$  denote the continuous distribution and density function of outcome  $x$  given that action  $a$  is undertaken by the agent, which are assumed sufficiently continuous differentiable for all  $x$  and  $a$ .

Since the principal–agent model was proposed, an efficient and conventional method for analyzing the problem (1) was the so-called first-order approach. But Mirrlees [1, 2] was the first scholar to point out that the analyzing method is generally incorrect. In the

subsequent paper, Mirrlees [3] gave the validity conditions for implying the first-order approach which is named the monotone likelihood ratio condition (MLRC) and the convexity of the distribution function condition (CDFC). Then Rogerson [4] offered a correct and much simpler proof than Mirrlees. Jewitt [5] provided some conditions which can be used for justifying the first-order approach in the multi-statistic case. For more literature about the first-order approach for solving principal–agent problems see, e.g., [6–14] and the references therein.

However, since the problem (1) is an infinite-dimensional nonconvex bilevel programming, lots of papers have focused on the theoretical analysis of the validity on first-order approach, few papers were presented for directly solving the principal–agent model. Prescott [15] computed solutions to moral-hazard programs using the Dantzig–Wolfe decomposition algorithm in a discrete condition, but which is required to be block angular in order to turn it into a linear program (LP) and as the cardinalities of the underlying sets increase, LP quickly grows in size making computation infeasible. Su and Judd [16] studied computational aspects of formulating moral-hazard problems as a mathematical program with equilibrium constraints (MPEC) and proposed a hybrid procedure that combined the linear programming (LP) approach with lotteries, but this algorithm only obtained the local convergence. Armstrong *et al.* [17] formulated two complementary generalized principal–agent models that incorporate features observed in a real-world contracting environment as MPEC and solved the resulting models by the state-of-the-art numerical algorithms. Cecchini *et al.* [18] solved numerically the principal–agent problems written as a linear-exponential-normal model by solving bilevel programming problems using the ellipsoid algorithm in the case of assuming the performance measures is linear. To solve the problem, Zhu *et al.* [19] firstly proposed a modified constraint shifting homotopy method by designing a piecewise linear contractual function under some typical risk averse utility functions and the typical distribution functions. Renner and Schmedders [20] reformulated the agent’s utility maximization problem as an equivalent system of equations and inequalities under the assumption that the agent’s expected utility is a rational function of the action and computed an approximate solution to the nonpolynomial problems transformed by the principal’s utility maximization problem using the polynomial optimization approach. For the general distribution functions, Zhu and Yu [21] presented a constraint set swelling homotopy method for computing the solution to its Karush–Kuhn–Tucker (KKT) systems under the case of designing piecewise linear contractual function by using the composite Simpson’s rule to approximately compute integration, and they proved the global convergence of the homotopy path under much weaker conditions. Moreover, Zhu and Yu [22] proposed another modified homotopy method for computing the solution to its KKT systems under requiring only an interior point and, not necessarily, a feasible initial approximation for the constraint shifting set. However, it is usually assumed that the discrete nodes are known.

To directly solve the problem (1) by designing contractual function, as in the literature, the conditions MLRC and CDFC are also assumed to hold in this paper: 1.  $f(x, a)$  satisfies MLRC, i.e.,  $\frac{f_a}{f}$  is nondecreasing in  $x$  for every  $a$ ; and 2.  $f(x, a)$  satisfies CDFC, i.e.,  $F_{aa}(x, a) \geq 0$ . In this respect, the incentive compatibility constraint can be replaced by the first-order condition and the principal–agent problem can be turned into the following equiv-

alently single-level nonconvex programming:

$$\begin{aligned}
 & \max_{s(x), a \in A} U(s(x), a), \\
 & \text{s.t. } V(s(x), a) \geq V_0 \\
 & \quad V_a(s(x), a) = 0.
 \end{aligned} \tag{2}$$

Then the solution to the principal–agent problems can be computed by solving the equivalent problem (2).

Since the combined homotopy method has global convergence and can be efficiently implemented for solving nonlinear programming, fixed point problems, variational inequalities, bilevel programming and other nonlinear problems, it has been paid much attention to since the 1990s; see, e.g., [23–33]. Since for the existing combined homotopy methods for solving the equivalent problem (2) it is usually assumed that the discrete nodes are known and the contract is designed as a piecewise linear function, to design a quadratic spline contractual function in the case of discretely unknown nodes, a new constraint shifting combined homotopy for solving the KKT systems of the principal–agent problem will be constructed and the global convergence will be proved in this paper.

In the paper, the standard model of principal–agent problems with designing quadratic spline contractual function is considered. In Sect. 2, the constraint shifting combined homotopy for the basic principal–agent problems is constructed and some lemmas from differential topology are introduced. In Sect. 3, the main results will be presented and the existence of a smooth path from any given initial point in shifted feasible set to a solution to the KKT systems is proven. In Sect. 4, a detailed predictor–corrector algorithm is presented.

## 2 Preliminaries

We assume that the wage contract is the following quadratic spline function. Setting the discrete nodes which are unknown in  $X = [\underline{x}, \bar{x}]$  satisfying  $\underline{x} = x_1 \leq x_2 \leq \dots \leq x_m \leq x_{m+1} = \bar{x}$ , the wage contract is

$$s(x) = \begin{cases} p_1x^2 + q_1x + r_1, & x_1 \leq x \leq x_2, \\ p_2x^2 + q_2x + r_2, & x_2 \leq x \leq x_3, \\ \vdots & \\ p_mx^2 + q_mx + r_m, & x_m \leq x \leq x_{m+1}, \end{cases}$$

which satisfies  $p_ix_{i+1}^2 + q_ix_{i+1} + r_i = p_{i+1}x_{i+1}^2 + q_{i+1}x_{i+1} + r_{i+1}$ ,  $2p_ix_{i+1} + q_i = 2p_{i+1}x_{i+1} + q_{i+1}$ ,  $i = 1, 2, \dots, m - 1$ , and  $p_i \leq 0$ ,  $i = 1, 2, \dots, m$ .

Then we can get the following expected utilities of the principal and agent, respectively:

$$\begin{aligned}
 U(s(x), a) &= \int_{x \in X} u(x - s(x))f(x, a) dx \\
 &= \int_{x \in X} (x - s(x))f(x, a) dx \\
 &= \sum_{i=1}^m \int_{x_i}^{x_{i+1}} (x - p_ix^2 - q_ix - r_i)f(x, a) dx
 \end{aligned}$$

and

$$\begin{aligned} V(s(x), a) &= \int_{x \in X} v(s(x))f(x, a) \, dx - a \\ &= \int_{x \in X} s(x)f(x, a) \, dx - a \\ &= \sum_{i=1}^m \int_{x_i}^{x_{i+1}} (p_i x^2 + q_i x + r_i)f(x, a) \, dx - a. \end{aligned}$$

Because the integral does not change the differentiability, the expected utilities of the principal and agent are also sufficiently smooth for all  $x$  and  $a$ .

The expected utilities of principal and agent are concave on action  $a$  under CDFC. Therefore, the relaxed Pareto-optimization programming of the principal–agent problem for designing the wage contract as quadratic spline function can be reformulated as follows:

$$\begin{aligned} \min_{(p_i, q_i, r_i) \in \mathbb{R}} & \sum_{i=1}^m \int_{x_i}^{x_{i+1}} [p_i x^2 + (q_i - 1)x + r_i]f(x, a) \, dx \\ \text{s.t.} & - \sum_{i=1}^m \int_{x_i}^{x_{i+1}} (p_i x^2 + q_i x + r_i)f(x, a) \, dx + a + V_0 \leq 0, \\ & a - \bar{a} \leq 0, \\ & -a + \underline{a} \leq 0, \tag{3} \\ & p_i \leq 0, \quad i = 1, 2, \dots, m, \\ & p_i x_{i+1}^2 + q_i x_{i+1} + r_i - p_{i+1} x_{i+1}^2 - q_{i+1} x_{i+1} - r_{i+1} = 0, \quad i = 1, 2, \dots, m - 1, \\ & 2p_i x_{i+1} + q_i - 2p_{i+1} x_{i+1} - q_{i+1} = 0, \quad i = 1, 2, \dots, m - 1, \\ & \sum_{i=1}^m \int_{x_i}^{x_{i+1}} (p_i x^2 + q_i x + r_i)f_a(x, a) \, dx - 1 = 0. \end{aligned}$$

For convenience, we give the following notation, respectively:

$$\begin{aligned} F(p, q, r, a, \zeta) &= \sum_{i=1}^m \int_{x_i}^{x_{i+1}} (p_i x^2 + q_i x + r_i - x)f(x, a) \, dx, \\ g_1(p, q, r, a, \zeta) &= p_1, \\ g_2(p, q, r, a, \zeta) &= p_2, \\ & \vdots \\ g_m(p, q, r, a, \zeta) &= p_m, \\ g_{m+1}(p, q, r, a, \zeta) &= - \sum_{i=1}^m \int_{x_i}^{x_{i+1}} (p_i x^2 + q_i x + r_i)f(x, a) \, dx + a + V_0, \\ g_{m+2}(p, q, r, a, \zeta) &= a - \bar{a}, \\ g_{m+3}(p, q, r, a, \zeta) &= -a + \underline{a}, \end{aligned}$$

$$\begin{aligned}
 h_1(p, q, r, a, \zeta) &= p_1x_2^2 + q_1x_2 + r_1 - p_2x_2^2 - q_2x_2 - r_2, \\
 h_2(p, q, r, a, \zeta) &= p_2x_3^2 + q_2x_3 + r_2 - p_3x_3^2 - q_3x_3 - r_3, \\
 &\vdots \\
 h_{m-1}(p, q, r, a, \zeta) &= p_{m-1}x_m^2 + q_{m-1}x_m + r_{m-1} - p_mx_m^2 - q_mx_m - r_m, \\
 h_m(p, q, r, a, \zeta) &= 2p_1x_2 + q_1 - 2p_2x_2 - q_2, \\
 h_{m+1}(p, q, r, a, \zeta) &= 2p_2x_3 + q_2 - 2p_3x_3 - q_3, \\
 &\vdots \\
 h_{2m-2}(p, q, r, a, \zeta) &= 2p_{m-1}x_m + q_{m-1} - 2p_mx_m - q_m, \\
 h_{2m-1}(p, q, r, a, \zeta) &= \sum_{i=1}^m \int_{x_i}^{x_{i+1}} (p_ix^2 + q_ix + r_i)f_a(x, a) dx - 1,
 \end{aligned}$$

where  $p = (p_1, p_2, \dots, p_m)^T$ ,  $q = (q_1, q_2, \dots, q_m)^T$ ,  $r = (r_1, r_2, \dots, r_m)$ ,  $a = a$ ,  $\zeta = (x_2, x_3, \dots, x_m)^T$ .

Then, when the wage contract is quadratic spline, the relaxed Pareto-optimization programming (3) can be written as follows:

$$\begin{aligned}
 &\min F(p, q, r, a, \zeta), \\
 &\text{s.t. } g_i(p, q, r, a, \zeta) \leq 0, \quad i \in \{1, 2, \dots, m + 3\}, \\
 &\quad h_j(p, q, r, a, \zeta) = 0, \quad j \in \{1, 2, \dots, 2m - 1\}.
 \end{aligned} \tag{4}$$

Let  $\mathbb{R}^n$ ,  $\mathbb{R}_+^n$ ,  $\mathbb{R}_{++}^n$  denote  $n$ -dimensional Euclidean space, nonnegative orthant and positive orthant of  $\mathbb{R}^n$ , respectively. It is well known that the solution of the optimization problem can be obtained from the KKT system for a convex nonlinear programming problem. But for a nonconvex nonlinear programming problem, we can only obtain the solution to its KKT system.

Therefore, the aim is to solve the following KKT system of the relaxed Pareto-optimization programming (4):

$$\begin{aligned}
 &\nabla F(p, q, r, a, \zeta) + \sum_{i=1}^{m+3} \nabla g_i(p, q, r, a, \zeta)y_i + \sum_{j=1}^{2m-1} \nabla h_j(p, q, r, a, \zeta)z_j = 0, \\
 &h_j(p, q, r, a, \zeta) = 0, \quad j = 1, 2, \dots, 2m - 1, \\
 &y_i g_i(p, q, r, a, \zeta) = 0, \quad y_i \geq 0, \quad g_i(p, q, r, a, \zeta) \leq 0, \quad i = 1, 2, \dots, m + 3,
 \end{aligned} \tag{5}$$

where  $y \in \mathbb{R}_+^{m+3}$ ,  $z \in \mathbb{R}^{2m-1}$ ,  $\nabla = (\frac{\partial}{\partial p}, \frac{\partial}{\partial q}, \frac{\partial}{\partial r}, \frac{\partial}{\partial a}, \frac{\partial}{\partial \zeta})^T$ .

Consider the shifted constraint functions be  $\tilde{g}_i(\theta, \mu) = g_i(\theta) - \mu^\sigma \tau$ , with  $\tau \in \mathbb{R}_{++}^{m+3}$  and  $\tilde{h}_j(\theta, \mu) = h_j(\theta) - \mu h_j(\theta^0)$ , which made  $\tilde{g}_i(\theta, 0) = g_i(\theta)$ , and  $\tilde{h}_j(\theta, 0) = h_j(\theta)$ . Therefore,  $\nabla \tilde{g}_i(\theta, \mu) = \nabla g_i(\theta)$  and  $\nabla \tilde{h}_j(\theta, \mu) = \nabla h_j(\theta)$ . Because of the properties of  $g_i(\theta)$  and  $h_j(\theta)$ ,  $\nabla \tilde{g}_i(\theta, \mu)$  and  $\nabla \tilde{h}_j(\theta, \mu)$  are also sufficiently smooth.

For convenience, the following notations will be used in the paper:

$$\theta = (p, q, r, a, \zeta)^T,$$

$$\begin{aligned} \Omega &= \{\theta | g_i(\theta) \leq 0, h_j(\theta) = 0, i \in \{1, 2, \dots, m + 3\}, j \in \{1, 2, \dots, 2m - 1\}\}, \\ \Omega_\mu &= \{\theta | \tilde{g}_i(\theta, \mu) \leq 0, \tilde{h}_j(\theta, \mu) = 0, i \in \{1, \dots, m + 3\}, j \in \{1, \dots, 2m - 1\}\}, \\ \Omega_\mu^0 &= \{\theta | \tilde{g}_i(\theta, \mu) < 0, \tilde{h}_j(\theta, \mu) = 0, i \in \{1, \dots, m + 3\}, j \in \{1, \dots, 2m - 1\}\}, \\ \partial\Omega_\mu &= \Omega_\mu \setminus \Omega_\mu^0, \quad \text{and} \\ I_\mu(\theta) &= \{i | \tilde{g}_i(\theta, \mu) = 0, i = 1, 2, \dots, m + 3\}. \end{aligned}$$

The following assumptions will be used in this paper. A vector  $\tau \in \mathbb{R}^{m+3}_{++}$  and an open subset  $V$  of the inequality constraint set  $\{\theta : \tilde{g}_i(\theta, 1) < 0\}$  are existed to satisfy the following conditions:

- (A1)  $\forall \mu \in [0, 1], \Omega_\mu$  is bounded and  $\Omega_\mu^0 \neq \emptyset$ .
- (A2)  $\forall \theta \in V, \nabla h(\theta)$  is a full column rank matrix. For any parameter  $\mu \in [0, 1]$  and  $\theta \in \Omega_\mu, (\nabla \tilde{g}_i(\theta, \mu)_{i \in I_\mu(\theta)}, \nabla h(\theta))$  is positive linearly independent at  $\theta$ , i.e.,  $\sum_{i \in I(\theta)} \alpha_i \nabla g_i(\theta) + \sum_{j=1}^m \beta_j \nabla h_j(\theta) = 0, \alpha_i \geq 0, \beta_j \in \mathbb{R} \Rightarrow \alpha_i = \beta_j = 0$ .
- (A3) (The normal cone condition)  $\forall \theta \in \partial\Omega_1$ , the normal cone of  $\Omega_1$  only meets  $\partial\Omega_1$  at  $\theta$ , i.e.,  $\forall \theta \in \Omega_1$ ,

$$\left\{ \theta + \sum_{i \in I_1(\theta)} \nabla \tilde{g}_i(\theta, 1) y_i + \nabla h(\theta) z \mid i \in I_1(\theta), y_i \geq 0, z \in \mathbb{R}^m \right\} \cap \Omega_1 = \{\theta\}.$$

In order to compute a solution to the KKT system of the relaxed Pareto-optimization programming (4), for any randomly chosen vector  $(\theta^0, \xi) \in V \times \mathbb{R}^{4m}$  and any given vector  $\eta \in \mathbb{R}^{m+3}_{++}$ , we construct the constraint shifting combined homotopy as follows:

$$H(w, w^0, \mu) = \begin{pmatrix} (1 - \mu)(\nabla F(\theta) + \nabla \tilde{g}(\theta, \mu)y + \nabla h(\theta)z + \mu(\theta - \theta^0) + \mu(1 - \mu)\xi) \\ Y\tilde{g}(\theta, \mu) + \mu\eta \\ h(\theta) - \mu h(\theta^0) \end{pmatrix}, \tag{6}$$

where  $w = (\theta, y, z)^T \in \mathbb{R}^{4m} \times \mathbb{R}^{m+3}_+ \times \mathbb{R}^{2m-1}, w^0 = (\theta^0, y^0, z^0)^T$ , and  $Y = \text{diag}(y_1, y_2, \dots, y_{m+3})$ . When  $\mu = 0$ , the homotopy equation  $H(w, w^0, 0) = 0$  turns to the KKT system (5) of the relaxed Pareto-optimization programming (4).

**Lemma 2.1** *Suppose that  $H(w, w^0, \mu)$  is defined as (6) and assumptions (A1), (A2) and (A3) hold, then the homotopy equation  $H(w, w^0, 1) = 0$  has a unique solution*

$$(\theta, y, z) = (\theta^0, y^0, z^0) = (\theta^0, -[\text{diag}(\tilde{g}(\theta^0, 1))]^{-1} \eta, 0).$$

*Proof* Let  $\bar{w} = (\bar{\theta}, \bar{y}, \bar{z})$  be a solution of  $H(w, w^0, 1)$ . From  $H(\bar{w}, w^0, 1) = 0$  and  $\bar{y} \geq 0$ , we get  $\bar{\theta} \in \Omega_1$ . In the following, we will prove  $\bar{\theta} = \theta^0$  by contradiction. We assume  $\bar{\theta} \neq \theta^0$ , which implies that  $\bar{z} \neq 0$ , together with the first equation of  $H(w, w^0, 1) = 0$ , we have  $\theta^0 = \bar{\theta} + \nabla h(\bar{\theta})\bar{z}$ , which contradicts with assumption (A2). Hence,  $\bar{\theta} = \theta^0$ . From assumption (A2) and  $\nabla h(\bar{\theta})\bar{z} = 0$ , we get  $\bar{z} = 0$ . From the second equation of  $H(\bar{w}, w^0, 1) = 0$  and  $\theta^0 \in \Omega_1^0$ , we have  $\bar{y} = -[\text{diag}(\tilde{g}(\theta^0, 1))]^{-1} \eta$ . Therefore, we obtain the result.  $\square$

The following lemmas from differential topology which can be found in Refs. [34–36] will be used in the next section. At first, let  $U \subset \mathbb{R}^n$  be an open set, let  $\phi: U \rightarrow \mathbb{R}^p$  be a  $C^\alpha$

( $\alpha > \max\{0, n - p\}$ ) mapping; we say that  $y \in \mathbb{R}^p$  is a regular value for  $\phi$  if

$$\text{Range}[\partial\phi(x)/\partial x] = \mathbb{R}^p, \quad \forall x \in \phi^{-1}(y).$$

**Lemma 2.2** *Let  $V \subset \mathbb{R}^n, U \subset \mathbb{R}^m$  be open sets, and let  $\phi: V \times U \rightarrow \mathbb{R}^k$  be a  $C^\alpha$  mapping, where  $\alpha > \max\{0, m - k\}$ . If  $0 \in \mathbb{R}^k$  is a regular value of  $\phi$ , then, for almost all  $a \in V, 0$  is a regular value of  $\phi_a = F(a, \cdot)$ .*

**Lemma 2.3** *Let  $\phi: U \subset \mathbb{R}^n \rightarrow \mathbb{R}^p$  be a  $C^\alpha$  ( $\alpha > \max\{0, n - p\}$ ). If  $0$  is a regular value of  $\phi$ , then  $\phi^{-1}(0)$  consists of some  $(n - p)$  dimensional  $C^\alpha$  manifolds.*

**Lemma 2.4** *A one-dimensional smooth manifold is diffeomorphic to a unit circle or a unit interval.*

### 3 Main result

For the sake of convenience, for any given  $w^0 \in V \times \mathbb{R}_{++}^{m+3} \times \mathbb{R}^{2m-1}$ ,  $H(w, w^0, \mu)$  in (6) is rewritten as  $H_{w^0}(w, \mu) = H(w, w^0, \mu)$ , and the zero-point set of  $H_{w^0}(w, \mu)$  is written as follows:

$$H_{w^0}^{-1}(0) = \{(w, \mu) \in \Omega_\mu \times \mathbb{R}_{++}^{m+3} \times \mathbb{R}^{2m-1} \times (0, 1] : H_{w^0}(w, \mu) = 0\}.$$

**Theorem 3.1** *Suppose that assumptions (A1) and (A2) hold for any chosen vector  $w^0 \in V$  and the homotopy is defined by (6), then, for almost all  $(w^0, \xi) \in V \times \mathbb{R}_{++}^{m+3} \times \mathbb{R}^{2m-1} \times \mathbb{R}^{4m}$ , the zero-point set  $H_{w^0}^{-1}(0)$  must contain a smooth curve  $\Gamma_{w^0}$  starting from  $(\theta^0, y^0, z^0, 1)$ . Besides, if assumption (A3) holds, then the smooth curve  $\Gamma_{w^0}$  terminates or approaches to the hyperplane  $\mu = 0$ . If  $(\bar{\theta}, \bar{y}, \bar{z}, 0)$  is an ending limit point of the smooth curve  $\Gamma_{w^0}$ , then  $\bar{w} = (\bar{\theta}, \bar{y}, \bar{z})$  is a solution to KKT system (5) of the relaxed Pareto-optimization programming (4).*

*Proof* Suppose that  $\tilde{H}(w, \theta^0, \xi, \mu): \Omega_\mu \times \mathbb{R}_{++}^{m+3} \times \mathbb{R}^{2m-1} \times V \times \mathbb{R}^{4m} \times (0, 1] \rightarrow \Omega_\mu \times \mathbb{R}_{++}^{m+3} \times \mathbb{R}^{2m-1}$  is the same map as  $H(w, w^0, \mu)$  but taking  $(\theta^0, \xi)$  as variate. Considering the following submatrix of the Jacobian  $D\tilde{H}(w, \theta^0, \xi, \mu)$ :

$$\frac{\partial \tilde{H}(w, \theta^0, \xi, \mu)}{\partial (\theta^0, y, \xi)} = \begin{pmatrix} -\mu I & (1 - \mu)\nabla \tilde{g}(\theta, \mu) & \mu(1 - \mu) \\ 0 & \text{diag}(\tilde{g}(\theta, \mu)) & 0 \\ -\mu \nabla h(\theta^0)^T & 0 & 0 \end{pmatrix}.$$

For all  $\mu \in (0, 1)$ , by the fact that  $\eta > 0$  and  $Y\tilde{g}(\theta, \mu) + \mu\eta = 0$ , we see that  $\text{diag}(\tilde{g}(\theta, \mu))$  is nonsingular. By assumption (A2), we see that the matrix  $\nabla h(\theta^0)^T$  is full row rank, which implies that  $\frac{\partial \tilde{H}(w, \theta^0, \xi, \mu)}{\partial (\theta^0, y, \xi)}$  is a matrix of full row rank for any  $\mu \in (0, 1)$ . Hence, the matrix  $D\tilde{H}(w, \theta^0, \xi, \mu)$  is full row rank for any solution of the equation  $\tilde{H}(w, \theta^0, \xi, \mu) = 0$  in  $\mathbb{R}_{++}^{m+3} \times \mathbb{R}^{2m-1} \times V \times \mathbb{R}^{4m} \times (0, 1)$ .

From the matrix

$$\frac{\partial H(w, w^0, \mu)}{\partial w} = \begin{pmatrix} \Pi & (1 - \mu)\nabla \tilde{g}(\theta, \mu) & \nabla h(\theta) \\ Y\nabla \tilde{g}(\theta, \mu)^T & \text{diag}(\tilde{g}(\theta, \mu)) & 0 \\ \nabla h(\theta)^T & 0 & 0 \end{pmatrix},$$

where  $\Pi = (1 - \mu)(\nabla^2 F(\theta) + \sum_{i=1}^{m+3} y_i \nabla^2 \tilde{g}_i(\theta, \mu)) + \sum_{j=1}^{2m-1} z_j \nabla^2 h_j(\theta) + \mu I$ , for any  $\mu \in (0, 1]$  and the chosen initial point  $w^0 = (\theta^0, y^0, z^0)$ , using  $z^0 = 0$ , we can obtain the following matrix:

$$\frac{\partial H(w^0, w^0, 1)}{\partial w} = \begin{pmatrix} I & 0 & \nabla h(\theta^0) \\ Y^0 \nabla \tilde{g}(\theta^0, 1)^T & \text{diag}(\tilde{g}(\theta^0, 1)) & 0 \\ \nabla h(\theta^0)^T & 0 & 0 \end{pmatrix},$$

which is also nonsingular. So, we can see that the matrix  $D\tilde{H}(w, \theta^0, \xi, \mu)$  is nonsingular for any  $\mu \in (0, 1]$ . Thus, 0 is a regular value of  $\tilde{H}(w, \theta^0, \xi, \mu)$ . By Lemma 2.2, we see that, for almost all  $(\theta^0, \xi) \in V \times \mathbb{R}^{4m}$ , 0 is a regular value of  $H(w, w^0, \mu)$ . From Lemma 2.3, if 0 is a regular value of  $H(w, w^0, \mu)$ ,  $\frac{\partial H(w^0, w^0, 1)}{\partial w}$  is nonsingular and by the fact of  $H(w^0, w^0, 1) = 0$ ,  $H_{w^0}^{-1}(0)$  must contain a smooth curve  $\Gamma_{w^0} \subset \Omega_\mu^0 \times \mathbb{R}_+^{m+3} \times \mathbb{R}^{2m-1} \times (0, 1]$  starting from  $(\theta^0, y^0, z^0, 1)$  and going into  $\Omega_1^0 \times \mathbb{R}_+^{m+3} \times \mathbb{R}^{2m-1} \times (0, 1)$  and terminating in the boundary of  $\Omega_\mu \times \mathbb{R}_+^{m+3} \times \mathbb{R}^{2m-1} \times [0, 1]$ . Then from Lemma 2.4,  $\Gamma_{w^0} \subset \Omega_\mu^0 \times \mathbb{R}_+^{m+3} \times \mathbb{R}^{2m-1} \times (0, 1]$  must be diffeomorphic to a unit circle or a unit interval  $[0, 1)$ . Since the matrix  $\frac{\partial H(w^0, w^0, 1)}{\partial w}$  is nonsingular, the smooth curve  $\Gamma_{w^0}$  cannot be diffeomorphic to a unit circle. Therefore,  $\Gamma_{w^0}$  must be diffeomorphic to  $[0, 1)$ .

When  $\mu \rightarrow 0$ , the limit points of  $\Gamma_{w^0}$  belong to  $\partial(\Omega_\mu \times \mathbb{R}_+^{m+3} \times \mathbb{R}^{2m-1} \times (0, 1))$ . Let  $(\bar{\theta}, \bar{y}, \bar{z}, \bar{\mu})$  be a limit point of  $\Gamma_{w^0}$  when  $\mu \rightarrow 0$ . Only the following five cases are possible:

- (i)  $(\bar{\theta}, \bar{y}, \bar{z}) \in \Omega_1 \times \mathbb{R}_+^{m+3} \times \mathbb{R}^{2m-1}$ ,  $\bar{\mu} = 1$ ,  $\|(\bar{y}, \bar{z})\| < \infty$ ;
- (ii)  $(\bar{\theta}, \bar{y}, \bar{z}) \in \Omega_{\bar{\mu}} \times \mathbb{R}_+^{m+3} \times \mathbb{R}^{2m-1}$ ,  $\bar{\mu} \in [0, 1]$ ,  $\|(\bar{y}, \bar{z})\| = \infty$ ;
- (iii)  $(\bar{\theta}, \bar{y}, \bar{z}) \in \Omega_{\bar{\mu}} \times \partial \mathbb{R}_+^{m+3} \times \mathbb{R}^{2m-1}$ ,  $\bar{\mu} \in (0, 1)$ ,  $\|(\bar{y}, \bar{z})\| < \infty$ ;
- (iv)  $(\bar{\theta}, \bar{y}, \bar{z}) \in \partial \Omega_{\bar{\mu}} \times \mathbb{R}_+^{m+3} \times \mathbb{R}^{2m-1}$ ,  $\bar{\mu} \in (0, 1)$ ,  $\|(\bar{y}, \bar{z})\| < \infty$ ;
- (v)  $(\bar{\theta}, \bar{y}, \bar{z}) \in \Omega \times \mathbb{R}_+^{m+3} \times \mathbb{R}^{2m-1}$ ,  $\bar{\mu} = 0$ ,  $\|(\bar{y}, \bar{z})\| < \infty$ .

From Lemma 2.1, the homotopy equation  $H(w, w^0, 1) = 0$  has only one simple solution  $w^0 = (\theta^0, y^0, z^0)$  in  $\Omega_1^0 \times \mathbb{R}_+^{m+3} \times \mathbb{R}^{2m-1}$ , and because the matrix  $\frac{\partial H(w^0, w^0, 1)}{\partial w}$  is nonsingular, case (i) is impossible.

If case (ii) happens, by assumption (A1), there must be a subsequence  $(\theta^k, y^k, z^k, \mu_k) \subset \Gamma_{w^0}$  such that  $\|(y^k, z^k)\| \rightarrow \infty$  and  $\mu_k \rightarrow \bar{\mu}$ ,  $\theta^k \rightarrow \bar{\theta}$  as  $k \rightarrow \infty$ . Only the following three subcases are possible: (1)  $\bar{\mu} = 1$ ; (2)  $\bar{\mu} \in (0, 1)$ ; (3)  $\bar{\mu} = 0$ .

From the first equation of (6), we have

$$(1 - \mu_k) \left( \nabla F(\theta^k) + \sum_{i=1}^{m+3} y_i^k \nabla \tilde{g}_i(\theta^k, \mu_k) \right) + \nabla h(\theta^k) z^k + \mu_k (\theta^k - \theta^0) + \mu_k (1 - \mu_k) \xi = 0. \tag{7}$$

(1) When  $\bar{\mu} = 1$ , if  $\|(1 - \mu^k)y^k, z^k\| < \infty$ , without loss of generality we can suppose that  $((1 - \mu_k)y_k, z^k) \rightarrow (\bar{y}, \bar{z})$ . Then  $\bar{y}_i = 0$  for  $i \notin I_1(\bar{\theta})$  from the second equation of (6). Taking the limit in (7) as  $k \rightarrow +\infty$ , we have

$$\begin{aligned} \theta^0 &= \bar{\theta} + \lim_{k \rightarrow \infty} [(1 - \mu_k)(\nabla F(\theta^k) + \nabla \tilde{g}(\theta^k, \mu_k)y^k) + \nabla h(\theta^k)z^k] \\ &= \bar{\theta} + h(\bar{\theta})\bar{z} + \lim_{k \rightarrow \infty} \sum_{i \in I_1(\bar{\theta})} (1 - \mu_k)y_i^k \nabla \tilde{g}_i(\theta^k, \mu_k) \end{aligned}$$



$$= \bar{\theta} + h(\bar{\theta})\bar{z} + \sum_{i \in I_1(\bar{\theta})} \bar{y}_i \nabla \tilde{g}_i(\bar{\theta}, 1), \tag{8}$$

which contradicts with assumption (A3).

If  $\|(1 - \mu^k)y^k, z^k\| \rightarrow \infty$ , the discussion is the same as case (2).

(2) When  $\bar{\mu} \in (0, 1)$ , without loss of generality, we can suppose that  $((1 - \mu^k)y^k, z^k)/\|(1 - \mu^k)y^k, z^k\| \rightarrow (\bar{\alpha}, \bar{\beta})$  with  $\|(\bar{\alpha}, \bar{\beta})\| = 1$  and  $\bar{\alpha}_i = 0$  for  $i \notin I_{\bar{\mu}}(\bar{\theta})$ . Dividing both sides of Eq. (7) by  $\|(1 - \mu^k)y^k, z^k\|$  and taking the limit, we have

$$\sum_{i \in I_{\bar{\mu}}(\bar{\theta})} \bar{\alpha}_i \nabla \tilde{g}_i(\bar{\theta}, \bar{\mu}) = 0,$$

which contradicts with assumption (A2).

(3) When  $\bar{\mu} = 0$ , without loss of generality, suppose that  $(y^k, z^k)/\|y^k, z^k\| \rightarrow (\bar{\alpha}, \bar{\beta})$  with  $\|(\bar{\alpha}, \bar{\beta})\| = 1$  and  $\bar{\alpha}_i = 0$  for  $i \notin I_{\bar{\mu}}(\bar{\theta})$ . Dividing both sides of Eq. (7) by  $\|(y^k, z^k)\|$  and taking the limit, we have

$$\sum_{i \in I_{\bar{\mu}}(\bar{\theta})} \bar{\alpha}_i \nabla \tilde{g}_i(\bar{\theta}, \bar{\mu}) + \sum_{j=1}^{2m-1} \bar{\beta}_j \nabla h_j(\bar{\theta}) = 0,$$

which contradicts with assumption (A2).

Therefore, from the discussions of (1), (2) and (3), we see that case (ii) is also impossible.

Now, we discuss that cases (iii) and (iv) are also impossible. For any given  $\eta \in \mathbb{R}_{++}^{m+3}$ , from the second equation  $\text{diag}(\tilde{g}(\bar{\theta}, \bar{\mu}))\bar{y} + \bar{\mu}\eta = 0$  of  $H(\bar{w}, w^0, \bar{\mu}) = 0$  as  $k \rightarrow \infty$ , we see that  $\bar{\mu} > 0$  and  $\bar{y} \in \partial \mathbb{R}_+^{m+3}$ , i.e.,  $\bar{y}_i = 0$  for some  $1 \leq i \leq m + 3$  cannot happen simultaneously. Therefore, case (iii) is also impossible. If  $\bar{y} > 0$  and  $\bar{\mu} > 0$ , from  $\text{diag}(\tilde{g}(\bar{\theta}, \bar{\mu}))\bar{y} + \bar{\mu}\eta = 0$ , we can get  $\text{diag}(\tilde{g}(\bar{\theta}, \bar{\mu})) < 0$ , which implies that case (iv) is impossible.

As a conclusion, case (v) is the only possible case. That is  $\Gamma_{w^0}$  must terminate in or approach to the hyperplane at  $\bar{\mu} = 0$  and  $\bar{w} = (\bar{\theta}, \bar{y}, \bar{z})$  is a solution to the KKT system (5) of principal-agent problems.  $\square$

### 4 Numerical algorithm

By Theorem 3.1, the homotopy equation (6) generates a smooth curve  $\Gamma_{w^0}$  for almost all  $(w^0, \mu) \in \Omega_1^0 \times \mathbb{R}_{++}^{m+3} \times \mathbb{R}^{2m-1} \times (0, 1]$  as  $\mu \rightarrow 0$ , one can find a solution to (5). Letting  $s$  be the arc length of  $\Gamma_{w^0}$ , we can parameterize the smooth curve  $\Gamma_{w^0}$  with respect to  $s$ , i.e.

$$\begin{aligned} H(w(s), w^0, \mu(s)) &= 0, \\ w(0) &= w^0, \quad \mu(0) = 1. \end{aligned} \tag{9}$$

By differentiating Eq. (9), we can get the following theorem.

**Theorem 4.1** *The smooth homotopy path  $\Gamma_{w^0}$  is determined by the following initial value problem to the ordinary differential equation:*

$$\begin{aligned} DH(w(s), w^0, \mu(s)) \begin{pmatrix} \dot{w} \\ \dot{\mu} \end{pmatrix} &= 0, \\ (w(0), \mu(0)) &= (w^0, 1), \end{aligned} \tag{10}$$

where  $DH$  is the derivative of  $H$ ,  $s$  is the arc length of the curve  $\Gamma_{w^0}$ , and the  $w$ -component of the solution point  $(w(s^*), \mu(s^*))$  of Eq. (10) is the solution to the KKT system (5) as  $\mu(s^*) = 0$ .

As regards how to trace numerically the homotopy path  $\Gamma_{w^0}$ , we can use the standard predictor–corrector procedure, which contains three steps: the first predictor step, the midway predictor step and the corrector step. The first predictor step is usually taken by computing the tangent direction, the midway predictor steps are usually taken by using secant directions, and the corrector steps are usually taken by Newton iterations for solving an augmented system.

For the first predictor step, the tangent vector at a point on the homotopy path  $\Gamma_{w^0}$  has two directions: one positive direction which makes the arc length  $s$  gradually increase, and one negative direction which makes the arc length  $s$  gradually decrease. Since the negative direction will lead the homotopy path  $\Gamma_{w^0}$  back to the initial point  $(w(0), \mu(0)) = (w^0, 1)$ , we must trace the homotopy path along the positive direction. By the basic theory of homotopy methods, the positive direction  $v$  at any point  $(w, \mu)$  of  $\Gamma_{w^0}$  must keep the sign of the following determinant invariant:

$$\left| \begin{array}{c} DH(w(s), w^0, \mu(s)) \\ v^T \end{array} \right|.$$

**Proposition 4.1** *If the homotopy path  $\Gamma_{w^0}$  is smooth, then the positive direction  $v^0$  at the initial point  $w^0$  satisfies*

$$\text{sign} \left| \begin{array}{c} DH(w^0, w^0, 1) \\ v^{0T} \end{array} \right| = (-1)^{(m+3)+(2m-1)+1}.$$

*Proof* We have the shifted constraint functions  $\tilde{g}_i(\theta, \mu) = g_i(\theta) - \mu^\sigma \tau$ ,  $\tau \in \mathbb{R}_{++}$  and  $\tilde{h}_j(\theta, \mu) = h_j(\theta) - \mu h_j(\theta^0)$ ,  $\nabla \tilde{g}_i(\theta, \mu) = \nabla g_i(\theta)$  and  $\nabla \tilde{h}_j(\theta, \mu) = \nabla h_j(\theta)$ . From the matrix

$$DH(w, w^0, \mu) = \frac{\partial H(w, w^0, \mu)}{\partial (w, \mu)} = \begin{pmatrix} \Pi & (1 - \mu)\nabla \tilde{g}(\theta, \mu) & \nabla h(\theta) & \Lambda \\ Y\nabla \tilde{g}(\theta, \mu)^T & \text{diag}(\tilde{g}(\theta, \mu)) & 0 & \Xi \\ \nabla h(\theta)^T & 0 & 0 & -h(\theta^0) \end{pmatrix},$$

where  $\Pi = (1 - \mu)(\nabla^2 F(\theta) + \sum_{i=1}^{m+3} y_i \nabla^2 \tilde{g}_i(\theta, \mu)) + \sum_{j=1}^{2m-1} z_j \nabla^2 h_j(\theta) + \mu I$ ,  $\Lambda = -\nabla F(\theta) - \nabla g(\theta)y + \theta - \theta^0 + (1 - 2\mu)\xi$  and  $\Xi = -\sigma \mu^{\sigma-1} Y \tau + \eta$ , for the initial point  $w^0 = (\theta^0, y^0, z^0)$ , and by using  $z^0 = 0$ , we can obtain

$$DH(w^0, w^0, 1) = \begin{pmatrix} I & 0 & \nabla h(\theta^0) & \Pi^0 \\ Y^0 \nabla \tilde{g}(\theta^0, 1)^T & \text{diag}(\tilde{g}(\theta^0, 1)) & 0 & \Xi^0 \\ \nabla h(\theta^0)^T & 0 & 0 & -h(\theta^0) \end{pmatrix} = (\Upsilon_1, \Upsilon_2),$$

where  $\Upsilon_1 \in \mathbb{R}^{[4m+(m+3)+(2m-1)] \times [4m+(m+3)+(2m-1)]}$ ,  $\Upsilon_2 \in \mathbb{R}^{[4m+(m+3)+(2m-1)] \times 1}$ . The tangent vector  $v^{0T} = (v_1^{0T}, v_2^0)$  of  $\Gamma_{w^0}$  at the point  $(w^0, 1)$  satisfies

$$(\Upsilon_1, \Upsilon_2) \begin{pmatrix} v_1^0 \\ v_2^0 \end{pmatrix} = 0,$$

where  $v_1^0 \in \mathbb{R}^{[4m+(m+3)+(2m-1)] \times 1}$  and  $v_2^0 \in \mathbb{R}$ . By a simple computation, we can get  $v_1^0 = -\gamma_1^{-1} \gamma_2 v_2^0$ . Then we can obtain the following determinant:

$$\begin{aligned} & \left| \begin{matrix} DH(w^0, w^0, 1) \\ v_1^0 \\ v_2^0 \end{matrix} \right| \\ &= \left| \begin{matrix} \gamma_1 & \gamma_2 \\ v_1^0 & v_2^0 \end{matrix} \right| = \left| \begin{matrix} \gamma_1 & \gamma_2 \\ -\gamma_2^T \gamma_1^T & 1 \end{matrix} \right| v_2^0 \\ &= \left| \begin{matrix} \gamma_1 & \gamma_2 \\ 0 & \gamma_2^T \gamma_1^{-T} \gamma_1^{-1} \gamma_2 + 1 \end{matrix} \right| v_2^0 = |\gamma_1| v_2^0 (\gamma_2^T \gamma_1^{-T} \gamma_1^{-1} \gamma_2 + 1) \\ &= \left| \begin{matrix} I & 0 & \nabla h(\theta^0) \\ Y^0 \nabla \tilde{g}(\theta^0, 1)^T & \text{diag}(\tilde{g}(\theta^0, 1)) & 0 \\ \nabla h(\theta^0)^T & 0 & 0 \end{matrix} \right| v_2^0 (\gamma_2^T \gamma_1^{-T} \gamma_1^{-1} \gamma_2 + 1) \\ &= \left| \begin{matrix} I & 0 & \nabla h(\theta^0) \\ Y^0 \nabla \tilde{g}(\theta^0, 1)^T & \text{diag}(\tilde{g}(\theta^0, 1)) & 0 \\ 0 & 0 & -\nabla h(\theta^0)^T \nabla h(\theta^0) \end{matrix} \right| v_2^0 (\gamma_2^T \gamma_1^{-T} \gamma_1^{-1} \gamma_2 + 1) \\ &= |-\nabla h(\theta^0)^T \nabla h(\theta^0)| \left| \begin{matrix} I & 0 \\ Y^0 \nabla \tilde{g}(\theta^0, 1)^T & \text{diag}(\tilde{g}(\theta^0, 1)) \end{matrix} \right| v_2^0 (\gamma_2^T \gamma_1^{-T} \gamma_1^{-1} \gamma_2 + 1) \\ &= (-1)^{2m-1} |\nabla h(\theta^0)^T \nabla h(\theta^0)| |\text{diag}(\tilde{g}(\theta^0, 1))| v_2^0 (\gamma_2^T \gamma_1^{-T} \gamma_1^{-1} \gamma_2 + 1). \end{aligned}$$

Since  $\tilde{g}(\theta^0, 1) < 0$ ,  $\gamma_2^T \gamma_1^{-T} \gamma_1^{-1} \gamma_2 + 1 > 0$ , and the last element  $v_2^0$  of the positive direction  $v^0$  should be negative, the sign of the determinant

$$\left| \begin{matrix} DH(w^0, w^0, 1) \\ v_1^0 \\ v_2^0 \end{matrix} \right| \text{ is } (-1)^{(m+3)+(2m-1)+1}.$$

The proof is complete. □

For completeness of this study, a detailed predictor–corrector algorithm is presented here.

**Algorithm 4.1**

*Step 1.* Initialization.

Given the risk averse utility  $v(\cdot)$ , the distribute density function  $f(x, a)$ , the minimum expected utility  $V_0$ , the number  $m$  of nodes and vector  $\tau \in \mathbb{R}_{++}^{m+3}$ , formulate  $F(\theta)$ ,  $g(\theta)$ ,  $h(\theta)$ , shifted constraint functions  $\tilde{g}(\theta, \mu)$ , the homotopy map  $H$  and its Jacobian. Given a randomly chosen vector  $\xi \in \mathbb{R}^{4m}$  and a positive vector  $\eta \in \mathbb{R}_{++}^{m+3}$ , set the accuracy parameters  $\varepsilon_1 \geq \varepsilon_2 > 0$ , the initial point  $\theta^0 \in V \subset \Omega^0(1)$ , the initial steplength  $\lambda_0 > 0$ , the minimum steplength  $\lambda_{\min}$ , the maximum steplength  $\lambda_{\max}$ , the maximum number  $\bar{N}$  of the corrector steps, the threshold value  $\epsilon_\alpha$  for the angle between two neighboring predictor directions, the step contraction factors  $\epsilon_1, \epsilon_2$  and  $\epsilon_3$ , the step expansion factors  $\epsilon_4$  and  $\epsilon_5$ , and the threshold value  $0 < \epsilon_\mu < 1$  for starting the end game. Set  $k = 0$ .

*Step 2.* The first predictor step.

If  $k = 0$ , set  $\hat{\lambda} = \lambda_0, \varepsilon = \varepsilon_1$ ;

Let  $v^{-1} = (0, \dots, 0, -1)^T \in \mathbb{R}^{4m+(m+3)+(2m-1)+1}$  and compute the predictor step  $v$  by

$$\begin{pmatrix} DH(w^0, w^0, 1) \\ (v^{-1})^T v \end{pmatrix} = -v^{-1}$$

Set  $\|v^0\| = \frac{v}{\|v\|}$ . Determine the smallest nonnegative integer  $i$  such that

$$w^{1,0}, \mu_{1,0} = (w^0, \mu_0) + \epsilon_3^i \hat{\lambda} v^0 \in \Omega(\mu_{1,0}) \times \mathbb{R}_+^{m+3} \times \mathbb{R}^{2m-1} \times (0, 1),$$

set  $\hat{\lambda} = \epsilon_3^i \hat{\lambda}$ ;

Go to Step 3;

Else, perform the midway predictor step.

Let  $v^k = ((w^k, \mu_k) - (w^{k-1}, \mu_{k-1})) / \|(w^k, \mu_k) - (w^{k-1}, \mu_{k-1})\|$ , determine the smallest nonnegative integer  $i$  such that

$$(w^{k+1,0}, \mu_{k+1,0}) = (w^k, \mu_k) + \epsilon_3^i \hat{\lambda} v^k \in \Omega(\mu_{k+1,0}) \times \mathbb{R}_+^{m+3} \times \mathbb{R}^{2m-1} \times (0, 1);$$

Set  $\hat{\lambda} = \epsilon_3^i \hat{\lambda}$ ;

Go to Step 3.

*Step 3.* The corrector step.

Set  $j = 0$ . Repeat; Compute the Newton step  $\hat{v}$  by solving the following augmented system:

$$\begin{pmatrix} DH(w^{k+1,j}, w^0, \mu_{k+1,j}) \\ (v^k)^T \end{pmatrix} \hat{v} = \begin{pmatrix} -H(w^{k+1,j}, w^0, \mu_{k+1,j}) \\ 0 \end{pmatrix}$$

Determine the smallest nonnegative integer  $i$  such that

$$(w^{k+1,j+1}, \mu_{k+1,j+1}) = (w^{k+1}, \mu_{k+1}) + \epsilon_3^i \hat{\lambda} \hat{v} \in \Omega(\mu_{k+1,j+1}) \times \mathbb{R}_+^{m+3} \times \mathbb{R}^{2m-1} \times (0, 1).$$

If  $\|H(w^{k+1,j+1}, w^0, \mu_{k+1,j+1})\|_\infty \leq \|H(w^{k+1,j}, w^0, \mu_{k+1,j})\|_\infty$

Set  $j = j + 1$ .

Else

Set  $j = \bar{N}$ ,  $(w^{k+1,j}, \mu_{k+1,j}) = (w^{k+1,0}, \mu_{k+1,0})$ , until

$$\|H(w^{k+1,j}, w^0, \mu_{k+1,j})\|_\infty \leq \epsilon \quad \text{or} \quad j = \bar{N};$$

Go to Step 4.

*Step 4.* The steplength strategy.

If  $j = \bar{N}$ ,  $\|H(w^{k+1,j}, w^0, \mu_{k+1,j})\|_\infty > \epsilon$ ,

Set  $\hat{\lambda} = \max\{\lambda_{\min}, \epsilon_2 \hat{\lambda}\}$  and  $(w^{k+1,0}, \mu_{k+1,0}) = (w^k, \mu_k) + \hat{\lambda} v^k$ ;

Go to Step 3;

Else

Set  $(w^{k+1}, \mu_{k+1}) = (w^{k+1,j}, \mu_{k+1,j})$ , adjust the steplength  $\hat{\lambda}$  as follows:

If  $(v^k)^T v^{k-1} < \epsilon_\alpha$ , set  $\hat{\lambda} = \max\{\lambda_{\min}, \epsilon_1 \hat{\lambda}\}$ ;

If  $j > 3$ , set  $\hat{\lambda} = \max\{\lambda_{\min}, \epsilon_2 \hat{\lambda}\}$ ;

If  $j = 2$ , set  $\hat{\lambda} = \min\{\lambda_{\max}, \epsilon_4 \hat{\lambda}\}$ ;

If  $j < 2$ , set  $\hat{\lambda} = \min\{\lambda_{\max}, \epsilon_5 \hat{\lambda}\}$ ;

If  $\mu_{k+1} < \epsilon_\mu$ , go to Step 5.

If  $\|H(w^{k+1}, w^0, 0)\|_\infty \leq \epsilon_2$  and  $w^{k+1}$  is feasible, then stop.  $\bar{w} = w^{k+1}$  is the computed solution to the KKT system (5). Thus,  $\bar{a} = a^k$  is the optimal action and  $s_i^{k+1}(x) = p_i^{k+1}x^2 + q_i^{k+1}x + r_i^{k+1}$ ,  $i = 1, 2, \dots, m$  is the quadratic spline contractual function.

Set  $\epsilon = \min\{\mu_{k+1}, \epsilon_1\}$ ,  $k = k + 1$ .

*Step 5.* The end game.

Set  $j = 0$ ,  $w^{k+1,0} = w^{k+1}$ ;

Repeat;

Compute the Newton step  $v_{\text{end}}$  by solving the equation

$$\frac{\partial H}{\partial w}(w^{k+1,0}, w^0, 0)v_{\text{end}} = -H(w^{k+1,j}, w^0, 0),$$

$$\text{set } w^{k+1,j+1} = w^{k+1,j} + v_{\text{end}},$$

$$j = j + 1;$$

Until,  $\|H(w^{k+1,j}, w^0, 0)\|_\infty \leq \epsilon_2$  or  $j = \bar{N}$ ;

If  $\|H(w^{k+1,j}, w^0, 0)\|_\infty \leq \epsilon_2$  and  $w^{k+1,j}$  is feasible, then stop.  $\bar{w} = w^{k+1,j}$  is the computed solution to the KKT system (5). Thus,  $\bar{a} = a^{k+1,j}$  is the optimal action and  $s_i^{k+1,j}(x) = p_i^{k+1,j}x^2 + q_i^{k+1,j}x + r_i^{k+1,j}$ ,  $i = 1, 2, \dots, m$  is the quadratic spline contractual function.

Else, set  $\epsilon_t = 0.1\epsilon_\mu$ .

## 5 Conclusions

To design a quadratic spline contractual function in the case of discretely unknown nodes, in this paper we constructed a modified constraint shifting homotopy algorithm for solving principal–agent problems. Then the existence of a globally convergent solution to KKT systems for the principal–agent problem with spline contractual function is proved under mild conditions. The proposed algorithm only requires that any initial point is in the shifted feasible set but not necessarily in the original feasible set. Our contribution in the paper is theoretical and the numerical simulations on performance for the CSCH can be implemented as in Refs. [21, 22]. For more discussions about algorithms for solving optimization problems, see, e.g., [36–38].

### Acknowledgements

The authors would like to thank the editor for help.

### Funding

The investigation which made possible this paper was supported by the Jilin Provincial Science and Technology Development Plan funded project (Grant No. 20170520050JH, 20160520105JH), and the Education Department Foundation of Jilin province (Grant No. JJKH20190742SK).

### Abbreviations

KKT, Karush–Kuhn–Tucker; MLRC, monotone likelihood ratio condition; CDfC, convexity of the distribution function condition; LP, linear program/programming; MPEC, mathematical program with equilibrium constraints.

### Availability of data and materials

Not applicable.

### Competing interests

The authors declare that they have no competing interests.

### Authors' contributions

All authors contributed equally to the writing of this paper. All authors read and approved the final manuscript.

### Author details

<sup>1</sup>School of Economics, Liaoning University, Liaoning, P.R. China. <sup>2</sup>School of Mathematical Sciences, Tianjin Polytechnic University, Tianjin, P.R. China. <sup>3</sup>Center for General Education, China Medical University, Taichung, Taiwan. <sup>4</sup>Gh. Mihoc-C. Iacob Institute of Mathematical Statistics and Applied Mathematics, Romanian Academy, Bucharest, Romania. <sup>5</sup>Department of Mathematics and Informatics, University "Politehnica" of Bucharest, Bucharest, Romania.

### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 16 August 2019 Accepted: 29 October 2019 Published online: 08 November 2019

### References

1. Mirrlees, J.: The theory of moral hazard and unobservable behavior—part I. Mimeo, Nuffield College, Oxford (1975)
2. Mirrlees, J.: The theory of moral hazard and unobservable behavior—part I. *Rev. Econ. Stud.* **66**(1), 3–21 (1999)
3. Mirrlees, J.: The implication of moral hazard for optimal insurance. In: Seminar Given at Conference Held in Honor of Karl Borch, Bergen, Norway, mimeo (1979)
4. Rogerson, W.P.: The first-order approach to principal–agent problems. *Economica* **53**(6), 1357–1367 (1985)
5. Jewitt, I.: Justifying the first-order approach to principal–agent problems. *Economica* **56**(5), 1177–1190 (1988)
6. Grossman, S.J., Hart, O.D.: An analysis of the principal–agent problem. *Econometrica* **51**(1), 7–45 (1983)
7. Sinclair-Desgagne, B.: The first-order approach to multi-signal principal–agent problems. *Econometrica* **62**(2), 459–465 (1994)
8. Alvi, E.: First-order approach to principal–agent problems: a generalization. *Geneva Risk Insur. Rev.* **22**, 59–65 (1997)
9. LiCalzi, M., Spaeter, S.: Distributions for the first-order approach to principal–agent problems. *Econ. Theory* **21**, 167–173 (2003)
10. Conlon, J.R.: Two new conditions supporting the first-order approach to multisignal principal–agent problems. *Econometrica* **77**(1), 249–278 (2009)
11. Abrahm, A., Koehne, S., Pavoni, N.: On the first-order approach in principal–agent models with hidden borrowing and lending. *J. Econ. Theory* **146**(4), 1331–1361 (2011)
12. Fagart, M.-C., Fluet, C.: The first-order approach when the cost of effort is money. *J. Math. Econ.* **49**, 7–16 (2013)
13. Bertola, G., Koeniger, W.: On the validity of the first-order approach with moral hazard and hidden assets. *Econ. Lett.* **124**(3), 402–405 (2014)
14. Kirkegaard, R.: A unifying approach to incentive compatibility in moral hazard problems. *Theor. Econ.* **12**(1), 25–51 (2017)
15. Prescott, E.S.: Computing solutions to moral-hazard programs using the Dantzig–Wolfe decomposition algorithm. *J. Econ. Dyn. Control* **28**, 777–800 (2004)
16. Su, C.-L., Judd, K.L.: Computation of moral-hazard problems. Mimeo, Kellogg School of Management, Northwestern University (October 30, 2007). <http://www.econ.sinica.edu.tw/upload/file/MPEC-MH.pdf>
17. Armstrong, C.S., Larcker, D.F., Su, C.L.: Endogenous selection and moral hazard in compensation contracts. *Oper. Res.* **58**(4), 1090–1106 (2010)
18. Cecchini, M., Ecker, J., Kupferschmid, M., Leitch, R.: Solving nonlinear principal–agent problems using bilevel programming. *Eur. J. Oper. Res.* **230**(2), 364–373 (2013)
19. Zhu, Z., Yu, B., Yang, L.: Globally convergent homotopy method for designing piecewise linear deterministic contractual function. *J. Ind. Manag. Optim.* **10**(3), 717–741 (2014)
20. Renner, P., Schmedders, K.: A polynomial optimization approach to principal–agent problems. *Econometrica* **83**(2), 729–769 (2015)
21. Zhu, Z., Yu, B.: Globally convergent homotopy algorithm for solving the KKT systems to the principal–agent bilevel programming. *Optim. Methods Softw.* **32**(1), 69–85 (2017)
22. Zhu, Z., Yu, B.: A modified homotopy method for solving the principal–agent bilevel programming problem. *Comput. Appl. Math.* **37**, 541–566 (2018)
23. Feng, G.C., Yu, B.: Combined homotopy interior point method for nonlinear programming problems. In: Fujita, H., Yamaguti, M. (eds.) *Adv. Numer. Math.: Proc. 2nd Japan–China Seminar Numer. Math. Lecture Notes in Numerical and Applied Analysis*, vol. 14, pp. 9–16. Kinokuniya, Tokyo (1995)
24. Lin, Z.H., Li, Y., Yu, B.: A combined homotopy interior point method for general nonlinear programming problems. *Appl. Math. Comput.* **80**, 209–224 (1996)
25. Lin, Z.H., Yu, B., Feng, G.C.: A combined homotopy interior point method for convex programming problem. *Appl. Math. Comput.* **84**, 193–211 (1997)
26. Feng, G.C., Lin, Z.H., Yu, B.: Existence of an interior pathway to a Karush–Kuhn–Tucker point of a nonconvex programming problem. *Nonlinear Anal.* **32**, 761–768 (1998)
27. Yu, B., Lin, Z.H.: Homotopy method for a class of nonconvex Brouwer fixed point problems. *Appl. Math. Comput.* **74**, 65–77 (1996)
28. Lin, Z.H., Li, Y.: Homotopy method for solving variational inequalities. *J. Optim. Theory Appl.* **100**, 207–218 (1999)
29. Xu, Q., Yu, B., Feng, G.C.: Homotopy method for solving variational inequalities in unbounded sets. *J. Glob. Optim.* **31**, 121–131 (2005)
30. Xu, Q., Yu, B.: Solving the Karush–Kuhn–Tucker system of a nonconvex programming problem on an unbounded set. *Nonlinear Anal.* **70**, 757–763 (2009)
31. Yu, B., Shang, Y.F.: Boundary moving combined homotopy method for nonconvex nonlinear programming. *J. Math. Res. Expo.* **26**, 831–834 (2006)
32. Shang, Y.F., Yu, B.: Boundary moving combined homotopy method for convex nonlinear programming and its convergence. *J. Jinlin Univ.* **44**, 357–361 (2006)
33. Yang, L., Yu, B., Xu, Q.: A constraint shifting homotopy method for general non-linear programming. *Optimization* **63**(4), 585–600 (2014)

34. Chow, S.N., Mallet-Paret, J., Yorke, J.A.: Finding zeroes of maps: homotopy me that are constructive with probability one. *Math. Comput.* **32**, 887–899 (1978)
35. Naber, G.L.: *Topological Method in Euclidean Space*. Cambridge University Press, London (1980)
36. Allgower, E.L., Georg, K.: *Introduction to Numerical Continuation Methods*. Classics in Applied Mathematics, vol. 45. SIAM, Philadelphia (2003)
37. Watson, L.T., Billups, S.C., Morgan, A.P.: Algorithm 652 hompack: a suite of codes for globally convergent homotopy algorithms. *ACM Trans. Math. Softw.* **13**, 281–310 (1987)
38. Yao, Y., Postolache, M., Zhu, Z.: Gradient methods with selection technique for the multiple-sets split feasibility problem. *Optimization*. <https://doi.org/10.1080/02331934.2019.1602772>

**Submit your manuscript to a SpringerOpen<sup>®</sup> journal and benefit from:**

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

---

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)

---