



Convergence, stability analysis, and solvers for approximating sublinear positone and semipositone boundary value problems using finite difference methods

Thomas Lewis^{a,1,*}, Quinn Morris^b, Yi Zhang^a

^a Department of Mathematics and Statistics, The University of North Carolina at Greensboro, Greensboro, NC 27402, USA

^b Department of Mathematical Sciences, Appalachian State University, Boone, NC 28608, USA

ARTICLE INFO

Article history:

Received 5 June 2020

Received in revised form 6 July 2021

MSC:

65N06

65N12

65N22

Keywords:

Finite difference methods

Convergence

Semipositone

Sublinear

Sub- and supersolutions

Nonuniqueness

ABSTRACT

Positone and semipositone boundary value problems are semilinear elliptic partial differential equations (PDEs) that arise in reaction–diffusion models in mathematical biology and the theory of nonlinear heat generation. Under certain conditions, the problems may have multiple positive solutions or even nonexistence of a positive solution. We develop analytic techniques for proving admissibility, stability, and convergence results for simple finite difference approximations of positive solutions to sublinear problems. We also develop guaranteed solvers that can detect nonuniqueness for positone problems and nonexistence for semipositone problems. The admissibility and stability results are based on adapting the method of sub- and supersolutions typically used to analyze the underlying PDEs. The new convergence analysis technique directly shows that all pointwise limits of finite difference approximations are solutions to the boundary value problem eliminating the possibility of false algebraic solutions plaguing the convergence of the methods. Most known approximation methods for positone and semipositone boundary value problems rely upon shooting techniques; hence, they are restricted to one-dimensional problems and/or radial solutions. The results in this paper will serve as a foundation for approximating positone and semipositone boundary value problems in higher dimensions and on more general domains using simple approximation methods. Numerical tests for known applied problems with multiple positive solutions are provided. The tests focus on approximating certain positive solutions as well as generating discrete bifurcation curves that support the known existence and uniqueness results for the PDE problem.

© 2021 Elsevier B.V. All rights reserved.

1. Introduction

We consider approximating positive solutions to the nonlinear elliptic partial differential equation (PDE)

$$-\Delta u = \lambda f(u) \quad \text{in } \Omega, \quad (1a)$$

$$u \geq 0 \quad \text{in } \Omega, \quad (1b)$$

* Corresponding author.

E-mail addresses: tlewis3@uncg.edu (T. Lewis), morrisqa@appstate.edu (Q. Morris), y_zhang7@uncg.edu (Y. Zhang).

¹ The work of this author was partially supported by the National Science Foundation (USA) grant DMS-2111059.

$$u = 0 \quad \text{on } \partial\Omega, \quad (1c)$$

where $\lambda > 0$ is a constant; Ω is an open, bounded, convex domain; and $f : \mathbb{R} \rightarrow \mathbb{R}$ is a continuous function. Throughout the paper we will assume that f is Lipschitz. When needed, we will specify additional assumptions on f corresponding to positivity, monotonicity, and sublinearity. i.e., we will assume that f satisfies

(H1) $f(w) > 0$ for all $w \geq c$ for some $c \geq 0$,

(H2) $f(0)$ is bounded,

(H3) f is nondecreasing,

(H4) $\lim_{w \rightarrow \infty} \frac{f(w)}{w} = 0$,

(H5) $f(w) = f(0)$ for all $w \leq 0$.

Problems satisfying conditions (H1)–(H3) with $f(0) > 0$ are referred to in the PDE literature as positone problems, whereas problems satisfying conditions (H1)–(H3) with $f(0) \leq 0$ are referred to as semipositone problems. Problems satisfying condition (H4) are referred to as sublinear at infinity. The assumption (H5) is used to extend f to be defined over $(-\infty, 0)$ with a constant extension and is helpful in our study of semipositone problems where sign-changing and negative solutions are expected. We refer the interested reader to [1] and [2] for a brief introduction to both positone and semipositone problems, respectively. We note that the results in this paper can be trivially extended to the nonautonomous problem $-\Delta u = \lambda h(x)f(u)$ for $h(x) \geq \eta > 0$ a bounded weight function.

Problems which are both semipositone and sublinear at infinity arise in mathematical biology and the theory of nonlinear heat generation, as will be discussed in Section 6. There is a long history of the study of positone and semipositone problems which are sublinear at infinity when Ω has a smooth boundary. It is known that when f is Lipschitz and positone, (1) has a positive solution for all $\lambda > 0$, and that if additionally f is concave, then the positive solution is unique (see [1]). Uniqueness when $\lambda \approx 0$ and $\lambda \gg 1$ has also been established for certain domains Ω and under suitable conditions on f (see [3,4]). When the problem is semipositone, (1) may not have a positive solution for λ small (see [2]). The mix of nonuniqueness and nonexistence makes approximating the positive solutions to such problems difficult. Even if a method is known to converge, choosing a good initial guess for a generic solver to find a positive solution can be nearly impossible since the negative or sign-changing solutions may be stable when viewing the solutions to the elliptic problem as steady-states of the corresponding parabolic PDE.

In the theoretical study of such PDEs, the method of sub- and supersolutions, which first appeared in [5], is often utilized. In short, the method constructs a sequence of functions from solutions to a linear PDE associated to the original PDE, which converge monotonically to a solution to the original nonlinear problem. We refer the interested reader to [6] for an overview of such methods applied to theoretical PDE problems. One obvious advantage of the method in finding positive solutions is that, given a positive subsolution, you are guaranteed positivity of the solution found through the usual monotone iteration scheme. Additionally, the method provides existence, comparison, and uniqueness/nonuniqueness results from the sequence of iterates. Analogous monotone iterative schemes have been applied to discrete versions of elliptic problems (see, for example, the work of C.V. Pao in [7–9]) yielding natural admissibility and stability results for approximation methods.

The goal of the paper is to provide a complete analysis for simple finite difference (FD) approximations of (1) utilizing a monotone iterative scheme and a novel convergence proof technique adopted from the numerical analysis of monotone finite difference methods for fully nonlinear elliptic PDEs. We emphasize simple methods so that (1) the methods can easily be applied, and (2) we can better understand the new analytic techniques before extending them to more complicated discretizations. The two primary difficulties that must be addressed in the analysis are the potential nonuniqueness of positive solutions to (1) and the potential nonexistence of positive solutions in the semilinear case. Thus, the method must be shown to be flexible enough to capture multiple solutions yet also account for the possibility of nonexistence. Even when Ω is smooth, (1) may have multiple positive solutions when f is not concave. Thus, we expect the nonlinear algebraic system of equations resulting from the discretization to have multiple solutions, some of which may be negative or sign-changing. Even worse, the system may have algebraic artifacts that do not correspond to a PDE solution. An example of this phenomena is when using the standard FD discretization for the Monge–Ampère equation in two dimensions there are $2^{(N-1)^2}$ solutions on an $N \times N$ grid despite the fact that the PDE problem typically has only two natural solutions with only one being a viscosity solution (see [10]). To ensure we are not capturing false solutions, we directly show that all convergent sequences of FD approximations converge to PDE solutions of (1), where (1b) is satisfied as long as we develop a way to choose appropriate sequences of positive FD approximations. Thus, we ensure that the number of discrete solutions matches the number of PDE solutions in limit. The convergence proof is highly flexible based on its minimal assumptions and can be applied for much more general nonlinearities. As such, the convergence proof techniques in this paper are expected to have applications to approximating positive solutions to a much broader class of nonlinear reaction–diffusion equations.

We also provide tools for proving admissibility and stability and develop an appropriate solver for finding positive solutions. The admissibility and stability proofs naturally extend PDE sub- and supersolution techniques to discrete problems. The proposed solver monotonically converges to a minimal positive solution when starting with a positive subsolution and monotonically converges to a maximal solution when starting with a positive supersolution. For

semipositone problems, the solver can be used to detect whether the associated nonlinear system of equations has a positive solution. For positone problems, the solver can be used to detect nonuniqueness of positive solutions. The solver can also be used to generate $(\lambda, \|\cdot\|_{\max})$ -bifurcation curves for counting the number of positive solutions. Lastly, we provide a simple efficient technique for generating both discrete sub- and supersolutions ensuring the existence of appropriate initial guesses.

A number of recent papers concerning the theoretical PDE study of existence and uniqueness of positive solutions to nonlinear elliptic problems include computational examples that use either quadrature methods or shooting methods (for example, [11,12]). Both methods are limited to one-dimensional problems. However, they are natural extensions of the theoretical techniques used to prove existence and uniqueness results in certain cases and they are relatively easy to implement. A last class of methods that is popular in the PDE community is the class of spectral methods proposed by Neuberger (see [13]) which have been highly optimized for speed and applied to various classes of nonlinear reaction–diffusion equations. However, the methods are heuristic and convergence results are still open.

The quadrature method (see [14]) transforms the differential equation into a nonlinear algebraic equation that relates λ and $\|u\|_{\infty}$. Thus, there is no discretization procedure and only a numerical nonlinear solver is needed. Unfortunately, the transformation requires symmetry of the solution that does not hold for nonautonomous problems such as $-u'' = \lambda h(x)f(u)$ for $h(x) > 0$ a given weight function. Therefore, shooting methods have become a method of choice, where shooting methods exhaustively search for all solutions of the given problem using a bisection method and an ordinary differential equation solver for a corresponding initial value problem. Typically the shooting parameter is the Neumann condition at one of the boundary points. The primary drawback of shooting methods is that they are limited to one-dimensional problems. However, when Ω is a ball or ring-shaped domain, then all solutions are known to be radially symmetric (see [15]), and thus can be recovered through an appropriate transformation to a one-dimensional problem. When Ω is not radially symmetric, however, approximations of nonradial solutions require more robust methods such as the FD method introduced in this paper. See Section 6.1 where we provide a non-radially symmetric example when $\Omega = (0, 1)^2$.

The standard numerical analysis approach for semilinear problems is outlined in [16]. The idea is to locally analyze the problem by showing a discrete approximation exists in a neighborhood of a given PDE solution. An immediate benefit of the approach is that it yields rates of convergence when assuming adequate regularity of the local PDE solution since it naturally extends techniques for linear problems. Unfortunately, by taking a local approach, the method does not address whether the scheme possesses algebraic solutions that cannot locally be mapped to a PDE solution leaving the question of algebraic artifacts open. A second benefit of the standard approach is that it naturally yields methods for approximating simple bifurcation points on the trivial branch as well as error estimates for bifurcation equations. Lastly, the approach has been used for the more general problem $-\Delta u + g(\lambda, u) = 0$ with less restrictive assumptions for the reaction term $g(\lambda, u)$ than we assume for f in (1). A drawback of the standard approach is that it relies on the difficult task of showing an approximation exists in a neighborhood of the PDE solution when compared to our approach that relies only on the simple construction of a sub- and supersolution. Other drawbacks include the fact that the standard approach does not directly address the need for positive solutions and the fact that the techniques considered have mostly gone unnoticed in the PDE community where the above alternative methods are used to approximate solutions and generate bifurcation curves for reaction–diffusion problems with multiple solutions.

More numerical methods and algorithms for finding multiple solutions of nonlinear PDEs can be found in [7,13,17–29] and the references therein. The methods in [7,17,18] are all variational-type methods that emphasize the positivity of the sought after solution. Alternatively, the other methods use a traditional discretization of the PDE and then emphasize the positivity when solving the system. In general, finding a positive solution can be difficult or highly inefficient. A benefit of our approach is that we have a simple efficient mechanism for enforcing the positivity condition while automatically detecting either nonexistence or nonuniqueness.

The remainder of the paper is organized as follows. The simple FD method for approximating (1) is formulated in Section 2. Section 3 contains the main convergence result for the paper. We prove existence, stability, and uniqueness results for the scheme in Section 4, and we provide a fixed-point solver in Section 5 that guarantees convergence to a positive solution, tests for multiple solutions, detects nonexistence for semipositone problems, and helps in the generation of bifurcation curves. In Section 6 we introduce examples of sublinear positone and semipositone problems that have multiple solutions and perform numerical tests that demonstrate the robustness of our proposed methods. Concluding remarks can be found in Section 7.

2. A simple finite difference method

In this section we formulate a simple FD method for approximating positive solutions to (1). For transparency, we assume Ω is a d -rectangle for the number of dimensions $d \geq 1$, i.e., $\Omega = (a_1, b_1) \times (a_2, b_2) \times \cdots \times (a_d, b_d)$. In future works we plan to extend the analytic techniques in this paper to finite difference methods on radial domains and finite element methods on polygonal domains as a way to naturally allow for approximating problems posed on more general domains. We will only consider grids that are uniform in each coordinate x_i , $i = 1, 2, \dots, d$. Let N_i be a positive integer and $h_i = \frac{b_i - a_i}{N_i - 1}$ for $i = 1, 2, \dots, d$. Define $\mathbf{h} = (h_1, h_2, \dots, h_d) \in \mathbb{R}^d$, $h = \max_{i=1,2,\dots,d} h_i$, $h_* = \min_{i=1,2,\dots,d} h_i$, $N = \prod_{i=1}^d N_i$, and $\mathbb{N}_N^d = \{\alpha = (\alpha_1, \alpha_2, \dots, \alpha_d) \mid 1 \leq \alpha_i \leq N_i, i = 1, 2, \dots, d\}$. Then, $|\mathbb{N}_N^d| = N$. We partition Ω into $\prod_{i=1}^d (N_i - 1)$

sub- d -rectangles with grid points $\mathbf{x}_\alpha = (a_1 + (\alpha_1 - 1)h_1, a_2 + (\alpha_2 - 1)h_2, \dots, a_d + (\alpha_d - 1)h_d)$ for each multi-index $\alpha \in \mathbb{N}_N^d$. We call $\mathcal{T}_h = \{\mathbf{x}_\alpha\}_{\alpha \in \mathbb{N}_N^d}$ a grid (set of nodes) for $\bar{\Omega}$.

Let $\{\mathbf{e}_i\}_{i=1}^d$ denote the canonical basis vectors for \mathbb{R}^d . Define the (second order) central difference operators for approximating second order partial derivatives by

$$\delta_{x_i, h_i}^2 v(\mathbf{x}) \equiv \frac{v(\mathbf{x} + h_i \mathbf{e}_i) - 2v(\mathbf{x}) + v(\mathbf{x} - h_i \mathbf{e}_i)}{h_i^2} \tag{2}$$

for a function v defined on \mathbb{R}^d and

$$\delta_{x_i, h_i}^2 V_\alpha \equiv \frac{V_{\alpha+\mathbf{e}_i} - 2V_\alpha + V_{\alpha-\mathbf{e}_i}}{h_i^2},$$

for a grid function V defined on the grid $\mathcal{T}_h \cap \Omega$. Also define the (second order) central discrete Laplacian operator by

$$\Delta_h \equiv \sum_{i=1}^d \delta_{x_i, h_i}^2. \tag{3}$$

The FD method is defined as finding a grid function $U_\alpha : \mathcal{T}_h \rightarrow \mathbb{R}$ such that

$$-\Delta_h U_\alpha = \lambda f(U_\alpha) \quad \text{if } \mathbf{x}_\alpha \in \mathcal{T}_h \cap \Omega, \tag{4a}$$

$$U_\alpha \geq 0 \quad \text{if } \mathbf{x}_\alpha \in \mathcal{T}_h \cap \Omega, \tag{4b}$$

$$U_\alpha = 0 \quad \text{if } \mathbf{x}_\alpha \in \mathcal{T}_h \cap \partial\Omega, \tag{4c}$$

where the main emphasis will be on how to identify solutions that satisfy (4b) or show that no such solutions exist. Formally, the method has a second order local truncation error.

Lastly we introduce the definitions for discrete sub- and supersolutions of (4). To this end, we call the grid function $\underline{U}_\alpha : \mathcal{T}_h \rightarrow \mathbb{R}$ a discrete subsolution if

$$-\Delta_h \underline{U}_\alpha \leq \lambda f(\underline{U}_\alpha) \quad \text{if } \mathbf{x}_\alpha \in \mathcal{T}_h \cap \Omega, \tag{5a}$$

$$\underline{U}_\alpha = 0 \quad \text{if } \mathbf{x}_\alpha \in \mathcal{T}_h \cap \partial\Omega, \tag{5b}$$

and we call the grid function $\bar{U}_\alpha : \mathcal{T}_h \rightarrow \mathbb{R}$ a discrete supersolution if

$$-\Delta_h \bar{U}_\alpha \geq \lambda f(\bar{U}_\alpha) \quad \text{if } \mathbf{x}_\alpha \in \mathcal{T}_h \cap \Omega, \tag{6a}$$

$$\bar{U}_\alpha \geq 0 \quad \text{if } \mathbf{x}_\alpha \in \mathcal{T}_h \cap \partial\Omega. \tag{6b}$$

Furthermore, \underline{U} is a positive discrete subsolution if $\underline{U}_\alpha \geq 0$ for all $\mathbf{x}_\alpha \in \mathcal{T}_h \cap \Omega$, and \bar{U} is a positive discrete supersolution if $\bar{U}_\alpha \geq 0$ for all $\mathbf{x}_\alpha \in \mathcal{T}_h \cap \Omega$. We will always be able to construct positive discrete supersolutions which can be used to detect the existence of a maximal (positive) solution to (4). For positive problems, we will construct positive discrete subsolutions which can be used to find the minimal (positive) solution to (4). In the admissibility analysis, we will construct uniformly bounded discrete sub- and supersolutions \underline{U} and \bar{U} . We will then show the problem (4a) and (4c) always has a solution uniformly bounded by \underline{U} and \bar{U} yielding a preliminary admissibility and stability result. We then consider how to find positive solutions that satisfy (4b).

3. Convergence

In this section we prove that any convergent sequence formed by solutions to the FD method (4) converges to a solution of the boundary value problem (1). Since it is unknown which solution the scheme will converge to, we completely avoid the issue of how to link a sequence of discrete solutions to an appropriate PDE solution when analyzing the error. Standard convergence analysis techniques for such problems can be found in [16] which considers the more abstract problem of approximating an infinite dimensional problem $F(\lambda, u) = 0$ with a discrete problem $F_h(\lambda, u_h) = 0$. Notationally, F_h can correspond to either the FD method or the finite element method. The idea is to locally analyze the problem by first establishing error estimates for linear problems and then using linearizations to produce error estimates for the nonlinear problem. In order to apply the techniques, one must first ensure a unique discrete solution exists in a neighborhood of the PDE solution.

To avoid the issues caused by the potential nonuniqueness of solutions to (1), we give a direct proof that shows any pointwise limit formed by solutions to (4) is a solution to (1). To avoid regularity considerations and to further extend the applicability of the proof, we will show that all limits are in fact viscosity solutions to (1). Thus, all of the limit functions satisfy the boundary value problem in a weaker sense. The result follows since viscosity solutions to (1) coincide with classical solutions to (1) when such smooth solutions exist. The convergence proof is a natural extension of the proof of Barles–Souganidis in [30] for approximating viscosity solutions of fully nonlinear boundary value problems using

monotone FD methods since the operator $-\Delta_{\mathbf{h}}$ is monotone. The extra details in the proof are due to the fact that the boundary value problem (1) does not satisfy a comparison principle. Thus, the existence of a continuous limiting function needs to be directly verified.

We begin by recording the definition of a viscosity solution to (1a) and (1c). Since the solutions to (1a) and (1c) are classical, we can assume the boundary condition is satisfied pointwise instead of in the viscosity sense. We will ensure (1b) is satisfied in Section 4. See [31] for more details about viscosity solutions to nonlinear elliptic problems.

Definition 3.1. (i) A continuous function $u : \overline{\Omega} \rightarrow \mathbb{R}$ with $u \leq 0$ on $\partial\Omega$ is called a *viscosity subsolution* of (1a) and (1c) if $\forall \varphi \in C^2(\overline{\Omega})$, when $u - \varphi$ has a local maximum at $\mathbf{x}_0 \in \Omega$, there holds

$$-\Delta\varphi(\mathbf{x}_0) - \lambda f(u(\mathbf{x}_0)) \leq 0.$$

(ii) A continuous function $u : \overline{\Omega} \rightarrow \mathbb{R}$ with $u \geq 0$ on $\partial\Omega$ is called a *viscosity supersolution* of (1a) and (1c) if $\forall \varphi \in C^2(\overline{\Omega})$, when $u - \varphi$ has a local minimum at $\mathbf{x}_0 \in \Omega$, there holds

$$-\Delta\varphi(\mathbf{x}_0) - \lambda f(u(\mathbf{x}_0)) \geq 0.$$

(iii) A continuous function $u : \overline{\Omega} \rightarrow \mathbb{R}$ is called a *viscosity solution* of (1a) and (1c) if u is both a viscosity subsolution and a viscosity supersolution of (1a) and (1c).

We next define a piecewise constant extension $u_{\mathbf{h}}$ for a given grid function $U \in S(\mathcal{T}_{\mathbf{h}})$ to be used in the statement of the convergence theorem. Let $\alpha \in \mathbb{N}_j$, and define B_α by

$$B_\alpha \equiv \prod_{i=1,2,\dots,d} \left[\mathbf{x}_\alpha - \frac{h_i}{2} \mathbf{e}_i, \mathbf{x}_\alpha + \frac{h_i}{2} \mathbf{e}_i \right],$$

where $\{\mathbf{e}_i\}_{i=1}^d$ denotes the canonical basis for \mathbb{R}^d . We define the piecewise constant extension $u_{\mathbf{h}}$ of a grid function U by

$$u_{\mathbf{h}}(\mathbf{x}) \equiv U_\alpha, \quad \mathbf{x} \in B_\alpha \tag{7}$$

for all $\mathbf{x} \in \Omega' \equiv \cup_{\alpha \in \mathbb{N}_j} B_\alpha \supset \overline{\Omega}$.

Using the uniform ℓ^∞ stability of the FD scheme (4a) and (4c) as will be verified below in Section 4 as well as the definition of the scheme itself, we guarantee the existence of a continuous pointwise limit function v in Lemma 3.1. Using a similar approach, we show any pointwise limit function is continuous in Corollary 3.1. The question addressed in Theorem 3.1 is whether the pointwise limits satisfy the original boundary value problem (1a) and (1c) or not. The result shows that any such limiting function v is in fact a viscosity solution to (1a) and (1c). If the sequence contains only positive solutions that also satisfy (4b), then we can ensure the limit is nonnegative. Thus, the FD method does not produce approximations that converge to false solutions, and it can be used to ensure the resulting PDE solution is nonnegative. Depending on the sequence, we can strengthen the bound to show the limit is strictly positive.

Lemma 3.1. Choose a mesh $\mathcal{T}_{\mathbf{h}}$, and let $\mathbf{h}_k = \frac{1}{k} \mathbf{h}$ for all $k \geq 1$. Let $U^{(k)} \in S(\mathcal{T}_{\mathbf{h}_k})$ be a solution to the FD scheme defined by (4a) and (4c). Let $u_{\mathbf{h}_k}$ be the piecewise constant extension of $U^{(k)}$ defined by (7). Then, there exist a continuous function $v : \overline{\Omega} \rightarrow \mathbb{R}$ and a subsequence $u_{\mathbf{h}_{k^*}}$ such that $u_{\mathbf{h}_{k^*}} \rightarrow v$ pointwise with $v = 0$ on $\partial\Omega$ provided the underlying scheme is ℓ^∞ stable.

Proof. For ease of presentation, all convergent sequences are understood as subsequences. Let $U^{(k)}$ be a sequence of solutions to (4a) and (4c), and let $u_{\mathbf{h}_k}$ be the corresponding sequence of piecewise constant functions. By the ℓ^∞ stability of the scheme, there exists a constant $C > 0$ independent of \mathbf{h}_k such that $\|u_{\mathbf{h}_k}\|_{L^\infty(\Omega)} \leq C$ for all k . Thus, there exist upper and lower semicontinuous functions \bar{v} and \underline{v} defined by

$$\bar{v}(\mathbf{x}) \equiv \limsup_{\substack{\xi \rightarrow \mathbf{x}, k \rightarrow \infty \\ \xi \in \overline{\Omega}}} u_{\mathbf{h}_k}(\xi), \quad \underline{v}(\mathbf{x}) \equiv \liminf_{\substack{\xi \rightarrow \mathbf{x}, k \rightarrow \infty \\ \xi \in \overline{\Omega}}} u_{\mathbf{h}_k}(\xi). \tag{8}$$

We show both \bar{v} and \underline{v} are continuous over $\overline{\Omega}$ and $\bar{v} = \underline{v} = 0$ over $\partial\Omega$. Choosing $v = \bar{v}$ or $v = \underline{v}$ and using the fact that the construction in (8) is analogous to the construction in [30], we can conclude that there exists a subsequence $u_{\mathbf{h}_k}$ that converges locally uniformly to v , and the result follows.

Suppose \bar{v} is not continuous at $\mathbf{x}_0 \in \Omega$. By the upper semi-continuity of \bar{v} , there exist a sequence $\{\mathbf{y}_k\} \subset \Omega$ and a direction x_i such that $\mathbf{y}_k \rightarrow \mathbf{x}_0$, $\bar{v}(\mathbf{y}_k) \rightarrow \bar{v}(\mathbf{x}_0)$, and $\bar{v}(\mathbf{y}_k)$ has a discontinuity in the x_i or $-x_i$ direction for all k . Choose $\epsilon > 0$ and $\sigma > 0$. Then we can assume there exist constants $c > 0$ and σ_k such that $0 < |\sigma_k| < \sigma$ and, for all k sufficiently large, there holds $|\bar{v}(\mathbf{y}_k) - \bar{v}(\mathbf{x}_0)| < \epsilon$ and $\bar{v}(\mathbf{x}_0) \geq \bar{v}(\mathbf{y}_k + \sigma_k \mathbf{e}_i) + c$. Also, there exists $\tau > 0$ such that $\bar{v}(\mathbf{x}_0) - \bar{v}(\mathbf{y}_k \pm \tau' \mathbf{e}_j) \geq -\epsilon$ for all $0 < \tau' < \tau$ for all $j \in \{1, 2, \dots, d\}$.

Let $h_j^{(k)}$ be the j th component of \mathbf{h}_k for all k . By the definition of \bar{v} in (8), it follows that there exist sequences $\mathbf{x}_k \rightarrow \mathbf{x}_0$ and $\mathbf{h}_k \rightarrow \mathbf{0}$, an index j^* , and a constant $c^* > 0$ such that

$$\mathbf{x}_k \in \mathcal{T}_{\mathbf{h}_k} \cap \Omega, \quad \lim_{k \rightarrow \infty} u_{\mathbf{h}_k}(\mathbf{x}_k) = \bar{v}(\mathbf{x}_0),$$

and

$$\lim_{k \rightarrow \infty} \left[u_{\mathbf{h}_k}(\mathbf{x}_k + h_{i^*}^{(k)} \mathbf{e}_{i^*}) - u_{\mathbf{h}_k}(\mathbf{x}_k) \right] = -c^*, \quad \lim_{k \rightarrow \infty} \left[u_{\mathbf{h}_k}(\mathbf{x}_k) - u_{\mathbf{h}_k}(\mathbf{x}_k - h_{i^*}^{(k)} \mathbf{e}_{i^*}) \right] \geq 0$$

or

$$\lim_{k \rightarrow \infty} \left[u_{\mathbf{h}_k}(\mathbf{x}_k + h_{i^*}^{(k)} \mathbf{e}_{i^*}) - u_{\mathbf{h}_k}(\mathbf{x}_k) \right] \leq 0, \quad \lim_{k \rightarrow \infty} \left[u_{\mathbf{h}_k}(\mathbf{x}_k) - u_{\mathbf{h}_k}(\mathbf{x}_k - h_{i^*}^{(k)} \mathbf{e}_{i^*}) \right] = c^*.$$

Furthermore,

$$\lim_{k \rightarrow \infty} \left[u_{\mathbf{h}_k}(\mathbf{x}_k + h_j^{(k)} \mathbf{e}_j) - u_{\mathbf{h}_k}(\mathbf{x}_k) \right] \leq 0, \quad \lim_{k \rightarrow \infty} \left[u_{\mathbf{h}_k}(\mathbf{x}_k) - u_{\mathbf{h}_k}(\mathbf{x}_k - h_j^{(k)} \mathbf{e}_j) \right] \geq 0$$

for all $j = 1, 2, \dots, d$ with $j \neq i^*$. Thus, by (2) and the quasiuniformity of \mathbf{h}_k , a subsequence can be chosen such that

$$\lim_{k \rightarrow \infty} \left(h_{i^*}^{(k)} \right)^2 \delta_{x_{i^*}, h_{i^*}^{(k)}}^2 u_{\mathbf{h}_k}(\mathbf{x}_k) \leq -c^*,$$

$$\lim_{k \rightarrow \infty} \left(h_{i^*}^{(k)} \right)^2 \delta_{x_j, h_j^{(k)}}^2 u_{\mathbf{h}_k}(\mathbf{x}_k) \leq 0$$

for all $j \neq i^*$. Hence,

$$\lim_{k \rightarrow \infty} \left(h_{i^*}^{(k)} \right)^2 \Delta_{\mathbf{h}_k} u_{\mathbf{h}_k}(\mathbf{x}_k) \leq -c^*,$$

and by the definition of the scheme, the continuity of f , and the boundedness of \bar{v} , there holds

$$0 = \lim_{k \rightarrow \infty} \left(h_{i^*}^{(k)} \right)^2 \left[-\Delta_{\mathbf{h}_k} u_{\mathbf{h}_k}(\mathbf{x}_k) - \lambda f(u_{\mathbf{h}_k}(\mathbf{x}_k)) \right] \geq c^* > 0,$$

a contradiction. Therefore, \bar{v} must be continuous at $\mathbf{x}_0 \in \Omega$.

Suppose \bar{v} is not continuous at $\mathbf{x}_0 \in \partial\Omega$. Since $u_{\mathbf{h}_k}(\mathbf{x}) = 0$ for all $\mathbf{x} \in \partial\Omega$, we must have $\bar{v}(\mathbf{x}_0) > 0$ with the limit coming from the interior of Ω . Thus, there exist sequences $\mathbf{x}_k \rightarrow \mathbf{x}_0$ and $\mathbf{h}_k \rightarrow \mathbf{0}$ such that $\mathbf{x}_k \in \mathcal{T}_{\mathbf{h}_k} \cap \Omega$ for all k and $\lim_{k \rightarrow \infty} u_{\mathbf{h}_k}(\mathbf{x}_k) = \bar{v}(\mathbf{x}_0)$. We can again choose the sequence to exploit the discontinuity at \mathbf{x}_0 . Since $\mathbf{x}_k \in \mathcal{T}_{\mathbf{h}_k} \cap \Omega$ for all k and $U^{(k)}$ solves (4), we have

$$0 = -\Delta_{\mathbf{h}_k} u_{\mathbf{h}_k}(\mathbf{x}_k) - \lambda f(u_{\mathbf{h}_k}(\mathbf{x}_k))$$

for all k . Therefore, we can use the same argument as in the case $\mathbf{x}_0 \in \Omega$ to arrive at a contradiction, and it follows that \bar{v} is continuous at $\mathbf{x}_0 \in \partial\Omega$ with $\bar{v}(\mathbf{x}_0) = 0$.

We can similarly show that \underline{v} must be continuous over $\bar{\Omega}$ with $\underline{v} = 0$ on $\partial\Omega$. The proof is complete. \square

Remark 3.1. The proof of Lemma 3.1 uses a construction that is inspired by the proof of Theorem 5.4 in [32] and similar to the contradiction derived in Subcase (iib) in the proof of Theorem 6.1 in [32]. Consequently, Lemma 3.1 has applications for extending the convergence results of [32] when the comparison principle does not hold.

Corollary 3.1. Choose a mesh $\mathcal{T}_{\mathbf{h}}$, and let $\mathbf{h}_k = \frac{1}{k} \mathbf{h}$ for all $k \geq 1$. Let $U^{(k)} \in S(\mathcal{T}_{\mathbf{h}_k})$ be a solution to the FD scheme defined by (4a) and (4c). Let $u_{\mathbf{h}_k}$ be the piecewise constant extension of $U^{(k)}$ defined by (7), and let $u_{\mathbf{h}_{k'}}'$ be a convergent subsequence such that $u_{\mathbf{h}_{k'}}' \rightarrow v : \bar{\Omega} \rightarrow \mathbb{R}$ pointwise. Then v is continuous and $v(\mathbf{x}) = 0$ over $\partial\Omega$ provided the underlying scheme is ℓ^∞ stable.

Proof. For ease of presentation, all convergent sequences are understood as subsequences. Define the corresponding upper and lower semicontinuous envelopes of v by

$$\bar{v}(\mathbf{x}) \equiv \limsup_{\xi \rightarrow \mathbf{x}} v(\xi), \quad \underline{v}(\mathbf{x}) \equiv \liminf_{\xi \rightarrow \mathbf{x}} v(\xi). \tag{9}$$

We show both \bar{v} and \underline{v} are continuous over $\bar{\Omega}$ from which it follows that $\bar{v} = \underline{v}$ with $\bar{v} = \underline{v} = 0$ over $\partial\Omega$. The result would follow since $\underline{v} \leq v \leq \bar{v}$.

Suppose \bar{v} is not continuous at $\mathbf{x}_0 \in \bar{\Omega}$. Then, there exist a sequence $\{\mathbf{y}_k\} \subset \bar{\Omega}$ and a direction x_i such that $\mathbf{y}_k \rightarrow \mathbf{x}_0$, $\bar{v}(\mathbf{y}_k) \rightarrow \bar{v}(\mathbf{x}_0)$, and $\bar{v}(\mathbf{y}_k)$ has a jump in the x_i or $-x_i$ direction for all k . Extend $u_{\mathbf{h}_k}$ by defining $u_{\mathbf{h}_k}(\mathbf{x}) = 0$ for all $\mathbf{x} \in \mathbb{R}^d \setminus \bar{\Omega}$. Then, by the definition of \bar{v} in (9) and the fact that $u_{\mathbf{h}_k} \rightarrow v$ pointwise, there exist sequences $\mathbf{x}_k \rightarrow \mathbf{x}_0$ and $\mathbf{h}_k \rightarrow \mathbf{0}$, an index i^* , and a constant $c^* > 0$ such that

$$\mathbf{x}_k \in \mathcal{T}_{\mathbf{h}_k}, \quad \lim_{k \rightarrow \infty} u_{\mathbf{h}_k}(\mathbf{x}_k) = \bar{v}(\mathbf{x}_0),$$

and

$$\lim_{k \rightarrow \infty} \left[u_{\mathbf{h}_k}(\mathbf{x}_k + h_{i^*}^{(k)} \mathbf{e}_{i^*}) - u_{\mathbf{h}_k}(\mathbf{x}_k) \right] = -c^*, \quad \lim_{k \rightarrow \infty} \left[u_{\mathbf{h}_k}(\mathbf{x}_k) - u_{\mathbf{h}_k}(\mathbf{x}_k - h_{i^*}^{(k)} \mathbf{e}_{i^*}) \right] \geq 0$$

or

$$\lim_{k \rightarrow \infty} [u_{\mathbf{h}_k}(\mathbf{x}_k + h_{i^*}^{(k)} \mathbf{e}_{i^*}) - u_{\mathbf{h}_k}(\mathbf{x}_k)] \leq 0, \quad \lim_{k \rightarrow \infty} [u_{\mathbf{h}_k}(\mathbf{x}_k) - u_{\mathbf{h}_k}(\mathbf{x}_k - h_{i^*}^{(k)} \mathbf{e}_{i^*})] = c^*.$$

Furthermore,

$$\lim_{k \rightarrow \infty} [u_{\mathbf{h}_k}(\mathbf{x}_k + h_j^{(k)} \mathbf{e}_j) - u_{\mathbf{h}_k}(\mathbf{x}_k)] \leq 0, \quad \lim_{k \rightarrow \infty} [u_{\mathbf{h}_k}(\mathbf{x}_k) - u_{\mathbf{h}_k}(\mathbf{x}_k - h_j^{(k)} \mathbf{e}_j)] \geq 0$$

for all $j = 1, 2, \dots, d$ with $j \neq i^*$. Thus, we can arrive at a contradiction using the same arguments as in the proof of Lemma 3.1. The proof is complete. \square

Theorem 3.1. Choose a mesh $\mathcal{T}_{\mathbf{h}}$, and let $\mathbf{h}_k = \frac{1}{k} \mathbf{h}$ for all $k \geq 1$. Let $U^{(k)} \in S(\mathcal{T}_{\mathbf{h}_k})$ be a solution to the FD scheme defined by (4a) and (4c). Let $u_{\mathbf{h}_k}$ be the piecewise constant extension of $U^{(k)}$ defined by (7), and let $u_{\mathbf{h}_{k'}}$ be a convergent subsequence such that $u_{\mathbf{h}_{k'}} \rightarrow v \in C(\overline{\Omega})$ pointwise as $k' \rightarrow \infty$ with $v = 0$ over $\partial\Omega$. Then v is a viscosity solution of (1a) and (1c) for any continuous reaction term f provided the underlying scheme is ℓ^∞ stable. Furthermore, v is a viscosity solution of (1) if $U^{(k)}$ is nonnegative.

Proof. For ease of presentation, all convergent sequences are understood as subsequences. Such a function v exists by Lemma 3.1. We verify v is a viscosity solution by verifying that it is both a viscosity subsolution and a viscosity supersolution.

To show v is a viscosity subsolution of (1), let $\varphi \in C^2(\overline{\Omega})$ such that $v - \varphi$ takes a strict local maximum at $\mathbf{x}_0 \in \Omega$. We first assume that $\varphi \in \mathcal{P}_2$, the set of all quadratic polynomials. Without a loss of generality, we assume $v(\mathbf{x}_0) = \varphi(\mathbf{x}_0)$ (after a translation in the dependent variable). Then there exists a ball, $B_{r_0}(\mathbf{x}_0) \subset \mathbb{R}^d$, centered at \mathbf{x}_0 with radius $r_0 > 0$ (in the C^0 metric) such that

$$v(\mathbf{x}) - \varphi(\mathbf{x}) < v(\mathbf{x}_0) - \varphi(\mathbf{x}_0) = 0 \quad \forall \mathbf{x} \in (B_{r_0}(\mathbf{x}_0) \cap \overline{\Omega}) \setminus \{\mathbf{x}_0\}. \tag{10}$$

Hence, there exists a sequence $\mathbf{x}_k \rightarrow \mathbf{x}_0$ such that

$$\limsup_{k \rightarrow \infty} \delta_{x_i, h_i^{(k)}}^2 u_{\mathbf{h}_k}(\mathbf{x}_k) \leq \delta_{x_i, h_i^{(k)}}^2 \varphi(\mathbf{x}_k) = \varphi_{x_i x_i}(\mathbf{x}_0) \tag{11}$$

for all $i = 1, 2, \dots, d$. Thus, by the definition of the scheme and the continuity of f , we have

$$0 = \liminf_{k \rightarrow \infty} [-\Delta_{\mathbf{h}_k} u_{\mathbf{h}_k}(\mathbf{x}_k) - \lambda f(u_{\mathbf{h}_k}(\mathbf{x}_k))] \geq \liminf_{k \rightarrow \infty} [-\Delta\varphi(\mathbf{x}_0) - \lambda f(\varphi(\mathbf{x}_0))] = -\Delta\varphi(\mathbf{x}_0) - \lambda f(\varphi(\mathbf{x}_0)).$$

We now consider the case of a general test function $\varphi \in C^2(\overline{\Omega})$. Recall that $v - \varphi$ is assumed to have a local maximum at \mathbf{x}_0 . Using Taylor's formula we write

$$\begin{aligned} \varphi(\mathbf{x}) &= \varphi(\mathbf{x}_0) + \nabla\varphi(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) + \frac{1}{2}(\mathbf{x} - \mathbf{x}_0)^T D^2\varphi(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) + o(|\mathbf{x} - \mathbf{x}_0|^2) \\ &\equiv p(\mathbf{x}) + o(|\mathbf{x} - \mathbf{x}_0|^2). \end{aligned}$$

For any $\sigma > 0$, we define the following quadratic polynomial:

$$\begin{aligned} p^\sigma(\mathbf{x}) &\equiv p(\mathbf{x}) + \sigma|\mathbf{x} - \mathbf{x}_0|^2 \\ &= \varphi(\mathbf{x}_0) + \nabla\varphi(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) + (\mathbf{x} - \mathbf{x}_0)^T [\sigma I_{d \times d} + \frac{1}{2} D^2\varphi(\mathbf{x}_0)](\mathbf{x} - \mathbf{x}_0). \end{aligned}$$

Trivially, $\nabla p^\sigma(\mathbf{x}) = \nabla\varphi(\mathbf{x}_0) + [2\sigma + D^2\varphi(\mathbf{x}_0)](\mathbf{x} - \mathbf{x}_0)$, $D^2 p^\sigma(\mathbf{x}) = 2\sigma I_{d \times d} + D^2\varphi(\mathbf{x}_0)$, and $\varphi(\mathbf{x}) - p^\sigma(\mathbf{x}) = o(|\mathbf{x} - \mathbf{x}_0|^2) - \sigma|\mathbf{x} - \mathbf{x}_0|^2 \leq 0$. Thus, $\varphi - p^\sigma$ has a local maximum at \mathbf{x}_0 and, therefore, $v - p^\sigma$ has a local maximum at \mathbf{x}_0 . It follows that $-\Delta p^\sigma(\mathbf{x}_0) - \lambda f(p^\sigma(\mathbf{x}_0)) \leq 0$. Taking $\liminf_{\sigma \rightarrow 0^+}$ and using the continuity of f we obtain

$$0 \geq \liminf_{\sigma \rightarrow 0^+} [-\Delta p^\sigma(\mathbf{x}_0) - \lambda f(p^\sigma(\mathbf{x}_0))] \geq \liminf_{\sigma \rightarrow 0^+} [-2d\sigma - \Delta\varphi(\mathbf{x}_0) - \lambda f(\varphi(\mathbf{x}_0))] = -\Delta\varphi(\mathbf{x}_0) - \lambda f(\varphi(\mathbf{x}_0)).$$

Thus, v is a viscosity subsolution of (1).

By following almost the same lines as those above we can show that if $v - \varphi$ takes a local minimum at $\mathbf{x}_0 \in \Omega$ for some $\varphi \in C^2(\overline{\Omega})$, then

$$0 \leq -\Delta\varphi(\mathbf{x}_0) - \lambda f(\varphi(\mathbf{x}_0)).$$

Hence, v is a viscosity supersolution of (1). The proof is complete. \square

Remark 3.2. Combining the results of Sections 3 and 4, we have the boundary value problem (1a) and (1c) has a viscosity solution. The existence proof holds for Ω a d -rectangle. Thus, it does not require $\partial\Omega$ is smooth. If f is positive, then we can guarantee $v > 0$ at all interior points yielding a solution to (1). If f is only semipositive, then we cannot guarantee v is nonnegative unless $\lambda > 0$ is sufficiently large similar to the classical case.

Remark 3.3. By construction, it follows that any limit point for the discrete problem (4a) and (4c) is a solution to the boundary value problem (1a) and (1c). Thus, each convergent sequence of grid functions produced by the FD scheme can naturally be associated with a particular PDE solution. Furthermore, using the techniques in [16], a discrete solution is guaranteed to exist in a neighborhood of a PDE solution. Thus, the FD method captures all PDE solutions (including the strictly positive ones) in limit without introducing any false solutions. Combining these observations with the convergence result to FD solutions that also satisfy (4b), we can ensure convergence to PDE solutions that also satisfy (1b).

Remark 3.4. The proofs of Lemma 3.1, Corollary 3.1, and Theorem 3.1 can easily be updated to non-uniformly refined meshes using the multi-limit technique found in [32].

4. Existence, stability, and uniqueness

We show that the FD method (4) has a solution for all $\lambda > 0$ when f is positone and the FD method (4a) and (4c) has a solution for all $\lambda > 0$ when f is semipositone. In Section 5 we will find positive solutions of (4a) and (4c) for appropriate λ values when f is semipositone. In this section, we propose a fixed point iteration that naturally utilizes discrete sub- and supersolutions to define a monotone, nonexpansive operator. The mapping will also yield ℓ^∞ -norm stability for solutions of (4a) and (4c). Lastly, we will show that the scheme has a unique solution for all $\lambda > 0$ sufficiently small when f is positone and no solution that also satisfies (4b) for all $\lambda > 0$ sufficiently small when $f(0) < 0$, consistent with the PDE theory. The following strongly uses the fact that the matrix representation of $-\Delta_{\mathbf{h}}$ with Dirichlet boundary conditions is a monotone matrix in the sense that the inverse matrix has all nonnegative components since $-\Delta_{\mathbf{h}}$ corresponds to an M-matrix with a minimal eigenvalue strictly bounded below by a positive constant independent of \mathbf{h} (see [33]). In other words, we rely upon a discrete maximum principle for Poisson’s equation.

4.1. Constructing a uniformly bounded positive discrete supersolution for (4)

Define $m_i = (b_i + a_i)/2$ for $i = 1, 2, \dots, d$ to be the midpoint of the domain Ω along the x_i direction, and define the quadratic function $\psi : \bar{\Omega} \rightarrow \mathbb{R}$ by

$$\psi(\mathbf{x}) = -\frac{1}{2d} \sum_{i=1}^d (x_i - m_i)^2 + \sum_{i=1}^d (b_i - a_i)^2.$$

Observe that $\psi(\mathbf{x}) > 0$ over $\bar{\Omega}$ with $\psi_{x_i x_i} = -\frac{1}{d}$. Consequently, $-\Delta_{\mathbf{h}}\psi = -\Delta\psi = 1$.

Define the grid function \bar{U} by $\bar{U}_\alpha = c\psi(\mathbf{x}_\alpha)$ for all $\mathbf{x}_\alpha \in \mathcal{T}_{\mathbf{h}}$. Then, there holds

$$-\Delta_{\mathbf{h}}\bar{U}_\alpha - \lambda f(\bar{U}_\alpha) = c - \lambda f(c\psi(\mathbf{x}_\alpha)) > 0$$

for all c sufficiently large by the sublinear growth of f and the boundedness of ψ . Thus, for $c \gg 0$, \bar{U} is a discrete supersolution of (4), where the value for c can be determined independent of \mathbf{h} . Furthermore, the choice for \bar{U} does not depend on whether f is positone or semipositone.

We can also construct a positive discrete supersolution for (4) for each fixed mesh $\mathcal{T}_{\mathbf{h}}$ using the following technique based on the principle eigenfunction of the discrete Laplacian operator. The supersolution can be used for the solvers in Section 5; however, a uniform stability estimate would require more assumptions regarding f . In particular, for $f(0) > 0$, the constant c below would be inversely dependent upon h_* .

Define the grid function \tilde{U} by $\tilde{U}_\alpha = c\phi_\alpha$ for all $\mathbf{x}_\alpha \in \mathcal{T}_{\mathbf{h}}$, where c is a positive constant and (λ_1, ϕ) denotes the principle eigenpair of the discrete Laplacian operator $-\Delta_{\mathbf{h}}$ over Ω . Then, $\lambda_1 > 0$ and $\phi_\alpha > 0$ for all $\mathbf{x}_\alpha \in \mathcal{T}_{\mathbf{h}} \cap \Omega$ with $\|\phi\|_{\ell^\infty(\mathcal{T}_{\mathbf{h}})}$ bounded independent of \mathbf{h} . Observe that

$$-\Delta_{\mathbf{h}}\tilde{U}_\alpha - \lambda f(\tilde{U}_\alpha) = \lambda_1 c\phi_\alpha - \lambda f(c\phi_\alpha) > 0$$

for c sufficiently large by the sublinear growth of f . Thus, for $c \gg 0$, \tilde{U} is a positive discrete supersolution of (4).

4.2. Constructing a uniformly bounded discrete subsolution for (4a) and (4c)

Define the grid function \underline{U} as the unique solution to

$$-\Delta_{\mathbf{h}}\underline{U}_\alpha = \lambda \min\{0, f(0)\} \quad \text{if } \mathbf{x}_\alpha \in \mathcal{T}_{\mathbf{h}} \cap \Omega, \tag{12a}$$

$$\underline{U}_\alpha = 0 \quad \text{if } \mathbf{x}_\alpha \in \mathcal{T}_{\mathbf{h}} \cap \partial\Omega. \tag{12b}$$

We consider two cases to show that \underline{U} is a discrete subsolution. First, suppose $f(0) \geq 0$. Then, we have $\underline{U}_\alpha = 0$ for all $\mathbf{x}_\alpha \in \mathcal{T}_{\mathbf{h}}$, and it follows that

$$-\Delta_{\mathbf{h}}\underline{U}_\alpha - \lambda f(\underline{U}_\alpha) = -\lambda f(0) \leq 0.$$

Next, suppose $f(0) < 0$. Then, we have $\underline{U}_\alpha \leq 0$ for all $\mathbf{x}_\alpha \in \mathcal{T}_h$, and, by (H3) and (H4), there holds

$$-\Delta_h \underline{U}_\alpha - \lambda f(\underline{U}_\alpha) = -\Delta_h \underline{U}_\alpha - \lambda f(0) = 0$$

for all $\mathbf{x}_\alpha \in \mathcal{T}_h \cap \Omega$. Thus, \underline{U} is a discrete subsolution of (4), and it is a positive discrete subsolution if $f(0) \geq 0$. Furthermore, there exists a constant $c \geq 0$ independent of h such that $0 \geq \underline{U}_\alpha \geq -c$ for all $\mathbf{x}_\alpha \in \mathcal{T}_h$.

4.3. Existence and stability for solutions of (4a) and (4c)

Theorem 4.1. *There exists at least one grid function U such that U solves (4a) and (4c). If f is positone, then U also satisfies (4b). Furthermore, there exists a constant C independent of h such that*

$$\|U\|_{\ell^\infty(\mathcal{T}_h)} < C,$$

i.e., the solution is ℓ^∞ -norm stable.

Proof. Let $S(\mathcal{T}_h)$ denote the space of all grid functions defined over \mathcal{T}_h . Choose $\rho > 0$, and define the mapping $\mathcal{M}_\rho : S(\mathcal{T}_h) \rightarrow S(\mathcal{T}_h)$ by

$$\widehat{U} = \mathcal{M}_\rho U, \tag{13}$$

where

$$\widehat{U}_\alpha = U_\alpha - \rho [-\Delta_h U_\alpha - \lambda f(U_\alpha)] \quad \text{if } \mathbf{x}_\alpha \in \mathcal{T}_h \cap \Omega, \tag{14a}$$

$$\widehat{U}_\alpha = 0 \quad \text{if } \mathbf{x}_\alpha \in \mathcal{T}_h \cap \partial\Omega. \tag{14b}$$

Clearly a fixed point of (13) is also a solution to (4a) and (4c).

By the definition of Δ_h , we have (14a) is increasing with respect to $U_{\alpha \pm \mathbf{e}_i}$ for all $i = 1, 2, \dots, d$. Furthermore, for $\rho < \frac{h_*^2}{2d}$, (14a) is increasing with respect to U_α . Thus, the mapping \mathcal{M}_ρ is monotone for all $\rho > 0$ sufficiently small.

Let \leq denote the partial ordering for vectors, and let $W \in S(\mathcal{T}_h)$ such that $\underline{U} \leq W \leq \bar{U}$, where \bar{U} is defined in Section 4.1 and \underline{U} is defined in Section 4.2. Define \vec{F}_0, \vec{F}^0 by

$$[\vec{F}_0]_\alpha \equiv f(\underline{U}_\alpha) = f(0), \quad [\vec{F}^0]_\alpha \equiv f(\bar{U}_\alpha).$$

There holds

$$\mathcal{M}_\rho W \geq \mathcal{M}_\rho \underline{U} = \underline{U} - \rho \left[-\Delta_h \underline{U} - \lambda \min\{\vec{0}, \vec{F}_0\} \right] = \underline{U},$$

$$\mathcal{M}_\rho W \leq \mathcal{M}_\rho \bar{U} = \bar{U} - \rho \left[-\Delta_h \bar{U} - \lambda \vec{F}^0 \right] \leq \bar{U}$$

over $\mathcal{T}_h \cap \Omega$ by the monotonicity of \mathcal{M}_ρ and the fact that \bar{U} is a discrete supersolution. Since $\mathcal{M}_\rho W = 0$ over $\mathcal{T}_h \cap \partial\Omega$, there holds $\underline{U} \leq \mathcal{M}_\rho W \leq \bar{U}$ over \mathcal{T}_h , and it follows by the Brouwer Fixed Point Theorem that \mathcal{M}_ρ has a fixed point U such that $\underline{U} \leq U \leq \bar{U}$. Furthermore, if f is positone, i.e., $f(0) > 0$, then, for all $\mathbf{x}_\alpha \in \mathcal{T}_h \cap \Omega$, there holds

$$U_\alpha = \mathcal{M}_\rho U_\alpha \geq \mathcal{M}_\rho \underline{U}_\alpha = \rho \lambda f(0) > 0.$$

Thus, U satisfies (4b), and it follows that U also satisfies (4b). Lastly, we have

$$\|U\|_{\ell^\infty(\mathcal{T}_h)} \leq \max \{ \|\bar{U}\|_{\ell^\infty(\mathcal{T}_h)}, \|\underline{U}\|_{\ell^\infty(\mathcal{T}_h)} \} < C$$

for some constant C independent of h . The proof is complete. \square

Remark 4.1. The mapping \mathcal{M}_ρ defined by (13) corresponds to an explicit pseudo timestepping iteration with a CFL condition $\rho \leq Ch_*^2$ for some constant C . Thus, the iteration embeds the nonlinear elliptic problem into a parabolic type problem. The fixed points would correspond to steady-state solutions of the parabolic problem.

Remark 4.2. The discrete subsolution \underline{U} was used to ensure the positivity of the solutions when f is positone. The discrete sub- and supersolutions together were used to define barriers that ensured the ℓ^∞ -norm stability of the solutions. The existence and stability results are only for solutions U such that $\underline{U} \leq U \leq \bar{U}$. If the boundary value problem has additional solutions, different choices for \underline{U} and \bar{U} are needed. If a positive discrete subsolution exists, it can be used to construct a lower bound that ensures (4b) is satisfied independent of whether f is positone or semipositone.

4.4. Uniqueness for solutions of (4) when λ is small and f is positone

Theorem 4.2. *Suppose the operator f in (1) is Lipschitz and positone. Then the FD method (4) has a unique solution for all $\lambda > 0$ sufficiently small independent of h .*

Proof. Suppose U and V are solutions to (4). Then, by the Mean Value Theorem, there exists a bounded function $\kappa \geq 0$ such that

$$-\Delta_{\mathbf{h}}(U_{\alpha} - V_{\alpha}) = \lambda [f(U_{\alpha}) - f(V_{\alpha})] = \lambda \kappa(\mathbf{x}_{\alpha})(U_{\alpha} - V_{\alpha}).$$

Thus, $W \equiv U - V$ is the solution to the linear problem

$$\begin{aligned} A_{\mathbf{h}}W_{\alpha} &\equiv (-\Delta_{\mathbf{h}} - \lambda\kappa(\mathbf{x}_{\alpha}))W_{\alpha} = 0 & \text{if } \mathbf{x}_{\alpha} \in \Omega, \\ W_{\alpha} &= 0 & \text{if } \mathbf{x}_{\alpha} \in \partial\Omega. \end{aligned}$$

Let A denote the matrix representation of $A_{\mathbf{h}}$, L denote the matrix representation of $-\Delta_{\mathbf{h}}$, and D denote the matrix representation of κ . Then, L is a symmetric positive definite matrix and D is a diagonal matrix. Thus, $A = L - \lambda D$ is nonsingular for all $\lambda > 0$ sufficiently small, and it follows that $W_{\alpha} = 0$ for all $\mathbf{x}_{\alpha} \in \mathcal{T}_{\mathbf{h}}$. Note that the bound for λ is independent of \mathbf{h} since A is a symmetric matrix with $L - \lambda D \geq (\lambda_0 - \lambda\kappa^*)I$ for $\lambda_0 > 0$ the minimal eigenvalue of L , which is bounded below independent of \mathbf{h} , and κ^* an upper bound for the Lipschitz constant of f which is determined by \underline{U} and \overline{U} to remove the dependence on W . The proof is complete. \square

Remark 4.3. Theorem 4.2 is comparable to a known uniqueness theorem for the PDE (1). See Section 2 in [3] for a discussion of such a result.

4.5. Nonexistence of a positive solution to (4) when λ is small and f is Semipositone

Theorem 4.3. Suppose $f(0) < 0$. Then the FD method (4) has no solution for all $\lambda > 0$ sufficiently small independent of \mathbf{h} .

Proof. We use a proof by contradiction. Suppose (4) has a solution U for all $\lambda > 0$. Let $c_0 > 0$ denote the constant in (H1) where $f(c_0) = 0$. By Theorem 4.1, we can assume there exists a constant $C > c_0$ such that $-C \leq U_{\alpha} \leq C$ for all $\mathbf{x}_{\alpha} \in \mathcal{T}_{\mathbf{h}}$.

Choose $\rho > 0$, and define the grid function $V^{(\rho)}$ by

$$-\Delta_{\mathbf{h}}V_{\alpha}^{(\rho)} = \rho \quad \text{if } \mathbf{x}_{\alpha} \in \mathcal{T}_{\mathbf{h}} \cap \Omega, \tag{15a}$$

$$V_{\alpha}^{(\rho)} = 0 \quad \text{if } \mathbf{x}_{\alpha} \in \mathcal{T}_{\mathbf{h}} \cap \partial\Omega. \tag{15b}$$

Then, there exists $\rho > 0$ such that $0 \leq V_{\alpha}^{(\rho)} < c_0$ for all $\mathbf{x}_{\alpha} \in \mathcal{T}_{\mathbf{h}}$.

Choose $\lambda > 0$ such that $\lambda f(C) < \rho$. Then,

$$-\Delta_{\mathbf{h}}U_{\alpha} = \lambda f(U_{\alpha}) \leq \lambda f(C) < \rho,$$

and it follows that $U < V^{(\rho)}$. Hence,

$$-\Delta_{\mathbf{h}}U_{\alpha} = \lambda f(U_{\alpha}) < \lambda f(c_0) = 0,$$

and we have $U_{\alpha} < 0$ for all $\mathbf{x}_{\alpha} \in \mathcal{T}_{\mathbf{h}} \cap \Omega$ contradicting the assumption that U satisfies (4b). The proof is complete. \square

5. Solving the algebraic system and forming bifurcation curves

We provide a guaranteed solver in this section that can be used to test for nonuniqueness for positone problems, to test for nonexistence for semipositone problems, and to assist in generating bifurcation curves. For $\lambda > 0$ fixed, the solver will converge monotonically provided the reaction term f in (1) is Lipschitz continuous and the initial guess is given by a discrete subsolution or a discrete supersolution. The results are an application of the more general results and techniques found in [34].

Consider the fixed point iteration

$$U^{(n+1)} = \mathcal{M}_K U^{(n)} \tag{16}$$

for all $n \geq 0$, where K is a Lipschitz constant for f in (1) and \mathcal{M}_K is defined such that

$$-\Delta_{\mathbf{h}}U_{\alpha}^{(n+1)} + \lambda K U_{\alpha}^{(n+1)} = \lambda f(U_{\alpha}^{(n)}) + \lambda K U_{\alpha}^{(n)} \quad \text{if } \mathbf{x}_{\alpha} \in \Omega, \tag{17a}$$

$$U_{\alpha}^{(n+1)} = 0 \quad \text{if } \mathbf{x}_{\alpha} \in \partial\Omega. \tag{17b}$$

Clearly the iteration is well-defined. Furthermore, a fixed point of (16) is also a solution to (4) if we can show (4b) is satisfied.

Lemma 5.1. Let U be a solution to (4a) and (4c). If $U^{(0)} \leq U$ is a subsolution of (4), then $U^{(1)} = \mathcal{M}_K U^{(0)}$ is a subsolution of (4) with $U^{(0)} \leq U^{(1)} \leq U$. If $U^{(0)} \geq U$ is a supersolution of (4), then $U^{(1)} = \mathcal{M}_K U^{(0)}$ is a supersolution of (4) with $U \leq U^{(1)} \leq U^{(0)}$.

Proof. Suppose that $U^{(0)} \leq U$ is a subsolution of (4). Observe that, by the definition of \mathcal{M}_K , there holds

$$-\Delta_{\mathbf{h}}U_{\alpha}^{(1)} + \lambda KU_{\alpha}^{(1)} = \lambda f(U_{\alpha}^{(0)}) + \lambda KU_{\alpha}^{(0)} \geq -\Delta_{\mathbf{h}}U_{\alpha}^{(0)} + \lambda KU_{\alpha}^{(0)}$$

for all $\mathbf{x}_{\alpha} \in \mathcal{T}_{\mathbf{h}} \cap \Omega$. Let M denote the standard matrix representation of $-\Delta_{\mathbf{h}} + \lambda K$. Then M is a monotone matrix and we have $U^{(1)} \geq U^{(0)}$.

We also have, for some $0 \leq \kappa \leq K$,

$$\begin{aligned} -\Delta_{\mathbf{h}}U_{\alpha}^{(1)} + \lambda KU_{\alpha}^{(1)} &= \lambda f(U_{\alpha}^{(1)}) + \lambda KU_{\alpha}^{(0)} + \lambda [f(U_{\alpha}^{(0)}) - f(U_{\alpha}^{(1)})] \\ &\leq \lambda f(U_{\alpha}^{(1)}) + \lambda KU_{\alpha}^{(0)} - \lambda \kappa [U_{\alpha}^{(0)} - U_{\alpha}^{(1)}] \\ &= \lambda f(U_{\alpha}^{(1)}) + \lambda \kappa U_{\alpha}^{(1)} + \lambda (K - \kappa) U_{\alpha}^{(0)} \end{aligned}$$

for all $\mathbf{x}_{\alpha} \in \mathcal{T}_{\mathbf{h}} \cap \Omega$. Subtracting $\lambda KU_{\alpha}^{(1)}$ from both sides and using the fact that $K \geq \kappa$ and $U^{(1)} \geq U^{(0)}$, it follows that

$$-\Delta_{\mathbf{h}}U_{\alpha}^{(1)} \leq \lambda f(U_{\alpha}^{(1)}) + \lambda (K - \kappa) (U_{\alpha}^{(0)} - U_{\alpha}^{(1)}) \leq \lambda f(U_{\alpha}^{(1)})$$

for all $\mathbf{x}_{\alpha} \in \mathcal{T}_{\mathbf{h}} \cap \Omega$. Therefore, $U^{(1)}$ is a subsolution of (4).

Lastly, since the function $\lambda(f + K)$ is monotone increasing and $U^{(0)} \leq U$, there holds

$$-\Delta_{\mathbf{h}}U_{\alpha}^{(1)} + \lambda KU_{\alpha}^{(1)} = \lambda f(U_{\alpha}^{(0)}) + \lambda KU_{\alpha}^{(0)} \leq \lambda f(U_{\alpha}) + \lambda KU_{\alpha} = -\Delta_{\mathbf{h}}U_{\alpha} + \lambda KU_{\alpha}$$

for all $\mathbf{x}_{\alpha} \in \mathcal{T}_{\mathbf{h}} \cap \Omega$. Thus, $U^{(1)} \leq U$.

The case when $U^{(0)} \geq U$ is a supersolution of (4) is similar. The proof is complete. \square

Theorem 5.1. Let U be a solution to (4). If $U^{(0)} \leq U$ is a subsolution of (4) with $U_{\alpha}^{(0)} \geq 0$ for all $\mathbf{x}_{\alpha} \in \mathcal{T} \cap \Omega$, then the sequence $U^{(n)}$ defined by (16) converges to a solution of (4). If $U^{(0)} \geq U$ is a supersolution of (4), then the sequence $U^{(n)}$ defined by (16) converges to a solution of (4).

Proof. Observe that, by Lemma 5.1, we have

$$0 \leq U_{\alpha}^{(0)} \leq U_{\alpha}^{(1)} \leq \dots \leq U_{\alpha}^{(n)} \leq U_{\alpha} \quad \text{or} \quad U_{\alpha}^{(0)} \geq U_{\alpha}^{(1)} \geq \dots \geq U_{\alpha}^{(n)} \geq U_{\alpha} \geq 0$$

for all $\mathbf{x}_{\alpha} \in \mathcal{T} \cap \Omega$ and $n \geq 1$. Thus, the sequence $U^{(n)}$ is monotone and bounded, and it follows that there exists $V \in S(\mathcal{T}_{\mathbf{h}})$ such that $U^{(n)} \rightarrow V$. Since V is nonnegative and V is a fixed point of (16), it follows that V is a solution to (4). \square

Remark 5.1. If f is positone, we can let $U_{\alpha}^{(0)} = 0 < U_{\alpha}$ for all $\mathbf{x}_{\alpha} \in \mathcal{T} \cap \Omega$ which implies $U^{(1)}$ is a subsolution with $U_{\alpha}^{(1)} > 0$ for all $\mathbf{x}_{\alpha} \in \mathcal{T} \cap \Omega$ since $f(0) > 0$.

Remark 5.2. If f is positone, the iteration can be used to determine if the nonlinear system (4) has a unique solution for a given value of λ . Indeed, letting $U^{(0)} = \underline{U}$, there exists a function U_1 such that $U^{(n)} \nearrow U_1$. Letting $U^{(0)} = \bar{U}$, there exists a function U_2 such that $U^{(n)} \searrow U_2$. If $U_2 > U_1$, then the system (4) has at least two solutions.

Remark 5.3. If f is semipositone with $f(0) < 0$, the iteration can be used to determine if the nonlinear system (4) has a solution for a given value of λ . By setting $U^{(0)} = \bar{U}$, we have $U^{(0)} \geq \bar{0}$, and the iteration will monotonically converge to U_2 which represents the maximal solution of (4a) and (4c) in the range of admissible solutions. If U_2 does not satisfy (4b), then we can conclude that (4) has no solution for the given value of λ .

The mapping \mathcal{M}_K can naturally be used to find a minimal solution and a maximal solution to (4a) and (4c). As such, the methods can be applied to assist in generating $(\lambda, \|\cdot\|_{\max})$ -bifurcation curves since they can be used to find the minimal branch and the maximal branch. In order to speed up the process, we use the method of continuation [35] when varying λ . Thus, for semipositone problems, most of the work goes into finding an initial point on the bifurcation curve. The proposed iteration can be used to find such a point when $\lambda > 0$ is sufficiently large after which more efficient continuation methods can be used.

When f is positone, we can more systematically resolve the bifurcation curve and speed up the continuation process. Suppose (λ, U) is a minimal solution to (4). Observe that, for all $\rho > 0$, there holds

$$-\Delta_{\mathbf{h}}U_{\alpha} = \lambda f(U_{\alpha}) < (\lambda + \rho)f(U_{\alpha})$$

since $f(U_{\alpha}) > 0$. Thus, $(\lambda + \rho, U)$ is a subsolution for (4) which should be closer to the solution corresponding to $\lambda + \rho$ than \underline{U} . Once a minimal solution has been found for λ , we have a natural initial guess for the minimal solution corresponding to $\lambda + \rho$. Analogously, a maximal solution corresponding to λ is a natural initial guess for the maximal solution corresponding to $\lambda - \rho$. Once the minimal and maximal branches have been found, convex combinations of the solutions can be used to find a third solution for a given λ from which point a Newton-based solver and the method of continuation can be used to complete the corresponding branch of the bifurcation curve. This technique is used in Section 6 to generate bifurcation curves for positone problems. Alternative methods for numerically generating bifurcation curves and more information

about the method of continuation can be found in [36], where the strength of our approach is that it guarantees finding an initial point on the bifurcation curve.

We lastly note that for both positone and semipositone problems the maximal branch can be found in parallel for various values of λ since the corresponding supersolutions can be found independently. We can similarly find the minimal branch for positone problems in parallel. Then continuation only needs to be used to connect the branches.

6. Applications and numerical experiments

In this section we consider various applications of the reaction–diffusion problem (1). We will focus on positone problems that have S -shaped $(\lambda, \|u\|_\infty)$ -bifurcation curves (see [37]). In particular, the models have unique solutions when λ is sufficiently small and when λ is sufficiently large, but can feature three solutions for a range of λ values. We will also consider a semipositone problem that has no positive solutions for λ small and a unique solution for λ large with some range of λ values for which two solutions exist.

For each application problem we will perform several numerical tests to gauge the accuracy of our proposed FD method. The tests show that the method can successively capture multiple solutions by simply changing the initial guess for the nonlinear solver. We also generate bifurcation curves that support the theoretical results for the example problems. All of the tests use either the iteration \mathcal{M}_K in Section 5 or Matlab's built-in nonlinear solver *fsolve* to solve (4a) and (4c). For minimal and maximal solutions, initial guesses for *fsolve* are provided after using several iterations of \mathcal{M}_K . Other solutions are found by varying the initial guess to *fsolve*.

The bifurcation curves are generated using the method of continuation as discussed in Section 5. We use a simple variant where we use a uniform mesh for λ and first find the maximal solutions and, for positone problems, the minimal solutions to identify the two primary branches. Once we have identified the region with three solutions for positone problems, we find a solution on the third branch and then do one more sweep of the continuation method to fill in the full branch. Similarly, for semipositone problems, we seek to find the point along the minimal branch where the system no longer has multiple positive solutions. The figures use breaks and different colors to identify the different “branches”. These could easily be connected by continuity and fixing a value for the approximate turning point to be on the λ mesh. In general, continuation methods have many variants that could be used to design more sophisticated continuation schemes that are adaptive and can better resolve turning points. However, our primary focus in this paper is on the convergence of the FD scheme, finding positive solutions, ruling out convergence to false solutions even when the underlying discrete problems have multiple solutions, and constructing guaranteed initial guesses.

We can see that our method exhibits optimal rates of convergence when approximating a single solution with λ fixed. We also approximate critical λ values when turning points exist in the bifurcation curve. By considering various coarse and fine meshes, we see how the turning points evolve for $h \rightarrow 0^+$. We observe that the number of solutions for a given λ value does not appear to change significantly as $h \rightarrow 0$ showing that the finite algebraic system of equations resulting from the FD method qualitatively preserves the multiplicity aspects of the original PDE problem.

6.1. The Perturbed Gelfand Problem

A standard example of sublinear positone problems is the perturbed Gelfand problem (see Boddington et al. [38]). The problem chooses the reaction term

$$f(u) = e^{\frac{au}{a+u}}, \quad (18)$$

where λ is the ignition parameter, $a > 0$ is the activation energy, and $u(x)$ is the dimensionless temperature. The reaction term is based on the Arrhenius reaction-rate law in irreversible chemical reaction kinetics. It is conjectured that there exist a value $a_0 > 0$ and corresponding values $\lambda^* > \lambda_* > 0$ such that (1) with f defined by (18) has three positive solutions whenever $a > a_0$ and $\lambda_* < \lambda < \lambda^*$. Illustrations of the corresponding bifurcation curves are given in Fig. 1. A graph of (18) for $\alpha = 9$ can be found in Fig. 2 from which we see the sublinear and positone nature of f . More information about multiplicity results for (1) with f defined by (18) can be found in [39].

6.1.1. Example 1: 1D Gelfand problem

We approximate the Gelfand problem with $\Omega = (0, 1)$. We see in Fig. 3 that as a increases, an S -shaped bifurcation curve is recovered, and we see how the S -shaped bifurcation curve for $a = 6$ evolves as $h \rightarrow 0^+$ in Fig. 4. All three approximations for $a = 6$ and $\lambda = 4$ are pictured for various values of h in Fig. 5 with approximate rates of convergence calculated in Fig. 6.

6.1.2. Example 2: 2D Gelfand problem

We approximate the Gelfand problem with $\Omega = (0, 1)^2$. We see in Fig. 7 that for $a = 6$ the problem has an S -shaped bifurcation curve. Fig. 8 illustrates the solver in Section 5. We can see that subsolutions can be used to find the minimal solution for each λ while supersolutions can be used to find the maximal solution for each λ . All three approximations for $a = 6$ and $\lambda = 8$ are pictured in Fig. 9.

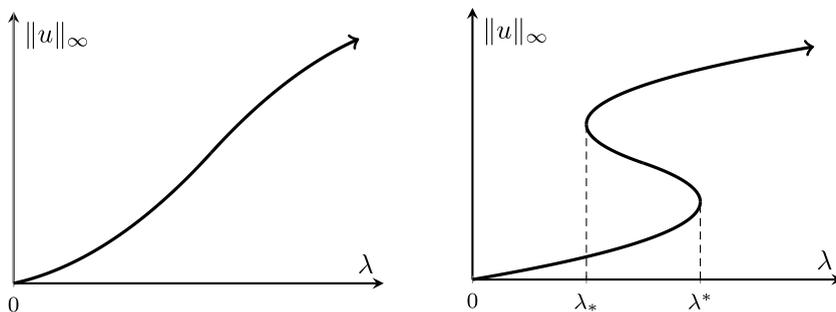


Fig. 1. Sketches of bifurcation curves for the perturbed Gelfand problem that support the conjecture of an S-shaped bifurcation curve when $a > a_0$ for some value a_0 . The curve on the left corresponds to $0 < a < a_0$ and the curve on the right corresponds to $a > a_0$.

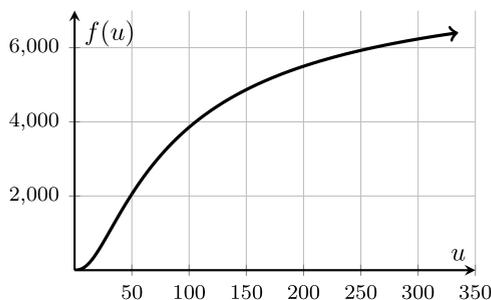


Fig. 2. The graph of f defined by (18) with $\alpha = 9$ which leads to an S-shaped bifurcation curve for (1) when restricting the problem to the interval $\Omega = (0, 1)$.

6.2. The model of Kernevez et al.

Another example with S-shaped bifurcation curves is given by the reaction term

$$f(u) = \frac{u}{1 + u + \beta u^2} \tag{19}$$

with $\Omega = (0, 1)$ and boundary condition $u(0) = u_0 = u(1)$ for β and u_0 positive constants. See [40] for more information about the model. When u_0 is sufficiently large, the corresponding reaction–diffusion problem has an S-shaped bifurcation curve. For the numerical tests we use the transformation $u \mapsto u_0 - u$ so that the model becomes

$$f(u) = \frac{u_0 - u}{1 + (u_0 - u) + \beta(u_0 - u)^2} \tag{20}$$

with $u(0) = 0 = u(1)$. The problem is also of interest since uniqueness results are still open in higher dimensions.

We note that the problem as represented by (20) is not strictly positone. The function f is positive only for $0 \leq u < u_0$, and the function is increasing on $[0, u_0 - \beta^{-1/2}] \subset [0, u_0]$. However, the problem is considered in the numerical tests section due to the existence of multiple positive solutions and the fact that the convergence analysis in Section 3 easily extends to this problem when assuming $0 \leq U \leq u_0$ for U the FD approximation, a result that is proved in Lemma 6.1 by updating the proof of Theorem 4.1. A graph of (20) for $\beta = 1$ and $u_0 = 10$ can be found in Fig. 10.

We now show that the Kernevez et al. model has a solution U such that $0 \leq U \leq u_0$.

Lemma 6.1. *There exists at least one grid function U such that U solves (4) with f defined by (20). Furthermore, $0 \leq U \leq u_0$.*

Proof. Again consider the mapping \mathcal{M}_ρ defined by (13). Then, for $\rho < \frac{1}{2} \min \left\{ \frac{h_\Omega^2}{2d}, \tau \right\}$ for some constant τ that depends on β and u_0 (based on maximizing $|f'(u)|$ on the interval $[0, u_0]$), (13) is monotone.

Let $W \in S(\mathcal{T}_h)$ such that $0 \leq W_\alpha \leq u_0$ for all $\mathbf{x}_\alpha \in \mathcal{T}_h$. Observe that

$$0 = [\mathcal{M}_\rho W]_\alpha \leq u_0 \quad \forall \mathbf{x}_\alpha \in \mathcal{T}_h \cap \partial\Omega.$$

Suppose $\mathbf{x} \in \mathcal{T}_h \cap \Omega$. Define the grid function $\mathbf{1}_h$ such that $\mathbf{1}_h(\mathbf{x}_\alpha) = 1$ for all $\mathbf{x}_\alpha \in \mathcal{T}_h$. Then,

$$\begin{aligned} [\mathcal{M}_\rho W]_\alpha &\geq [\mathcal{M}_\rho(0 \cdot \mathbf{1}_h)]_\alpha = 0 - \rho(-\Delta_h[0 \cdot \mathbf{1}_h]_\alpha - \lambda f(0)) = \rho \lambda f(0) > 0, \\ [\mathcal{M}_\rho W]_\alpha &\leq [\mathcal{M}_\rho(u_0 \cdot \mathbf{1}_h)]_\alpha = u_0 - \rho(-\Delta_h[u_0 \cdot \mathbf{1}_h]_\alpha - \lambda f(u_0)) = u_0 + \rho \lambda f(u_0) = u_0. \end{aligned}$$

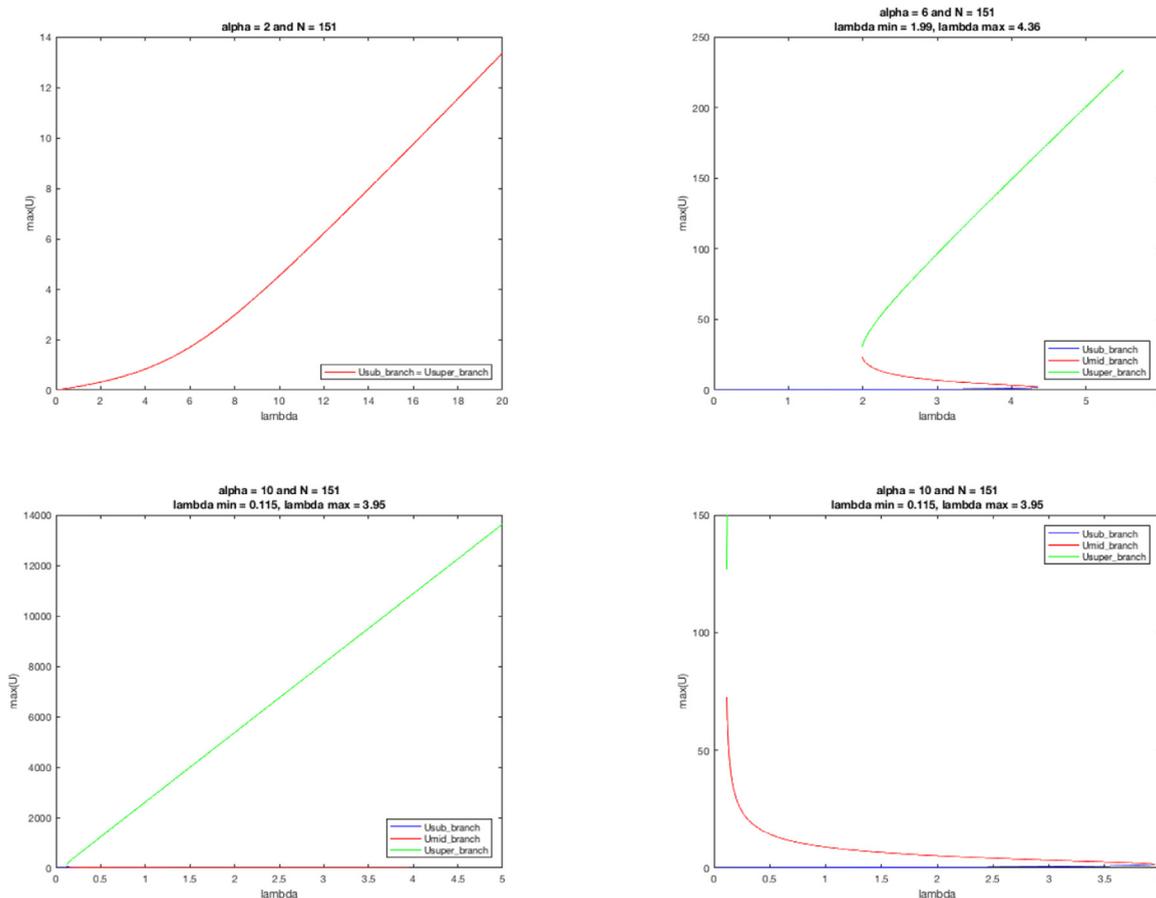


Fig. 3. Bifurcation curves for Application 6.1 when $N = 151$ and $a = 2$ (top left), $a = 6$ (top right), and $a = 10$ (bottom). The bottom right graph is zoomed in on the y-axis to better show the lower two branches of the bifurcation curve. There are no turning points when $a = 2$. The turning points for $a = 6$ are at $\lambda = 1.99$ and $\lambda = 4.36$. The turning points for $a = 10$ are at $\lambda = 0.115$ and $\lambda = 3.95$.

Thus, $0 \leq [\mathcal{M}_\rho W]_\alpha \leq u_0$ for all $\mathbf{x}_\alpha \in \mathcal{T}_h$, and it follows that \mathcal{M}_ρ has a fixed point U with $0 \leq U_\alpha \leq u_0$ for all $\mathbf{x}_\alpha \in \mathcal{T}_h \cap \Omega$. The proof is complete. \square

6.2.1. Example 3: 1D Kernevez problem

We approximate the Kernevez et al. model with $\Omega = (0, 1)$. We see in Fig. 11 how the S-shaped bifurcation curve for $u_0 = 10$ and $\beta = 1$ evolves as $h \rightarrow 0^+$. In Fig. 12, we first graph all three approximations for $u_0 = 10$, $\beta = 1$, and $\lambda = 285$. We also plot the unique solution for various large values of λ to see how the solution to the model changes as $\lambda \rightarrow \infty$. We can see that, as $\lambda \rightarrow \infty$, we expect $U \rightarrow u_0$ throughout Ω consistent with the bifurcation curve in Fig. 11 and the plots of U for various values of λ in Fig. 12. Lastly, we see the optimal convergence rates for all three solutions when $\lambda = 285$ in Fig. 13.

6.2.2. Example 4: 2D Kernevez problem

We approximate the Kernevez et al. model with $\Omega = (0, 1)^2$. We see in Fig. 14 that for $u_0 = 10$ the problem has an S-shaped bifurcation curve. Fig. 15 illustrates the solver in Section 5. We can see that subsolutions can be used to find the minimal solution for each λ while supersolutions can be used to find the maximal solution for each λ . The maximal and minimal approximations for $u_0 = 10$ and $\lambda = 560$ are pictured in Fig. 16. The experiment provides strong evidence in support of the conjecture that the PDE has multiple solutions in higher dimensions.

6.3. Logistic growth with constant yield harvesting

The last application that we consider is the semipositone reaction term corresponding to logistic growth with constant yield harvesting, i.e.,

$$f(u) = u(M - u) - c \tag{21}$$

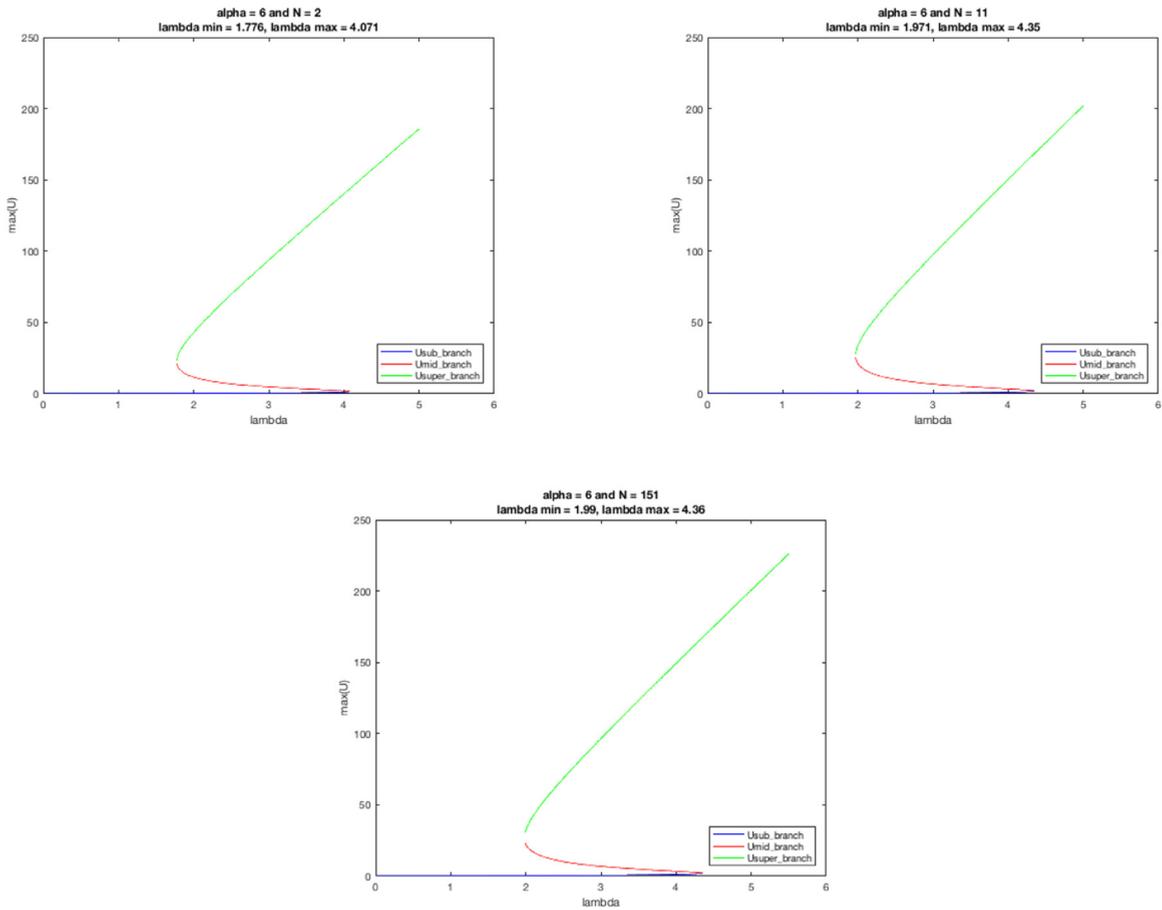


Fig. 4. Bifurcation curves for Application 6.1 when $a = 6$ and $N = 2$ (top left), $N = 11$ (top right), and $N = 151$ (bottom). The turning points for $N = 2$ are at $\lambda = 1.776$ and $\lambda = 4.071$. The turning points for $N = 11$ are at $\lambda = 1.971$ and $\lambda = 4.35$. The turning points for $N = 151$ are at $\lambda = 1.99$ and $\lambda = 4.36$.

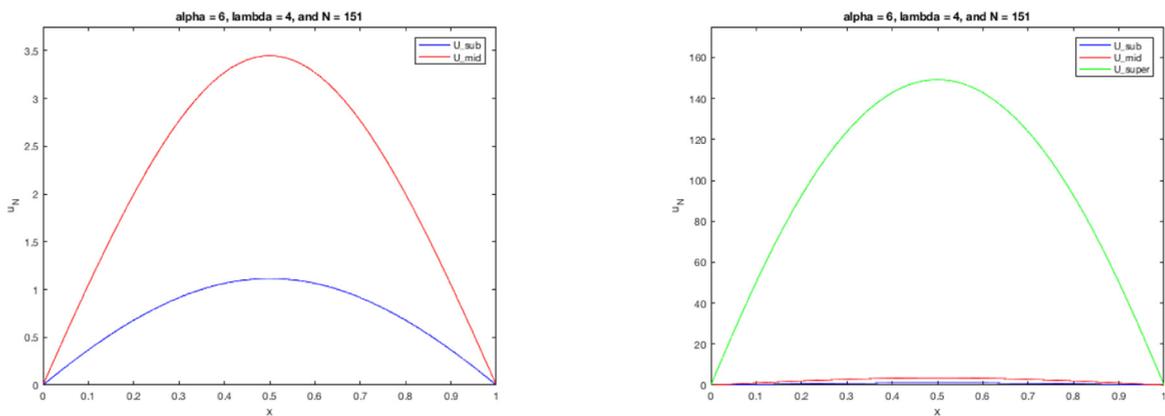


Fig. 5. Approximation curves for Application 6.1 when $\lambda = 4$, $N = 151$, and $a = 6$. The second graph is zoomed in on the y-axis to better show the lower two solutions.

for $M > 0$ the carrying capacity and $c > 0$ the harvesting term. In applications, we assume $0 \leq u \leq M$ and $M^2 - 4c > 0$. See [41] for more information about the model and the PDE analysis.

h	Umin Error	Order	Umid Error	Order	Umax Error	Order
2.00e-01	5.62e-02		4.12e-01		3.24e+00	
1.00e-01	1.29e-02	2.12	1.02e-01	2.01	1.19e+00	1.44
5.00e-02	3.12e-03	2.05	2.59e-02	1.98	3.25e-01	1.88
2.50e-02	7.74e-04	2.01	6.51e-03	2.00	7.72e-02	2.07
1.25e-02	1.93e-04	2.00	1.63e-03	2.00	1.87e-02	2.04
6.25e-03	4.83e-05	2.00	4.07e-04	2.00	4.65e-03	2.01
3.13e-03	1.21e-05	2.00	1.02e-04	2.00	1.16e-03	2.00
1.56e-03	3.02e-06	2.00	2.54e-05	2.00	2.90e-04	2.00

Fig. 6. Simulated rates of convergence for all three solutions of Application 6.1 when $a = 6$ and $\lambda = 4$.

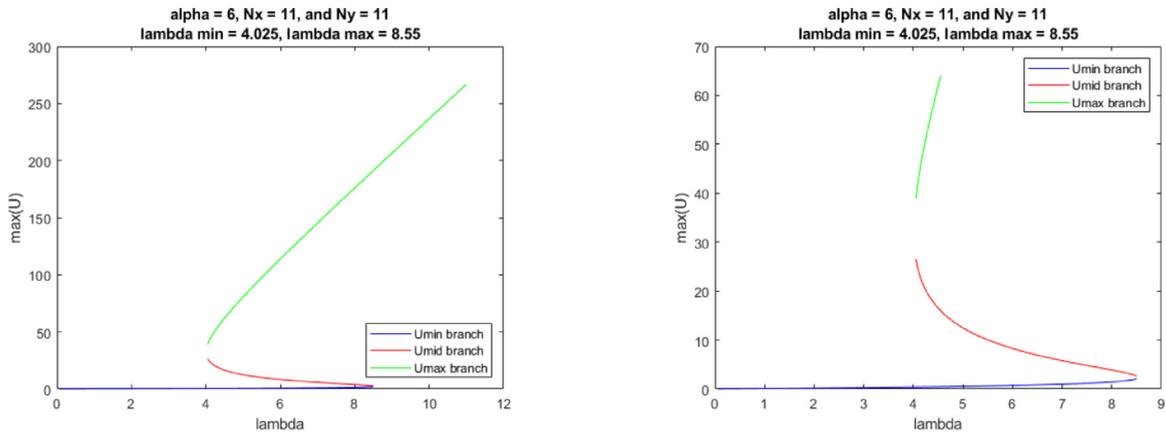


Fig. 7. Bifurcation curves for Application 6.1 in two dimensions when $h_x = 1/11$, $h_y = 1/11$, and $a = 6$. The second graph is zoomed in on the y -axis to better show the lower two branches of the bifurcation curve. The turning points are at $\lambda = 4.025$ and $\lambda = 8.55$.

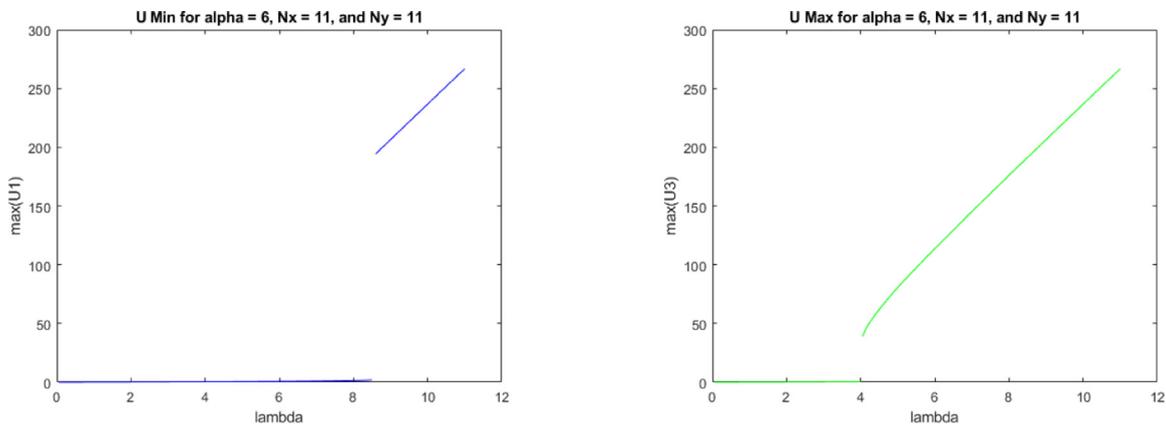


Fig. 8. Bifurcation curves generated using the monotone iteration in Section 5 for Application 6.1 in two dimensions when $h_x = 1/11$, $h_y = 1/11$, and $a = 6$. The first graph uses initial guesses based on subsolutions. The second graph uses initial guesses based on supersolutions.

We note that the problem as represented by (1) and (21) is not strictly semipositone. Instead, we consider the problem

$$\mathcal{L}u \equiv -\Delta u + \lambda \kappa u = \lambda \tilde{f}(u) \quad \text{in } \Omega, \tag{22a}$$

$$u \geq 0 \quad \text{in } \Omega, \tag{22b}$$

$$u = 0 \quad \text{on } \partial\Omega, \tag{22c}$$

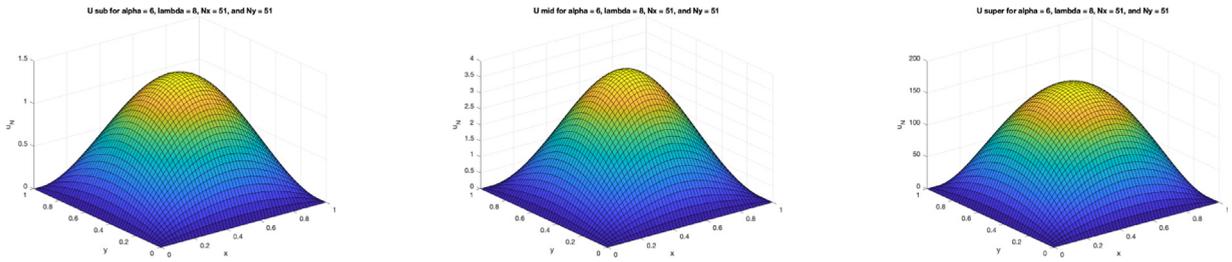


Fig. 9. The three solutions for Application 6.1 in two dimensions when $h_x = 1/51$, $h_y = 1/51$, $a = 6$, and $\lambda = 8$.

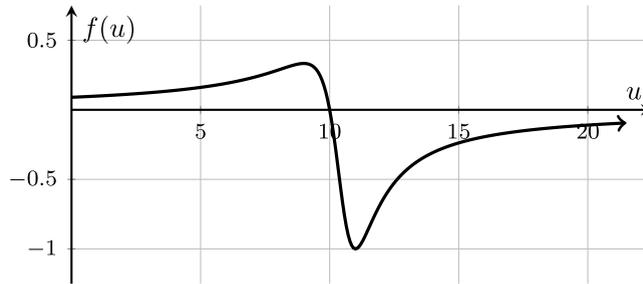


Fig. 10. The graph of f defined by (20) with $\beta = 1$ and $u_0 = 10$ which leads to an S-shaped bifurcation curve for (1) on the interval $\Omega = (0, 1)$.

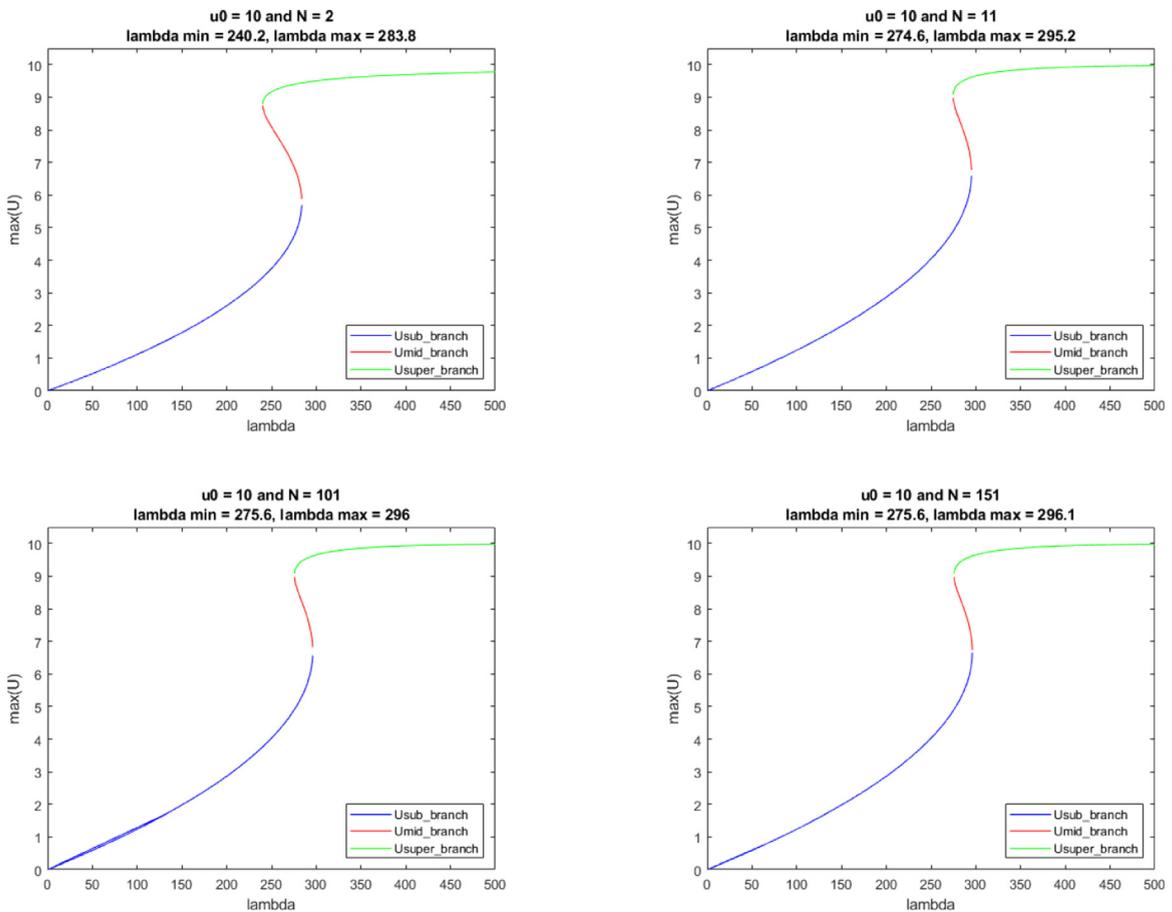


Fig. 11. Bifurcation curves for Application 6.2 when $N = 2$ (top left), $N = 11$ (top right), $N = 101$ (bottom left), and $N = 151$ (bottom right) with $\beta = 1$ and $u_0 = 10$. The turning points for $N = 2$ are at $\lambda = 240.2$ and $\lambda = 283.8$. The turning points for $N = 11$ are at $\lambda = 274.6$ and $\lambda = 295.2$. The turning points for $N = 101$ are at $\lambda = 275.6$ and $\lambda = 296$. The turning points for $N = 151$ are at $\lambda = 275.6$ and $\lambda = 296.1$.

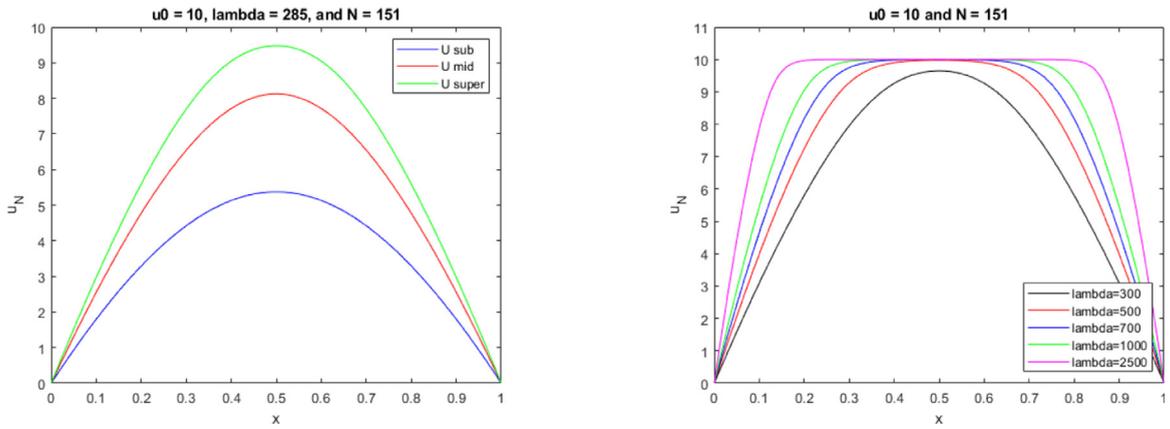


Fig. 12. Plots of solutions of Application 6.2 when $\beta = 1$ and $u_0 = 10$ for various values of λ . The left plot shows all three approximations when $\lambda = 285$. The right plot shows the unique solutions for $\lambda = 300, 500, 700, 1000, 2500$, where we can see that the solution converges towards the constant u_0 as $\lambda \rightarrow \infty$.

h	Umin Error	Order	Umid Error	Order	Umax Error	Order
2.00e-01	2.03e-01		7.48e-01		2.38e-01	
1.00e-01	4.49e-02	2.18	1.79e-01	2.06	8.42e-02	1.50
5.00e-02	1.08e-02	2.06	4.53e-02	1.98	2.23e-02	1.92
2.50e-02	2.67e-03	2.01	1.14e-02	2.00	5.57e-03	2.00
1.25e-02	6.66e-04	2.00	2.84e-03	2.00	1.39e-03	2.00
6.25e-03	1.66e-04	2.00	7.11e-04	2.00	3.48e-04	2.00
3.13e-03	4.16e-05	2.00	1.78e-04	2.00	8.70e-05	2.00
1.56e-03	1.04e-05	2.00	4.45e-05	2.00	2.17e-05	2.00

Fig. 13. Simulated Rates of convergence for all three solutions of Application 6.2 when $\beta = 1$, $u_0 = 10$, and $\lambda = 285$.

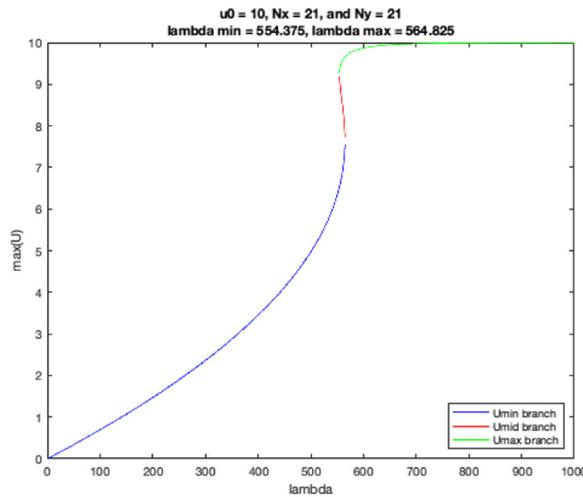


Fig. 14. Bifurcation curve for Application 6.2 in two dimensions when $h_x = 1/21$, $h_y = 1/21$, and $u_0 = 10$. The turning points are at $\lambda = 554.375$ and $\lambda = 564.825$.

with the reaction term

$$\tilde{f}(u) = \begin{cases} -c & \text{if } u < 0, \\ \kappa u + u(M - u) - c & \text{if } 0 \leq u \leq M, \\ \kappa M - c & \text{if } u > M \end{cases} \tag{23}$$

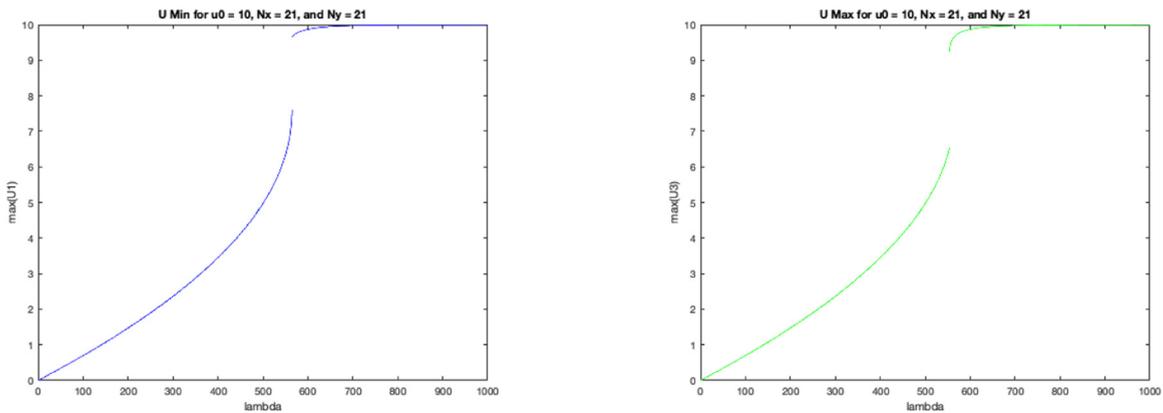


Fig. 15. Bifurcation curves generated using the monotone iteration in Section 5 for Application 6.2 in two dimensions when $h_x = 1/21, h_y = 1/21,$ and $u_0 = 10$. The first graph uses initial guesses based on subsolutions. The second graph uses initial guesses based on supersolutions.

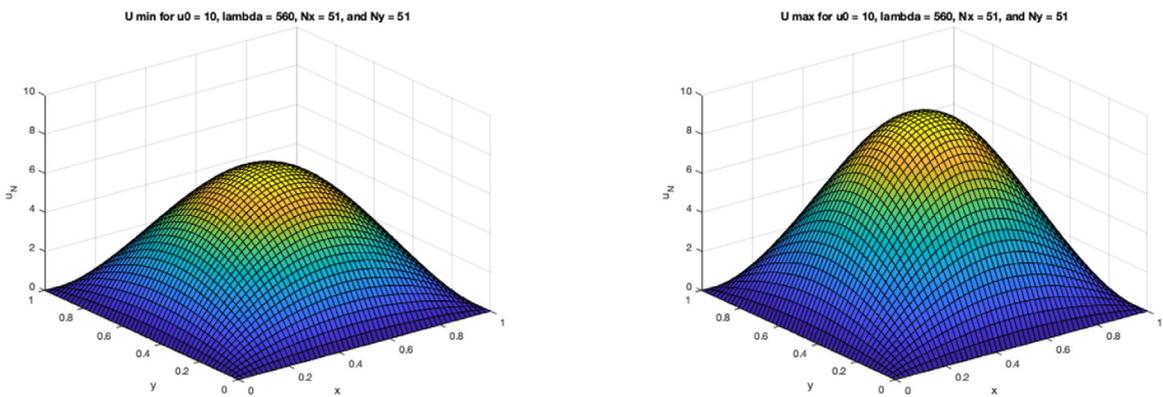


Fig. 16. The maximal and minimal solutions for Application 6.2 in two dimensions when $h_x = 1/51, h_y = 1/51, u_0 = 10,$ and $\lambda = 560$.

for $\kappa \geq M$ and note that (22) and (23) are semipositone with $-\Delta$ replaced by the elliptic linear operator \mathcal{L} . A graph of (21) for $M = 5$ and $c = 1$ and the corresponding graph of (23) for $M = 5, c = 1,$ and $\kappa = 6$ can be found in Fig. 17.

Since the discrete operator L_h corresponding to \mathcal{L} is monotone, we can straightforwardly extend the convergence result in Theorem 3.1; the admissibility, stability, and nonexistence results for semipositone problems in Section 4; and the solver in Section 5 to the simple finite difference scheme (4) with $-\Delta_h$ replaced by L_h and f replaced by \tilde{f} . Due to the constant extension used in (23) and the fact that solutions to the FD method $L_h U_\alpha = C$ paired with zero Dirichlet boundary conditions are uniformly bounded independent of h , we can construct a uniformly bounded positive discrete supersolution similar to how we constructed the discrete subsolution in (12). However, we can also choose $\bar{U} = M$ as the supersolution since $L_h M = \lambda \kappa M \geq \lambda \kappa M - \lambda c = \tilde{f}(M)$.

We could analogously choose \underline{U} to solve (12) and $\bar{U} = M$ and work with f directly as we did in Section 6.2 for the model of Kernevez et al. to directly show that the unmodified scheme has a solution U with $\underline{U} \leq U \leq \bar{U}$ and that the scheme has no nonnegative solution for λ sufficiently small. We use this approach in the numerical experiments.

6.3.1. Example 5: 1D logistic problem with harvesting

We approximate the logistic growth model with constant yield harvesting with $\Omega = (0, 1), M = 5,$ and $c = 1$. We see in Fig. 18 that the problem has no positive solutions for λ sufficiently small, two positive solutions for a certain range of λ values, and a unique positive solution for λ sufficiently large. In Fig. 19 we graph the solutions for various values of λ and observe the change from positive to sign-changing solutions for λ sufficiently large along the lower branch of the bifurcation curve.

7. Conclusion

The paper provides a blue print for proving admissibility and stability as well as convergence results when a PDE has multiple solutions. The admissibility and stability analysis assumed the underlying reaction–diffusion equation was positone or semipositone with sublinear growth. Assuming the underlying FD scheme is stable, the convergence proof

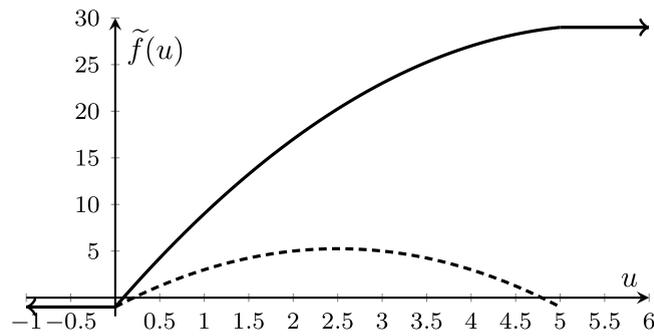


Fig. 17. The graph of f and \tilde{f} defined by (21) and (23) with $M = 5$, $c = 1$, and $\kappa = 6$.

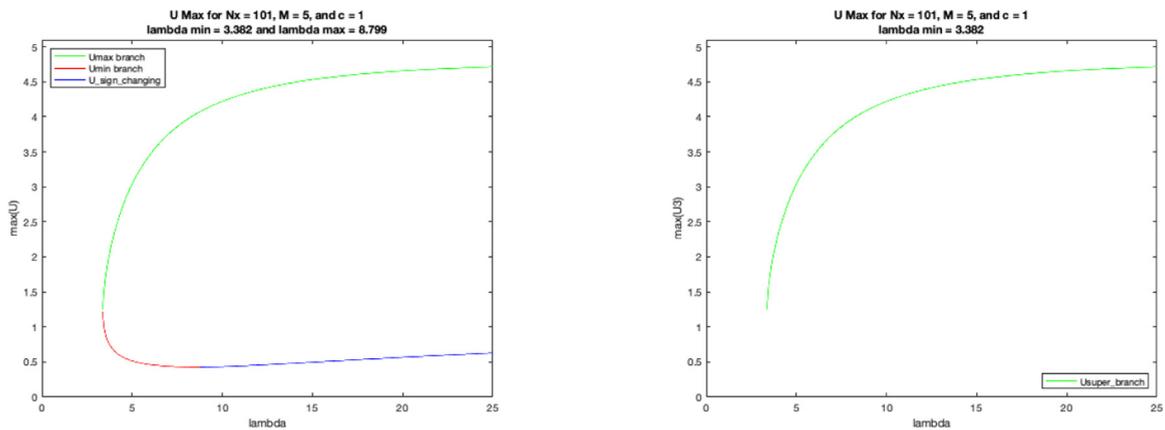


Fig. 18. Bifurcation curves for Application 6.3 when $N = 101$ with $M = 5$ and $c = 1$. The green curve represents the maximal solutions while the red curve represents the nonuniqueness of positive solutions. The blue curve corresponds to sign-changing solutions after the termination of the red branch. The critical λ values are $\lambda = 3.382$ and $\lambda = 8.799$. The second plot is generated using the monotone iteration in Section 5 to capture only the maximal solutions and determine the critical value for λ for nonexistence.

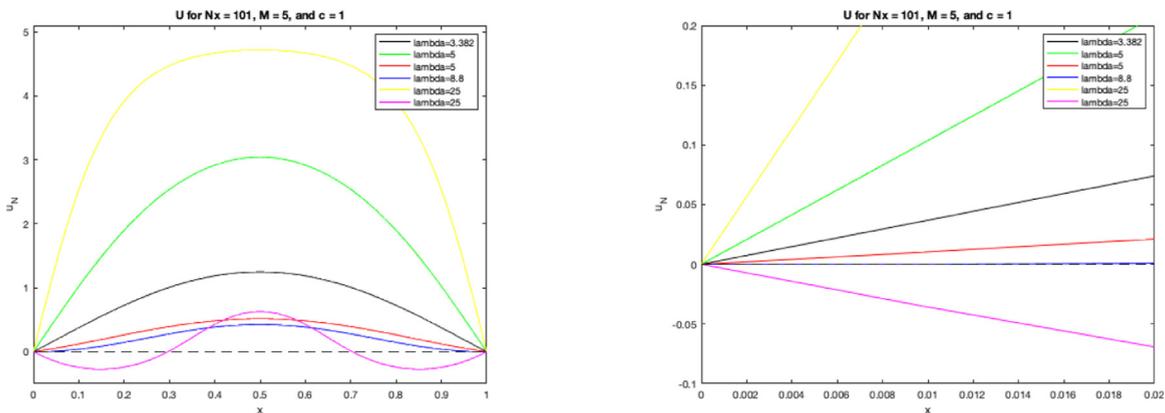


Fig. 19. Plots of solutions for Application 6.3 when $N = 101$ with $M = 5$ and $c = 1$ for various λ values. The black curve is the solution when $\lambda = 3.382$. The green and red curves are both for $\lambda = 5$. The blue curve is the positive solution for $\lambda = 8.8$. The yellow and pink curves are for $\lambda = 25$ where the yellow curve is the positive solution and the pink curve is a sign-changing solution. The right plot shows how the semipositone structure forces the outward normal derivative to transition from negative to positive as the positive solutions transform to sign-changing solutions along the lower branch of the bifurcation curve.

only required continuity of the reaction term. Thus, the convergence analysis applies to a much wider class of semilinear problems, and, going forward, only the admissibility and stability of the scheme needs to be verified for more general choices of f . The paper also provides a guaranteed solver that can find nonnegative solutions or show that no such solution

exists. Initial guesses are not restricted to be in a neighborhood of a PDE solution making the methods much more robust for application problems, and the initial points are easy to construct.

The convergence analysis provided in Section 3 is sufficiently flexible so as to be applicable to many classes of problems for which monotone approximation methods exist and for which stability bounds can be derived using the notion of discrete sub- and supersolutions. Given that monotone FD methods exist for fully nonlinear degenerate elliptic problems [30,42], nonlocal equations [43], and problems with nonlinear boundary conditions [44] (to name a few), it is possible that our results may be extended to these cases as well. Further, for any theoretical setting in which a method of sub- and supersolutions exists, it may be possible to replicate our results.

Nonlinear reaction–diffusion problems naturally arise in mathematical biology and many other applications. Non-uniqueness can be a major hurdle when trying to approximate such problems, especially when only the positive solution is relevant. This paper shows that simple FD methods yield convergent approximations that can reliably be used in applications. The paper also provides a solver that can test for uniqueness and nonexistence. Thus, the simple FD methods can also be used to reliably and efficiently generate bifurcation curves for studying existence and uniqueness results of the underlying PDE. In a future work we plan to extend the work to radial finite difference methods and finite element methods to better handle a more general class of domains.

References

- [1] P.L. Lions, On the existence of positive solutions for semilinear elliptic equations, *SIAM Rev.* 24 (1982) 442–467.
- [2] A. Castro, R. Shivaji, Non-negative solutions for a class of non-positone problems, *Proc. Roy. Soc. Edinburgh* 108A (1988) 291–302.
- [3] R. Shivaji, Uniqueness results for a class of positone problems, *Nonlinear Anal.* 7 (2) (1983) 223–230.
- [4] A. Castro, R. Shivaji, Uniqueness of positive solutions for a class of elliptic boundary value problems, *Proc. Roy. Soc. Edinburgh Sect. A* 98A (1984) 267–269.
- [5] G. Dragoni, Il problema dei valori ai limiti studiato in grande per gli integrali di una equazione differenziale del secondo ordine, *G. Mat. Battaglini* 69 (1931) 77–112.
- [6] C. De Coster, P. Habets, Upper and lower solutions in the theory of ODE boundary value problems: Classical and recent results, in: F. Zanolin (Ed.), *Non Linear Analysis and Boundary Value Problems for Ordinary Differential Equations*, Springer Vienna, Vienna, 1996, pp. 1–78, http://dx.doi.org/10.1007/978-3-7091-2680-6_1.
- [7] C.V. Pao, Block monotone iterative methods for numerical solutions of nonlinear elliptic equations, *Numer. Math.* 72 (1995) 239–262.
- [8] C.V. Pao, Monotone iterative methods for finite difference system of reaction-diffusion equations, *Numer. Math.* 46 (4) (1985) 571–586.
- [9] C.V. Pao, Comparison methods and stability analysis of reaction diffusion systems, in: *Comparison Methods and Stability Theory*, CRC Press, 1994, pp. 277–292.
- [10] X. Feng, R. Glowinski, M. Neilan, Recent developments in numerical methods for fully nonlinear second order partial differential equations, *SIAM Rev.* 55 (2013) 205–267.
- [11] J. Goddard II, Q. Morris, R. Shivaji, B. Son, Bifurcation curves for some singular and nonsingular problems with nonlinear boundary conditions, *Electron. J. Differential Equations* 2018 (26) (2018) 1–12.
- [12] P. Korman, Y. Li, T. Ouyang, A simplified proof of a conjecture for the perturbed Gelfand equation from combustion theory, *J. Differential Equations* 263 (5) (2017) 2874–2885.
- [13] J.M. Neuberger, J.W. Swift, Newton's method and morse index for semilinear elliptic PDEs, *Internat. J. Bifur. Chaos Appl. Sci. Engrg.* 11 (2001) 801–820.
- [14] T. Laetsch, The number of solutions of a nonlinear two point boundary value problem, *Indiana Univ. Math. J.* 20 (1970/71) 1–13.
- [15] B. Gidas, W.-M. Ni, L. Nirenberg, Symmetry and related properties via the maximum principle, *Comm. Math. Phys.* 68 (1979) 209–243.
- [16] M. Crouzeix, J. Rappaz, On Numerical Approximation in Bifurcation Theory, in: *RMA Collection Recherches en Mathématiques Appliquées*, Springer, 1990.
- [17] Y.S. Choi, P.J. McKenna, A mountain pass method for the numerical solutions of semilinear elliptic problems, *Nonlinear Anal. TMA* 20 (1993) 417–437.
- [18] Z.H. Ding, D. Costa, G. Chen, A high-linking algorithm for sign-changing solutions of semilinear elliptic equations, *Nonlinear Anal. TMA* 38 (1999) 151–172.
- [19] J.M. Neuberger, N. Sieben, J.W. Swift, Symmetry and automated branch following for a semilinear elliptic PDE on a fractal region, *SIAM J. Appl. Dyn. Syst.* 5 (2006) 476–507.
- [20] Y. Li, J.X. Zhou, A minimax method for finding multiple critical points and its application to semilinear PDEs, *SIAM J. Sci. Comput.* 23 (2002) 840–865.
- [21] C. Chen, Z. Xie, Search-extension method for multiple solutions of nonlinear problem, *Comput. Math. Appl.* 47 (2004) 327–343.
- [22] Z. Xie, C. Chen, Y. Xu, An improved search-extension method for computing multiple solutions of semilinear PDEs, *IMA J. Numer. Anal.* 25 (2005) 549–576.
- [23] X. Zhang, J. Zhang, B. Yu, Eigenfunction expansion method for multiple solutions of semilinear elliptic equations with polynomial nonlinearity, *SIAM J. Numer. Anal.* 51 (2013) 2680–2699.
- [24] E.L. Allgower, D.J. Bates, A.J. Sommese, C.W. Wampler, Solution of polynomial system derived from differential equations, *Computing* 76 (2006) 1–10.
- [25] E.L. Allgower, S.G. Cruceanu, S. Tavener, Application of numerical continuation to compute all solutions of semilinear elliptic equations, *Adv. Geom.* 9 (2009) 371–400.
- [26] W. Hao, J.D. Hauenstein, B. Hu, A.J. Sommese, A bootstrapping approach for computing multiple solutions of differential equations, *J. Comput. Appl. Math.* 258 (2014) 181–190.
- [27] H. Zhu, Z. Li, Z. Yang, Newton's method based on bifurcation for solving multiple solutions of nonlinear elliptic equations, *Math. Methods Appl. Sci.* 36 (2013) 2208–2223.
- [28] P.E. Farrell, A. Birkinsson, S.W. Funke, Deflation techniques for finding distinct solutions of nonlinear partial differential equations, *SIAM J. Sci. Comput.* 37 (2015) A2026–A2045.
- [29] Y. Wang, W. Hao, G. Lin, Two-level spectral methods for nonlinear elliptic equations with multiple solutions, *SIAM J. Sci. Comput.* 40 (2018) B1180–B1205.
- [30] G. Barles, P.E. Souganidis, Convergence of approximation schemes for fully nonlinear second order equations, *Asymptotic Anal.* 4 (1991) 271–283.

- [31] M.G. Crandall, H. Ishii, P.L. Lions, User's guide to viscosity solutions of second order partial differential equations, *Bull. Amer. Math. Soc.* 27 (1992) 1–67.
- [32] X. Feng, T. Lewis, A narrow-stencil finite difference method for approximating viscosity solutions of Hamilton-Jacobi-Bellman equations, *SIAM J. Numer. Anal.* 59 (2021) 886–924.
- [33] R.J. Plemmons, M-matrix characterizations. I – nonsingular M-matrices, *Linear Algebra Appl.* 18 (1977) 175–188.
- [34] H. Amann, Fixed point equations and nonlinear eigenvalue problems in ordered Banach spaces, *SIAM Rev.* 18 (1976) 620–709.
- [35] R. Seydel, *Practical Bifurcation and Stability Analysis*, in: *Interdisciplinary Applied Mathematics*, Springer-Verlag, New York, 2010.
- [36] H.B. Keller, Numerical solution of bifurcation and nonlinear eigenvalue problems, in: *Applications of Bifurcation Theory (Proc. Advanced Sem., Univ. Wisconsin, Madison, Wis., 1976)*, Publ. Math. Res. Center, 1977, pp. 359–384, No. 38.
- [37] K.J. Brown, M.M.A. Ibrahim, R. Shivaji, S-shaped bifurcation curves, *Nonlinear Anal.* 5 (5) (1981) 475–486.
- [38] T. Boddington, P. Gray, C. Robinson, Thermal explosion and the disappearance of criticality at small activation energies: Exact results for the slab, *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* 368 (1979) 441–461.
- [39] K.-C. Hung, S.-H. Wang, A theorem on S-shaped bifurcation curve for a positone problem with convex-concave nonlinearity and its applications to the perturbed Gelfand problem, *J. Differential Equations* 251 (2011) 223–237.
- [40] J.P. Kernevez, G. Joly, M.C. Duban, B. Bunow, D. Thomas, Hysteresis, oscillations and pattern formation in realistic immobilized enzyme systems, *J. Math. Biol.* 7 (1979) 41–56.
- [41] S. Oruganti, J. Shi, R. Shivaji, Diffusive logistic equation with constant yield harvesting. I. Steady states, *Trans. Amer. Math. Soc.* 354 (9) (2002) 3601–3619.
- [42] A. Oberman, Convergent difference schemes for degenerate elliptic and parabolic equations: Hamilton-Jacobi equations and free boundary problems, *SIAM J. Numer. Anal.* 44 (2) (2006) 879–895.
- [43] C.V. Pao, Monotone iterative methods for numerical solutions of nonlinear integro-elliptic boundary problems, *Appl. Math. Comput.* 186 (2) (2007) 1624–1642.
- [44] C.V. Pao, X. Lu, Block monotone iterative method for semilinear parabolic equations with nonlinear boundary conditions, *SIAM J. Numer. Anal.* 47 (6) (2010) 4581–4606.