NUMERICAL VARIATIONAL METHODS APPLIED TO CYLINDER BUCKLING[*]

JIŘÍ HORÁK[†], GABRIEL J. LORD[‡], AND MARK A. PELETIER[§]

Abstract. We review and compare different computational variational methods applied to a system of fourth order equations that arises as a model of cylinder buckling. We describe both the discretization and implementation, in particular how to deal with a one-dimensional null space. We show that we can construct many different solutions from a complex energy surface. We examine numerically convergence in the spatial discretization and in the domain size. Finally we give a physical interpretation of some of the solutions found.

1. Introduction. We describe complementary approaches to finding solutions of systems of fourth order elliptic PDEs. The techniques are applied to a problem that arises in the classic treatment of an isotropic cylindrical shell under axial compression, but they are also applicable to a wide range of problems such as waves on a suspension bridge [4, 5], the Fučík spectrum of the Laplacian [6], and the formation of microstructure [12, 3].

The cylindrical shell offers a computationally challenging and physically relevant problem with a complex energy surface. We take as our model for the shell the Von Kármán–Donnell equations, which can be rescaled (see the appendix and [7]) to the form

$$\Delta^2 w + \lambda w_{xx} - \phi_{xx} - 2\,[w, \phi] = 0, \tag{1.1}$$
$$\Delta^2 \phi + w_{xx} + [w, w] = 0, \tag{1.2}$$

where the bracket is defined as

$$[u, v] = \frac{1}{2} u_{xx} v_{yy} + \frac{1}{2} u_{yy} v_{xx} - u_{xy} v_{xy}. \tag{1.3}$$

The function w is a scaled inward radial displacement measured from the unbuckled (fundamental) state, ϕ is the Airy stress function, and $\lambda \in (0, 2)$ is a load parameter. The unknowns w and ϕ are defined on a two-dimensional spatial domain $\Omega = (-a, a) \times (-b, b)$, where $x \in (-a, a)$ is the axial and $y \in (-b, b)$ is the tangential coordinate. Since the y-domain $(-b, b)$ represents the circumference of the cylinder, the following boundary conditions are prescribed:

$$w \text{ is periodic in } y \quad \text{and} \quad w_x = (\Delta w)_x = 0 \text{ at } x = \pm a, \tag{1.4a}$$
$$\phi \text{ is periodic in } y \quad \text{and} \quad \phi_x = (\Delta \phi)_x = 0 \text{ at } x = \pm a, \tag{1.4b}$$

as shown in Figure 1 (i), (ii).

[†]Department of Mathematics, Universität Köln, 50931 Köln, Germany (jhorak@math.uni-koeln.de).

[‡]Department of Mathematics, Heriot-Watt University, Edinburgh EH14 4AS, UK (g.j.lord@ma.hw.ac.uk).

[§]Department of Mathematics and Computer Science, Technische Universiteit Eindhoven, Eindhoven S60 OMB, The Netherlands (m.a.peletier@tue.nl).
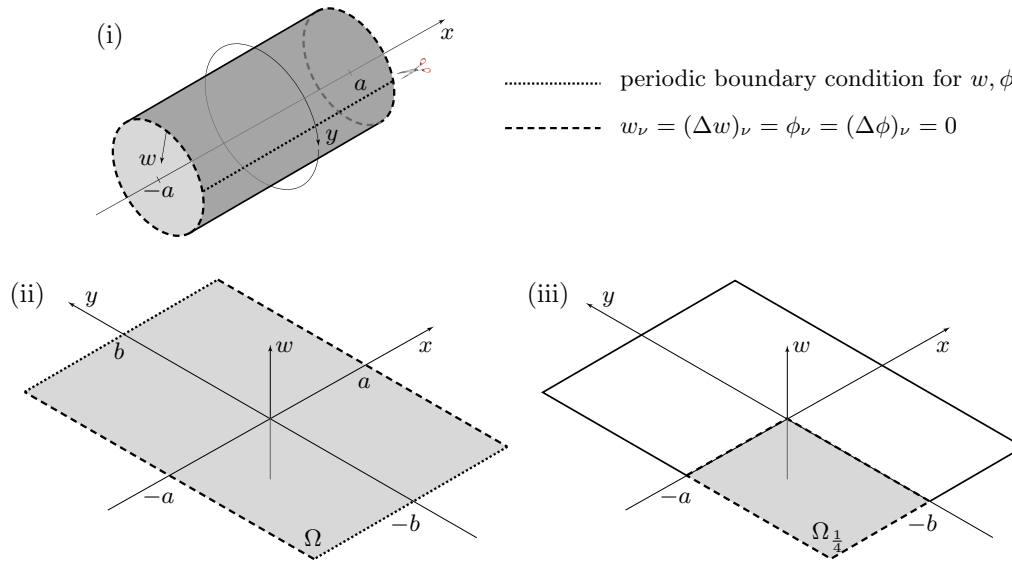
FIG. 1. (i) *The geometry of the cylinder;* (ii) *the computational domain and the boundary conditions;* (iii) *one quarter of the domain and the corresponding boundary conditions.*

1.1. Functional setting. We search for weak solutions w, ϕ of (1.1)–(1.4) in the space

$$X = \left\{ \psi \in H^2(\Omega) : \psi_x(\pm a, \cdot) = 0, \ \psi \text{ is periodic in } y, \text{ and } \int_\Omega \psi = 0 \right\}$$

with the norm

$$\|w\|_X^2 = \int_\Omega \left(\Delta w^2 + \Delta \phi_1^2 \right),$$

where $\phi_1 \in H^2(\Omega)$ is the unique solution of

$$(1.5) \qquad \Delta^2 \phi_1 = -w_{xx}, \qquad \phi_1 \text{ satisfies (1.4b)}, \qquad \text{and} \qquad \int_\Omega \phi_1 = 0.$$

This norm is equivalent to the H^2-norm on X, and with the appropriate inner product $\langle \cdot, \cdot \rangle_X$ the space X is a Hilbert space. Alternatively, if the load parameter $\lambda \in (0, 2)$ is fixed, another norm,

$$\|w\|_{X,\lambda}^2 = \int_\Omega \left(\Delta w^2 + \Delta \phi_1^2 - \lambda w_x^2 \right),$$

can be used. Because of the estimate

$$\int_\Omega w_x^2 = -\int_\Omega w w_{xx} = \int_\Omega w \Delta^2 \phi_1 = \int_\Omega \Delta w \Delta \phi_1 \leq \frac{1}{2} \int_\Omega \Delta w^2 + \frac{1}{2} \int_\Omega \Delta \phi_1^2 = \frac{1}{2} \|w\|_X^2,$$

it is equivalent to $\|\cdot\|_X$ and hence also to the H^2-norm on X. The corresponding inner product will be denoted $\langle \cdot, \cdot \rangle_{X,\lambda}$.

Equations (1.1)–(1.2) are related to the stored energy E, the average axial shortening S, and the total potential given by

$$(1.6) \quad E(w) := \frac{1}{2} \int_\Omega \left(\Delta w^2 + \Delta \phi^2 \right), \qquad S(w) := \frac{1}{2} \int_\Omega w_x^2, \qquad F_\lambda = E - \lambda S.$$

Note that the function ϕ in (1.6) is determined from w by solving (1.2) with boundary conditions (1.4b). All the functionals E, S, and F_λ belong to $C^1(X)$, i.e., are continuously Fréchet differentiable.

The fact that (1.1) is a reformulation of the stationarity condition $F_\lambda' = E' - \lambda S' = 0$ will be important in section 2, and we therefore briefly sketch the argument. It is easy to see that

$$S'(w) \cdot h = -\int_\Omega w_{xx} h.$$

For $E'(w) \cdot h$, let $w, \phi \in X$ solve (1.2) and $h, \psi \in X$ solve $\Delta^2 \psi = -h_{xx} - [h, h] - 2[w, h]$. Then, assuming sufficient regularity on w,

$$E(w + h) - E(w) = \int_\Omega \Delta w \Delta h + \frac{1}{2} \int_\Omega (\Delta h)^2 + \int_\Omega \Delta \phi \Delta \psi + \frac{1}{2} \int_\Omega (\Delta \psi)^2$$

$$= \int_\Omega \left(h \Delta^2 w - h \phi_{xx} - 2[w, h] \phi \right) + \frac{1}{2} \int_\Omega (\Delta h)^2 + \frac{1}{2} \int_\Omega (\Delta \psi)^2 - \int_\Omega [h, h] \phi,$$

where we used integration by parts several times. The last three integrals are $O(\|h\|_X^2)$ for $\|h\|_X \to 0$, and it can be shown by integration by parts that

$$(1.7) \qquad\qquad \int_\Omega [w, h] \, \phi = \int_\Omega h \, [w, \phi].$$

Therefore

$$F_\lambda'(w) \cdot h = E'(w) \cdot h - \lambda S'(w) \cdot h = \int_\Omega h \left(\Delta^2 w + \lambda w_{xx} - \phi_{xx} - 2 \left[w, \phi \right] \right).$$

1.2. Review of some variational numerical methods. We now describe the variational methods used to find numerical approximations of critical points of the total potential F_λ. In our numerical experiments these methods are accompanied by Newton's method and continuation. The advantage of this approach is that it combines the knowledge of global features of the energy landscape with local ones of a neighborhood of a critical point. The details related to spatial discretization will be discussed in section 2, the Newton-based methods in section 3.

1.2.1. Steepest descent method (SDM). Let the load parameter $\lambda \in (0, 2)$ be fixed; we work in a discretized version of $(X, \langle \cdot, \cdot \rangle_{X,\lambda})$. We try to minimize the total potential F_λ by following its gradient flow. We solve the initial value problem

$$\frac{d}{dt} w(t) = -\nabla_\lambda F_\lambda(w(t)), \qquad w(0) = w_0,$$

with a suitable starting point w_0 on some interval $(0, T]$. This problem is then discretized in t.

In [7] it was shown that $w = 0$ is a local minimizer of F_λ. Indeed, if $\|w_0\|_{X,\lambda}$ is small, the numerical solution $w(t)$ converges to zero as t tends to infinity. If, on
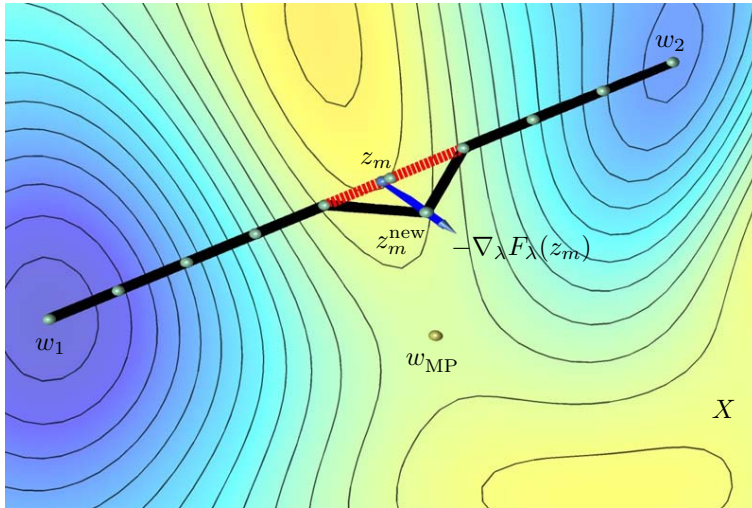
FIG. 2. *Deforming the path in the main loop of the MPA: point z_m is moved a small distance in the direction $-\nabla_\lambda F_\lambda(z_m)$ and becomes z_m^{new}. This step is repeated until the mountain-pass point w_{MP} is reached.*

the other hand, $\|w_0\|_{X,\lambda}$ is large, the numerical solution $w(t)$ stays bounded away from zero. In most of our experiments, the numerical algorithm did not converge for $t \to \infty$ in the large norm case. The only exception for a relatively small value of λ will be mentioned later in section 5.3. Nevertheless, for a sufficiently large computational domain Ω and a sufficiently large $t > 0$ we obtain $F_\lambda(w(t)) < 0$. Such a state $w(t)$ is needed for the mountain-pass algorithm as explained below. Existence of this state was also proved in [7].

1.2.2. Mountain-pass algorithm (MPA). The algorithm was first proposed in [2] for a second order elliptic problem in one dimension and extended in [5] to a fourth order problem in two dimensions. We give a brief description of the algorithm here.

Let the load parameter $\lambda \in (0, 2)$ be fixed; we work again in a discretized version of $(X, \langle \cdot, \cdot \rangle_{X,\lambda})$. We denote $w_1 = 0$ the local minimum of F_λ and take a point w_2 such that $F_\lambda(w_2) < 0$ (in practice this point is found using the SDM). We take a discretized path $\{z_m\}_{m=0}^p$ connecting $z_0 = w_1$ with $z_p = w_2$. After finding the point z_m at which F_λ is maximal along the path, this point is moved a small distance in the direction of the steepest descent $-\nabla_\lambda F_\lambda(z_m)$. Thus the path has been deformed and the maximum of F_λ lowered. This deforming of the path is repeated until the maximum along the path cannot be lowered anymore: a critical point w_{MP} has been reached. Figure 2 illustrates the main idea of the method.

The MPA is local in its nature. The numerical solution w_{MP} it finds has the mountain-pass property in a certain neighborhood only. The choice of the path endpoint w_2 may influence to which critical point the algorithm converges. Different choices of w_2 are in turn achieved by choosing different initial points w_0 in the SDM.

1.2.3. Constrained steepest descent method (CSDM). We fix the amount of shortening S of the cylinder. This is often considered as what actually occurs in experiments. We work now in a discretized version of $(X, \langle \cdot, \cdot \rangle_X)$. Let $C > 0$ be a

fixed number and define a set of w with constant shortening:

$$(1.8) \qquad \mathscr{M} = \{w \in X \; : \; S(w) = C\}.$$

Critical points of E under this constraint are critical points of F_λ, where λ is a Lagrange multiplier. The simplest such points are local minima of the stored energy E on $\mathscr{M}$. We need to follow the gradient flow of E on $\mathscr{M}$; hence we solve the initial value problem

$$\frac{d}{dt}w(t) = -P_{w(t)}\nabla E(w(t)), \qquad w(0) = w_0 \in \mathscr{M}$$

for $t > 0$. P_w denotes the orthogonal projection in X on the tangent space of $\mathscr{M}$ at $w \in \mathscr{M}$:

$$P_w u = u - \frac{\langle \nabla S(w), u \rangle_X}{\|\nabla S(w)\|_X^2} \nabla S(w).$$

The details of the algorithm can be found in [4]. The initial value problem is solved by repeating the following two steps: given a point $w \in \mathscr{M}$ find $\bar{w} = w - \Delta t P_w \nabla E(w)$, where $\Delta t > 0$ is small, and define $w_{\text{new}} = c\bar{w}$, where the scaling coefficient c is chosen so that $w_{\text{new}} \in \mathscr{M}$. The algorithm is stopped when $\|P_w \nabla E(w)\|_X$ is smaller than a prescribed tolerance. The corresponding load is given by

$$\lambda = \frac{\langle \nabla S(w), \nabla E(w) \rangle_X}{\|\nabla S(w)\|_X^2}.$$

1.2.4. Constrained mountain-pass algorithm (CMPA). Let $C > 0$ and let $\mathscr{M}$ be the set of w with constant shortening given in (1.8). We would like to find mountain-pass points of E on $\mathscr{M}$. The method has been described in [4] in detail. We need two local minima w_1, w_2 of E on $\mathscr{M}$, which can be found using the CSDM. The algorithm is then similar to the MPA. We take a discretized path $\{z_m\}_{m=0}^p \subset \mathscr{M}$ connecting $z_0 = w_1$ with $z_p = w_2$. After finding the point z_m at which E is maximal along the path, this point is moved a small distance in the tangent space to $\mathscr{M}$ at z_m in the direction of the steepest descent $-P_{z_m}\nabla E(z_m)$ and then scaled (as in the CSDM) to come back to $\mathscr{M}$. Thus the path has been deformed on $\mathscr{M}$ and the maximum of E lowered. This deforming of the path is repeated until the maximum along the path cannot be lowered anymore: a mountain-pass point of E on $\mathscr{M}$ has been reached. The load λ is computed as in the CSDM.

The choice of the endpoints w_1 and w_2 will in general influence to which critical point the algorithm converges.

1.3. Computational domains. We consider the problem on the domain Ω (Figure 1 (ii)) both without further restraints and under a symmetry assumption, which reduces the computational complexity. In the latter case we assume

$$(1.9) \qquad \begin{aligned} w(x,y) &= w(-x,y) = w(x,-y) \\ \phi(x,y) &= \phi(-x,y) = \phi(x,-y) \end{aligned} \qquad \text{for } (x,y) \in \Omega.$$

By looking for solutions $w, \phi \in X$ that satisfy (1.9) we effectively reduce the domain to one quarter: $\Omega_{\frac{1}{4}} = (-a,0) \times (-b,0)$ as shown in Figure 1 (iii). One needs to solve (1.1)–(1.2) only on $\Omega_{\frac{1}{4}}$ with the boundary conditions

$$(1.10) \qquad w_\nu = (\Delta w)_\nu = \phi_\nu = (\Delta \phi)_\nu = 0 \quad \text{on } \partial\Omega_{\frac{1}{4}},$$

where ν denotes the outward normal direction to the boundary. Hence we search for weak solutions of (1.1)–(1.2), (1.10) in the space

$$X_{\frac{1}{4}} = \left\{ \psi \in H^2(\Omega_{\frac{1}{4}}) : \psi_\nu = 0 \text{ on } \partial\Omega_{\frac{1}{4}}, \text{ and } \int_{\Omega_{\frac{1}{4}}} \psi = 0 \right\}.$$

We can then use (1.9) to extend these functions to the whole Ω.

We have performed numerical experiments both with and without the symmetry assumption. For the sake of simplicity we will give a detailed description of the numerical methods for the second case only where the boundary conditions are the same on all sides of $\Omega_{\frac{1}{4}}$. The first case with periodic conditions on two sides of Ω is very similar and will be briefly mentioned in Remark 2.1.

1.4. Solving the biharmonic equation. In order to obtain ϕ for a given w, one has to solve (1.2); to compute the norm of w, one has to solve (1.5). Both problems are of the form

$$(1.11) \qquad \Delta^2\psi = f \text{ in } \Omega_{\frac{1}{4}}, \qquad \psi_\nu = (\Delta\psi)_\nu = 0 \text{ on } \partial\Omega_{\frac{1}{4}}, \qquad \int_{\Omega_{\frac{1}{4}}} \psi = 0,$$

where $f \in L^1(\Omega_{\frac{1}{4}})$ is given. If $\int_{\Omega_{\frac{1}{4}}} f = 0$, then (1.11) has a unique weak solution ψ in $X_{\frac{1}{4}}$. It is a straightforward calculation to verify that the right-hand sides of equations in (1.2) and (1.5) have zero average.

In the discretization described below, problem (1.11) is treated as a system:

$$(1.12) \qquad \begin{array}{l} -\Delta u = f \\ -\Delta v = u \end{array} \text{ in } \Omega_{\frac{1}{4}}, \qquad u_\nu = v_\nu = 0 \text{ on } \partial\Omega_{\frac{1}{4}}, \qquad \int_{\Omega_{\frac{1}{4}}} u = \int_{\Omega_{\frac{1}{4}}} v = 0.$$

The system has a unique weak solution $(u,v) \in (H^1(\Omega_{\frac{1}{4}}))^2$. Since the domain $\Omega_{\frac{1}{4}}$ has no reentrant corners, Theorem 1.4.5 of [9] guarantees that $v \in H^2(\Omega_{\frac{1}{4}})$ and therefore that the two formulations are equivalent.

2. Finite difference discretization. We discretize the domain $\Omega_{\frac{1}{4}}$ by a uniform mesh $(x_m, y_n) \in \Omega_{\frac{1}{4}}$ with M points in the x-direction and N points in the y-direction:

$$\begin{aligned} x_m &= -a + (m - \tfrac{1}{2})\Delta x, & m &\in \{1 \dots, M\}, \\ y_n &= -b + (n - \tfrac{1}{2})\Delta y, & n &\in \{1 \dots, N\}, \end{aligned}$$

where $\Delta x = a/M$, $\Delta y = b/N$. We represent the values of some function w on $\Omega_{\frac{1}{4}}$ at these points by a vector $w = (w_i)_{i=1}^{MN}$, where $w_i = w(x_m, y_n)$ and $i = (n-1)M + m$. In our notation we will not distinguish between w as a function and w as a corresponding vector. The vector w can also be interpreted as a block vector with N blocks, each containing M values of a single row of the mesh. We introduce the following conventions for notation:

- For two matrices $A^M = (a_{ij})_{i,j=1}^M$ and $B^N = (b_{k\ell})_{k,\ell=1}^N$ we define $A^M \otimes B^N := (b_{k\ell} A^M)_{k,\ell=1}^N$, which is an $N \times N$ block matrix, and each block is an $M \times M$ matrix.
- For two vectors $u = (u_i)_{i=1}^{MN}$, $v = (v_i)_{i=1}^{MN}$ we define $u \odot v = (u_i v_i)_{i=1}^{MN}$, i.e., a product of the components.

To discretize second derivatives we use the standard central finite differences (with Neumann boundary conditions [10]). Let Id^M denote the $M \times M$ identity matrix and define another $M \times M$ matrix:

$$A_2^M = \begin{bmatrix} 1 & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ & & & -1 & 1 \end{bmatrix}.$$

The second derivatives $-\partial_{xx}$, $-\partial_{yy}$ and the biharmonic operator Δ^2 with the appropriate boundary conditions are approximated by

$$\boldsymbol{A}_{xx} = \frac{1}{\Delta x^2} A_2^M \otimes \mathrm{Id}^N, \qquad \boldsymbol{A}_{yy} = \frac{1}{\Delta y^2} \mathrm{Id}^M \otimes A_2^N, \qquad \boldsymbol{A}_{\Delta^2} = (\boldsymbol{A}_{xx} + \boldsymbol{A}_{yy})^2,$$

respectively.

2.1. Discretization of $\boldsymbol{E}$, $\boldsymbol{S}$, and the bracket $[\cdot, \cdot]$. Supposing that we can solve the discretized version of (1.2),

$$(2.1) \qquad\qquad \boldsymbol{A}_{\Delta^2} \phi - \boldsymbol{A}_{xx} w + [w, w]_2 = \boldsymbol{0},$$

we can also evaluate the energy E and the shortening S:

$$(2.2) \quad E(w) = 2 \left(w^T \boldsymbol{A}_{\Delta^2} w + \phi^T \boldsymbol{A}_{\Delta^2} \phi \right) \Delta x \, \Delta y, \qquad S(w) = 2 \left(w^T \boldsymbol{A}_{xx} w \right) \Delta x \, \Delta y.$$

In order to solve (2.1) we need to be able to solve the biharmonic equation and choose a discretization of the bracket $[\cdot, \cdot]$. This bracket appears in the equations in two different roles: in (1.2) the bracket is part of the mapping $w \mapsto \phi$, and therefore of the definition of the energy E; in (1.1), which represents the stationarity condition $E' - \lambda S' = 0$, the bracket appears as a result of differentiating E with respect to w and applying partial integration. As a result, we need to use two different forms of discretization for the two cases.

In both cases the bracket requires a discretization of the mixed derivative ∂_{xy}. One choice is to use one-sided finite differences. Define $M \times M$ matrices

$$(2.3) \qquad A_{1L}^M = \begin{bmatrix} 0 & & & \\ -1 & 1 & & \\ & \ddots & \ddots & \\ & & -1 & 1 \end{bmatrix}, \qquad A_{1R}^M = \begin{bmatrix} -1 & 1 & & \\ & \ddots & \ddots & \\ & & -1 & 1 \\ & & & 0 \end{bmatrix}.$$

We choose either left or right-sided differences represented by these matrices, respectively, and we let A_1^M denote our choice (cf. section 5.1). The derivatives ∂_x, ∂_y, and $-\partial_{xy}$ are approximated by

$$(2.4) \qquad \boldsymbol{A}_x = \frac{1}{\Delta x} A_1^M \otimes \mathrm{Id}^N, \qquad \boldsymbol{A}_y = \frac{1}{\Delta y} \mathrm{Id}^M \otimes A_1^N, \qquad \boldsymbol{A}_{xy} = -\boldsymbol{A}_x \boldsymbol{A}_y.$$

For the definition of ϕ in terms of w (see (1.2)) we choose

$$(2.5) \qquad\qquad [w, w]_2 = (\boldsymbol{A}_{xx} w) \odot (\boldsymbol{A}_{yy} w) - (\boldsymbol{A}_{xy} w) \odot (\boldsymbol{A}_{xy} w),$$

and the corresponding choice for (1.1) is

$$(2.6) \quad [w, \phi]_1 = \tfrac{1}{2} \boldsymbol{A}_{yy} \left\{ (\boldsymbol{A}_{xx} w) \odot \phi \right\} + \tfrac{1}{2} \boldsymbol{A}_{xx} \left\{ (\boldsymbol{A}_{yy} w) \odot \phi \right\} - \boldsymbol{A}_{xy}^T \left\{ (\boldsymbol{A}_{xy} w) \odot \phi \right\} .$$

These two definitions are related in the sense given in (1.7): $[w, h]_2^T \phi = h^T [w, \phi]_1$ for all h.

With these definitions the partial derivatives of discretized E and S with respect to the components of w are given by

$$(2.7) \qquad E'(w) = \boldsymbol{A}_{\Delta^2} w + \boldsymbol{A}_{xx} \phi - 2[w, \phi]_1 , \qquad S'(w) = \boldsymbol{A}_{xx} w .$$

2.2. Solving the discretized biharmonic equation. Matrix $\boldsymbol{A}_{\Delta^2}$ is symmetric and has a one-dimensional null space: $\boldsymbol{A}_{\Delta^2} \mathbf{1} = \mathbf{0}$, where $\mathbf{1}$ and $\mathbf{0}$ are vectors with MN-components which are all one and all zero, respectively. The same is true for $\boldsymbol{A}_{xx}$ and $\boldsymbol{A}_{yy}$. For a given vector f we would like to solve

$$(2.8) \qquad \boldsymbol{A}_{\Delta^2} \psi = f , \qquad \mathbf{1}^T \psi = 0 .$$

A unique solution exists if and only if f has zero average, i.e., $\mathbf{1}^T f = 0$. So we must verify that the discretized versions of the right-hand sides in (1.2), (1.5) satisfy this condition. Let $w \in \mathbb{R}^{MN}$; then

$$\mathbf{1}^T \boldsymbol{A}_{xx} w = 0 ,$$
$$(2.9) \qquad \mathbf{1}^T [w, w]_2 = (\boldsymbol{A}_{xx} w)^T (\boldsymbol{A}_{yy} w) - (\boldsymbol{A}_{xy} w)^T (\boldsymbol{A}_{xy} w)$$
$$(2.10) \qquad = w^T \boldsymbol{A}_{xx} \boldsymbol{A}_{yy} w - w^T \boldsymbol{A}_{xy}^T \boldsymbol{A}_{xy} w = 0 ,$$

where the last equality holds because $\boldsymbol{A}_x^T \boldsymbol{A}_x = \boldsymbol{A}_{xx}$ and $\boldsymbol{A}_y^T \boldsymbol{A}_y = \boldsymbol{A}_{yy}$, and because the x- and y-matrices commute. We have, in fact, shown that the integration-by-parts formula from the continuous case holds for our choice of spatial discretization. This is not true for an arbitrary discretization but is key for a successful scheme.

The inverse of matrix $\boldsymbol{A}_{\Delta^2}$ on the subspace of vectors with zero average, denoted with a slight abuse of notation by $\boldsymbol{A}_{\Delta^2}^{-1}$, can be found, for example, using the fast cosine transform described below in section 2.4.

2.3. Computing the gradient. The variational methods of this paper are based on a steepest descent flow and modifications of this algorithm. The direction of the steepest descent of E at a point $w \in X$ is opposite to the gradient of E at w. The gradient is the Riesz representative of the Fréchet derivative, and hence we need to find a vector $u \in X$, such that $E'(w) \cdot v = \langle u, v \rangle$ for all $v \in X$. The inner product is either $\langle \cdot, \cdot \rangle_X$ or $\langle \cdot, \cdot \rangle_{X,\lambda}$, and hence the gradient depends on the choice of the inner product. We use the notation $u = \nabla E(w)$ for the gradient in $(X, \langle \cdot, \cdot \rangle_X)$ and $u = \nabla_\lambda E(w)$ for the gradient in $(X, \langle \cdot, \cdot \rangle_{X,\lambda})$. To find the discretized version of the gradient, we first need to discretize the inner product.

Let $u, v \in \mathbb{R}^{MN}$, $\mathbf{1}^T u = \mathbf{1}^T v = 0$. The inner product is evaluated in the following way:

$$\langle u, v \rangle_{X,\lambda} = 4 \left(u^T \boldsymbol{A}_{\Delta^2} v + \phi_1^{u\,T} \boldsymbol{A}_{\Delta^2} \phi_1^v - \lambda u^T \boldsymbol{A}_{xx} v \right) \Delta x \, \Delta y$$

$$= 4 \left(u^T \left(\boldsymbol{A}_{\Delta^2} + \boldsymbol{A}_{xx} \boldsymbol{A}_{\Delta^2}^{-1} \boldsymbol{A}_{xx} - \lambda \boldsymbol{A}_{xx} \right) v \right) \Delta x \, \Delta y ,$$

where ϕ_1^u, ϕ_1^v are solutions of the discretized version of (1.5) with w replaced by u and v, respectively, and we assume that we work on $\Omega_{\frac{1}{4}}$. For $w \in \mathbb{R}^{MN}$, $\mathbf{1}^T w = 0$ the Riesz representative of $E'(w)$ given in (2.7) is computed as

$$(2.11) \qquad \nabla_\lambda E(w) = \left(\boldsymbol{A}_{\Delta^2} + \boldsymbol{A}_{xx} \boldsymbol{A}_{\Delta^2}^{-1} \boldsymbol{A}_{xx} - \lambda \boldsymbol{A}_{xx}\right)^{-1} E'(w).$$

As in the case of $\boldsymbol{A}_{\Delta^2}^{-1}$ we have abused notation here since the inverse makes sense only on a subspace of vectors with zero average. It can be easily verified that $\mathbf{1}^T E'(w) = 0$. The numerical evaluation of $\nabla_\lambda S$ and of the $\langle \cdot, \cdot \rangle_X$-gradients is similar.

2.4. Fourier coordinates. In Fourier coordinates most of the finite difference operators become diagonal matrices. This increases the efficiency of the numerical algorithm and makes it possible to easily find the inverse of matrices like $\boldsymbol{A}_{\Delta^2}$. See, for example, [1].

On a uniform grid, it is standard procedure to apply some form of the fast Fourier transform to diagonalize finite difference matrices like A_2^M (see, e.g., [11]). Due to the Neumann boundary conditions (1.10) we need to employ the fast cosine transform. We define the $M \times M$ matrices

$$C_{\mathrm{f}}^M = \frac{1}{\sqrt{2M}}\left(2\cos\frac{(i-1)(2j-1)\pi}{2M}\right)_{i,j=1}^M, \quad C_{\mathrm{b}}^M = \frac{1}{\sqrt{2M}}\left[\begin{array}{c|c} 1 & \\ \vdots & \left(2\cos\frac{(2i-1)(j-1)\pi}{2M}\right)_{i=1,j=2}^{M,M} \\ 1 & \end{array}\right],$$

which have the following properties:

$$C_{\mathrm{f}}^M C_{\mathrm{b}}^M = \mathrm{Id}^M, \qquad C_{\mathrm{f}}^M A_2^M C_{\mathrm{b}}^M = \Lambda^M,$$

where $\Lambda^M = \mathrm{diag}(2 - 2\cos\frac{(m-1)\pi}{M})_{m=1}^M$. Hence they are inverses of each other and diagonalize A_2^M.

We further define the matrices

$$\boldsymbol{C}_{\mathrm{f}} = C_{\mathrm{f}}^M \otimes C_{\mathrm{f}}^N, \qquad \boldsymbol{C}_{\mathrm{b}} = C_{\mathrm{b}}^M \otimes C_{\mathrm{b}}^N,$$

which diagonalize $\boldsymbol{A}_{xx}$, $\boldsymbol{A}_{yy}$, and $\boldsymbol{A}_{\Delta^2}$:

$$(2.12) \qquad \boldsymbol{C}_{\mathrm{f}}\boldsymbol{A}_{xx}\boldsymbol{C}_{\mathrm{b}} = \boldsymbol{\Lambda}_{xx}, \qquad \boldsymbol{C}_{\mathrm{f}}\boldsymbol{A}_{yy}\boldsymbol{C}_{\mathrm{b}} = \boldsymbol{\Lambda}_{yy}, \qquad \boldsymbol{C}_{\mathrm{f}}\boldsymbol{A}_{\Delta^2}\boldsymbol{C}_{\mathrm{b}} = \boldsymbol{\Lambda}_{\Delta^2},$$

where the diagonal matrices are given by

$$(2.13) \quad \boldsymbol{\Lambda}_{xx} = \frac{1}{\Delta x^2}\Lambda^M \otimes \mathrm{Id}^N, \qquad \boldsymbol{\Lambda}_{yy} = \frac{1}{\Delta y^2}\mathrm{Id}^M \otimes \Lambda^N, \qquad \boldsymbol{\Lambda}_{\Delta^2} = (\boldsymbol{\Lambda}_{xx} + \boldsymbol{\Lambda}_{yy})^2.$$

For a vector $w \in \mathbb{R}^{MN}$ we introduce its Fourier coordinates $\hat{w}$ by

$$\hat{w} = \boldsymbol{C}_{\mathrm{f}}w \qquad\qquad w = \boldsymbol{C}_{\mathrm{b}}\hat{w}.$$

We note that $\mathbf{1}^T w = 0$ if and only if the first component of $\hat{w}$ is zero.

Most of the computations involved in the variational methods described in section 1.2 can be done in the Fourier coordinates. One needs to go back to the original coordinates only when evaluating the brackets (2.5) and (2.6), because they are nonlinear and involve the discretized mixed derivative operator $\boldsymbol{A}_{xy}$.

2.5. Alternative discretization of $-\partial_{xy}$. The fast Fourier transform provides us with another discretization of the mixed derivative which is not biased to the left or right. In an analogy to (2.13) and (2.12) we define

$$\mathbf{\Lambda}_{xy} = \tfrac{1}{\Delta x \,\Delta y}\sqrt{\Lambda^M} \otimes \sqrt{\Lambda^N}, \qquad \tilde{\mathbf{A}}_{xy} = \mathbf{S}\mathbf{\Lambda}_{xy}\mathbf{C}_{\mathrm{f}},$$

where $\mathbf{S}$ is the fast sine transform matrix

$$\mathbf{S} = S^M \otimes S^N, \quad S^M = \tfrac{1}{\sqrt{2M}}\left[\begin{array}{c|c} 0 & \\ \vdots & \left(2\sin\frac{(2i-1)(j-1)\pi}{2M}\right)^{M,M}_{i=1,j=2} \\ 0 & \end{array}\right].$$

Property (2.10) also holds with $\mathbf{A}_{xy}$ replaced by $\tilde{\mathbf{A}}_{xy}$.

Remark 2.1. When discretizing the problem on the full domain Ω with boundary conditions (1.4), we need to use different matrices in the x and y-directions. In the x-direction we use the matrices described above; in the y-direction to discretize the second derivatives, for example, we use

$$A_2 = \begin{bmatrix} 2 & -1 & & & -1 \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ -1 & & & -1 & 2 \end{bmatrix}.$$

In this direction the fast Fourier transform is used instead of the fast cosine/sine transform.

3. Newton's method. We use Newton's method in two different ways. The first is to improve the numerical approximations obtained by the variational numerical methods. Since these are sometimes slow to converge, it is often faster to stop such an algorithm early and use its result as an initial guess for Newton's method. The second use for Newton's method is as part of a numerical continuation algorithm (see section 3.3).

3.1. Newton's method for given load parameter λ. This method can be used to improve solutions obtained by the MPA. Let $\lambda \in (0,2)$ be given. We are solving

$$(3.1) \qquad \mathscr{G}(w,\phi) = \begin{bmatrix} \mathscr{G}_1 \\ \mathscr{G}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{A}_{\Delta^2}w - \lambda\mathbf{A}_{xx}w + \mathbf{A}_{xx}\phi - 2[w,\phi]_1 \\ -\mathbf{A}_{\Delta^2}\phi + \mathbf{A}_{xx}w - [w,w]_2 \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}$$

for w and ϕ with zero average using Newton's method. The matrix we need to invert is

$$(3.2) \qquad \mathscr{G}'(w,\phi) = \begin{bmatrix} \frac{\partial\mathscr{G}_1}{\partial w} & \frac{\partial\mathscr{G}_1}{\partial\phi} \\ \frac{\partial\mathscr{G}_2}{\partial w} & \frac{\partial\mathscr{G}_2}{\partial\phi} \end{bmatrix} = \begin{bmatrix} \mathbf{A}_{\Delta^2} - \lambda\mathbf{A}_{xx} - 2\mathbf{B}_1 & \mathbf{A}_{xx} - 2\mathbf{B}_2 \\ \mathbf{A}_{xx} - 2\mathbf{B}_2^T & -\mathbf{A}_{\Delta^2} \end{bmatrix},$$

where

$$\mathbf{B}_1 = \frac{\partial}{\partial w}[w,\phi]_1 = \frac{1}{2}\mathbf{A}_{xx}(\operatorname{diag}\phi)\mathbf{A}_{yy} + \frac{1}{2}\mathbf{A}_{yy}(\operatorname{diag}\phi)\mathbf{A}_{xx} - \mathbf{A}_{xy}^T(\operatorname{diag}\phi)\mathbf{A}_{xy},$$

$$\mathbf{B}_2 = \frac{\partial}{\partial\phi}[w,\phi]_1 = \frac{1}{2}\left(\frac{\partial}{\partial w}[w,w]_2\right)^T$$

$$= \frac{1}{2}\mathbf{A}_{xx}(\operatorname{diag}\mathbf{A}_{yy}w) + \frac{1}{2}\mathbf{A}_{yy}(\operatorname{diag}\mathbf{A}_{xx}w) - \mathbf{A}_{xy}^T(\operatorname{diag}\mathbf{A}_{xy}w).$$

From the properties of $\boldsymbol{A}_{xx}, \boldsymbol{A}_{yy}, \boldsymbol{A}_{xy}, \boldsymbol{A}_{\Delta^2}$ it follows that matrix $\mathscr{G}'(w, \phi)$ is symmetric and singular, and its null space is spanned by $\left[\begin{smallmatrix}1\\0\end{smallmatrix}\right], \left[\begin{smallmatrix}0\\1\end{smallmatrix}\right]$.

To describe how to find the inverse of $\mathscr{G}'(w, \phi)$ on a subspace orthogonal to its null space, we introduce a new notation for the four blocks of $\mathscr{G}'(w, \phi)$ from (3.2):

$$\mathscr{G}'(w, \phi) = \left[\begin{array}{cc} \boldsymbol{G}_{11} & \boldsymbol{G}_{12} \\ \boldsymbol{G}_{12}^T & \boldsymbol{G}_{22} \end{array}\right].$$

For given vectors u, η with zero average we need to find vectors v, ζ with zero average such that

$$(3.3) \qquad \left[\begin{array}{cc} \boldsymbol{G}_{11} & \boldsymbol{G}_{12} \\ \boldsymbol{G}_{12}^T & \boldsymbol{G}_{22} \end{array}\right] \left[\begin{array}{c} v \\ \zeta \end{array}\right] = \left[\begin{array}{c} u \\ \eta \end{array}\right].$$

Let the tilde denote the block of the first $MN - 1$ rows and columns of a matrix and $MN - 1$ components of a vector. The matrix $\left[\begin{array}{cc} \tilde{\boldsymbol{G}}_{11} & \tilde{\boldsymbol{G}}_{12} \\ \tilde{\boldsymbol{G}}_{12}^T & \tilde{\boldsymbol{G}}_{22} \end{array}\right]$ is symmetric, nonsingular, and sparse. It can be inverted by a linear sparse solver. System (3.3) is then solved in the following steps:

$$\left[\begin{array}{c} r \\ \rho \end{array}\right] := \left[\begin{array}{cc} \tilde{\boldsymbol{G}}_{11} & \tilde{\boldsymbol{G}}_{12} \\ \tilde{\boldsymbol{G}}_{12}^T & \tilde{\boldsymbol{G}}_{22} \end{array}\right]^{-1} \left[\begin{array}{c} \tilde{u} \\ \tilde{\eta} \end{array}\right], \qquad \begin{array}{l} s := -\frac{1}{MN}\tilde{\mathbf{1}}^T r, \\[2mm] \sigma := -\frac{1}{MN}\tilde{\mathbf{1}}^T \rho, \end{array} \qquad \begin{array}{l} v = \left[\begin{array}{c} r + s\tilde{\mathbf{1}} \\ s \end{array}\right], \\[2mm] \eta = \left[\begin{array}{c} \rho + \sigma\tilde{\mathbf{1}} \\ \sigma \end{array}\right]. \end{array}$$

3.2. Newton's method for given S. This method can be used to improve solutions obtained by the CSDM and the CMPA. Let $C > 0$ be given. We are looking for numerical solutions of (1.1)–(1.2) in the set $\mathscr{M}$ defined by (1.8). Hence we are solving

$$(3.4) \qquad \left[\begin{array}{c} \boldsymbol{A}_{\Delta^2}w - \lambda\boldsymbol{A}_{xx}w + \boldsymbol{A}_{xx}\phi - 2[w, \phi]_1 \\ -\boldsymbol{A}_{\Delta^2}\phi + \boldsymbol{A}_{xx}w - [w, w]_2 \\ -\frac{1}{2}w^T\boldsymbol{A}_{xx}w + C/(4\Delta x\Delta y) \end{array}\right] = \left[\begin{array}{c} \mathbf{0} \\ \mathbf{0} \\ 0 \end{array}\right]$$

for w and ϕ with zero average and λ using Newton's method. The approach is very similar to that described in the previous section; the resulting matrix is symmetric and has just one more row and column than the matrix in (3.2).

3.3. Continuation. To follow branches of solutions (λ, w) of (3.1) we adopt a continuation method described in [8]. We introduce a parameter $s \in \mathbb{R}$ by adding a constraint—pseudoarclength normalization (in the $(\lambda, \|w\|_X)$-plane). For a given value of s we are solving

(3.5)

$$\bar{\mathscr{G}}(w, \phi, \lambda) = \left[\begin{array}{c} \mathscr{G}_1 \\ \mathscr{G}_2 \\ \mathscr{G}_3 \end{array}\right] = \left[\begin{array}{c} \boldsymbol{A}_{\Delta^2}w - \lambda\boldsymbol{A}_{xx}w + \boldsymbol{A}_{xx}\phi - 2[w, \phi]_1 \\ -\boldsymbol{A}_{\Delta^2}\phi + \boldsymbol{A}_{xx}w - [w, w]_2 \\ \theta\langle\dot{w}_0, w - w_0\rangle_X + (1 - \theta)\dot{\lambda}_0(\lambda - \lambda_0) - (s - s_0) \end{array}\right] = \left[\begin{array}{c} \mathbf{0} \\ \mathbf{0} \\ 0 \end{array}\right]$$

for w, ϕ with zero average and the load λ, where the value of $\theta \in (0, 1)$ is fixed (e.g., $\theta = \frac{1}{2}$). We assume that we are given a value s_0, an initial point (λ_0, w_0) on

the branch, and an approximate direction $(\dot{\lambda}_0, \dot{w}_0)$ of the branch at this point (an approximation of the derivative $\frac{d}{ds}(\lambda(s), w(s))|_{s=s_0}$).

System (3.5) is solved for a discrete set of values of s in some interval (s_0, s_1) by Newton's method. Then a new initial point on the branch is defined by setting $w_0 = w(s_1)$, $\lambda_0 = \lambda(s_1)$, $s_0 = s_1$, a new direction $(\dot{\lambda}_0, \dot{w}_0)$ at this point is computed, and the process is repeated. The matrix we need to invert in Newton's method is

$$(3.6) \quad \bar{\mathscr{G}}'(w, \phi, \lambda) = \begin{bmatrix} \mathscr{G}'(w, \phi, \lambda) & g \\ h^T & d \end{bmatrix},$$

$$g = \begin{bmatrix} -\boldsymbol{A}_{xx} w \\ \boldsymbol{0} \end{bmatrix}, \quad h = \begin{bmatrix} 4\theta \boldsymbol{A}_{\langle,\rangle} \dot{w}_0 \, \Delta x \Delta y \\ \boldsymbol{0} \end{bmatrix}, \quad d = (1 - \theta)\dot{\lambda}_0,$$

where $\boldsymbol{A}_{\langle,\rangle} = \boldsymbol{A}_{\Delta^2} + \boldsymbol{A}_{xx}\boldsymbol{A}_{\Delta^2}^{-1}\boldsymbol{A}_{xx}$.

Solving a linear system with this matrix amounts to solving system (3.3) for two right-hand sides. For a given $u \in \mathbb{R}^{2MN}$ with $[\mathbf{1}^T \ \mathbf{1}^T]u = 0$ and a given $\eta \in \mathbb{R}$ we want to solve

$$(3.7) \qquad \begin{bmatrix} \mathscr{G}'(w, \phi, \lambda) & g \\ h^T & d \end{bmatrix} \begin{bmatrix} v \\ \zeta \end{bmatrix} = \begin{bmatrix} u \\ \eta \end{bmatrix}$$

for v with $[\mathbf{1}^T \ \mathbf{1}^T]v = 0$ and ζ. System (3.7) is solved in the following steps:

$$\text{solve}: \quad \begin{aligned} \mathscr{G}'(w, \phi, \lambda)v_1 &= g, \\ \mathscr{G}'(w, \phi, \lambda)v_2 &= u, \end{aligned} \qquad \zeta = \frac{\eta - h^T v_2}{d - h^T v_1}, \qquad v = v_2 - \zeta v_1.$$

Remark 3.1. Note that in this implementation we simply follow a solution of the equation; there is no guarantee that this remains a local minimum, a mountain-pass solution or constrained mountain-pass solution (cf. Figure 12).

Remark 3.2. Newton's method and continuation have been implemented using only a one-sided discretization of the mixed derivative ∂_{xy} and only on the domain $\Omega_{\frac{1}{4}}$ assuming symmetry (1.9). The alternative discretization of ∂_{xy} described in section 2.5 uses the fast cosine/sine transform. The resulting matrix $\tilde{\boldsymbol{A}}_{xy}$ is not sparse, and therefore we would obtain a dense block $\boldsymbol{G}_{12}$ in system (3.3) which would prevent us from using a sparse solver.

On the full domain Ω we assume periodicity of w and ϕ in the y-direction. Hence for a discretization with a small step Δy the matrix we invert when solving (3.3) would become close to singular. The shift in the y-direction is prevented by assuming the symmetry $w(x, y) = w(x, -y)$, $\phi(x, y) = \phi(x, -y)$.

4. Numerical solutions. We fix the size of the domain and the size of the space step for the following numerical computations: $a = b = 100$, $\Delta x = \Delta y = 0.5$. We obtain solutions using the variational techniques SDM, MPA, CSDM, and CMPA. Table 1 provides a summary of which discretization was used in which algorithm.

4.1. A mountain-pass solution on the full domain Ω. The first numerical experiments were done on the full domain Ω, without symmetry restrictions, and with the unbiased (Fourier) discretization of ∂_{xy} (section 2.5). For a fixed load $\lambda = 1.4$ we computed a mountain-pass solution using the MPA (section 1.2.2). For endpoints we used $w_1 = 0$ and a second point w_2 obtained by the SDM (here the initial point for the SDM was chosen to have a single peak centered at $x = y = 0$). The graph of the

TABLE 1
Summary of which spatial discretization was used in the different numerical methods.

	Mixed derivative ∂_{xy}		Computational domain	
	One-sided	Fourier	Full	1/4
Variational methods	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$
Newton/continuation	$\checkmark$			$\checkmark$
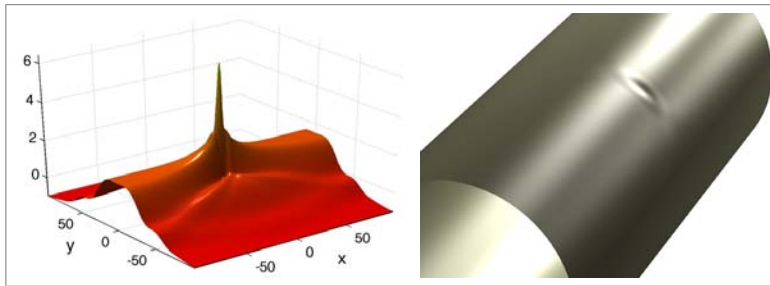
FIG. 3. *Mountain-pass solution for $\lambda = 1.4$ found using the MPA on the full domain Ω with ∂_{xy} discretized using the fast Fourier transform.*

solution w_{MP} is shown in Figure 3 (left). The figure on the right shows w_{MP} rendered on a cylinder (see the appendix for details on the scaling of the geometry), and we see it has the form of a single dimple. The value of shortening for this solution is $S(w_{\mathrm{MP}}) = 14.93529$.

Alternatively, if we apply the CSDM with $S = 14.93529$ and use a function with a single peak in the center of the domain as the initial condition w_0, we also obtain the same solution w_{MP}, this time as a local minimizer of E under constrained S.

We remark that although the MPA and the CSDM have a local character, we have not found any numerical mountain-pass solution with the total potential F_λ smaller than $F_\lambda(w_{\mathrm{MP}})$ for $\lambda = 1.4$. Similarly, using the CSDM we have not found any solution with energy E smaller than $E(w_{\mathrm{MP}})$ under the constraint $S = 14.93529$. We briefly discuss the physical relevance of this solution in section 6.

4.2. Solutions under symmetry restrictions. The solution w_{MP} of Figure 3 satisfies the symmetry property (1.9). In the computations described below we enforced this symmetry and worked on the quarter domain $\Omega_{\frac{1}{4}}$, thus reducing the complexity of the problem. In order to improve the variational methods by combining them with Newton's method, we also discretized the mixed derivative ∂_{xy} using left-sided finite differences. The influence of this choice on the numerical solution is described in section 5.1.

4.2.1. CSDM. We first fixed $S = 40$ and used the CSDM to obtain constrained local minimizers of E as shown in Table 2. They are ordered according to the increasing value of stored energy E. Their graphs and renderings on a cylinder are shown in Figure 4. Solution 1.1 is similar to the single dimple solution w_{MP} described above and according to Table 2 it has, indeed, the smallest value of E.

4.2.2. MPA. We then used the MPA for fixed $\lambda = 1.4$ and various choices of w_2 to obtain the local mountain-pass points of F_λ as shown in Table 3. They are ordered according to the increasing value of the total potential F_λ. The shape of their graph is very similar to that of the CSDM solutions discussed above and depicted in

TABLE 2
Numerical solutions obtained by the CSDM on $\Omega_{\frac{1}{4}}$ with ∂_{xy} discretized using left-sided finite differences. Graphs are shown in Figure 4.

CSDM	λ	S	E	F_λ	Same shape as **MPA**
1.1	1.108121	40	56.85636	12.53151	2.1
1.2	1.299143	40	62.76150	10.79577	2.2
1.3	1.316146	40	63.21646	10.57063	2.3
1.4	1.311687	40	63.64083	11.17334	2.4
1.5	1.309586	40	63.70623	11.32278	2.5
1.6	1.328997	40	64.00875	10.84889	2.6
1.7	1.344898	40	64.52244	10.72651	2.7

TABLE 3
Numerical solutions obtained by the MPA on $\Omega_{\frac{1}{4}}$ with ∂_{xy} discretized using left-sided finite differences.

MPA	λ	S	E	F_λ	Same Shape as **CSDM**
2.1	1.4	17.73822	29.42997	4.596460	1.1
2.2	1.4	29.85121	49.08882	7.297132	1.2
2.3	1.4	31.28849	51.39952	7.595635	1.3
2.4	1.4	31.41723	52.01893	8.034809	1.4
2.5	1.4	31.22992	51.84074	8.118852	1.5
2.6	1.4	32.77491	54.15818	8.273314	1.6
2.7	1.4	34.19888	56.56472	8.686284	1.7

TABLE 4
Numerical solutions obtained by the CMPA/Newton on $\Omega_{\frac{1}{4}}$ with ∂_{xy} discretized using left-sided finite differences. Graphs are shown in Figures 5 and 6.

CMPA	λ	S	E	F_λ	Endpoints
3.1	1.310815	40	63.98996	11.55737	1.2, 1.4
3.2	1.332112	40	64.38609	11.10161	1.3, 1.6
3.3	1.447626	40	66.49032	8.585294	1.2, 1.3 or 1.3, 1.4
3.4	1.440841	40	66.72079	9.087129	1.1, 1.4
3.5	1.447594	40	66.97057	9.066810	1.1, 1.6
3.6	1.484790	40	68.20637	8.814758	1.3, 1.5
3.7	1.477769	40	68.23274	9.121955	1.1, 1.5*
3.8	1.482261	40	68.41086	9.120428	1.1, 1.7
3.9	1.413917	40	68.56697	12.01028	1.1, 1.2
3.10	1.520975	40	68.83818	7.999162	1.2, 1.5
3.11	1.475705	40	69.00087	9.972652	1.1, 1.5*
3.12	1.532000	40	69.27379	7.993781	1.4, 1.5†
3.13	1.527108	40	69.35834	8.274019	1.2, 1.6
3.14	1.551762	40	69.47838	7.407904	1.4, 1.5†
3.15	1.547955	40	69.68292	7.764712	1.6, 1.7
3.16	1.539785	40	69.78487	8.193480	1.5, 1.7
3.17	1.546480	40	69.85900	7.999795	1.3, 1.7
3.18	1.549780	40	70.16253	8.171339	1.2, 1.7
3.19	1.561117	40	70.74117	8.296474	1.4, 1.7

Figure 4, so we do not show their graphs here. Solution 2.1 is again the single dimple solution, and the table shows that it has the smallest value of F_λ.

4.2.3. CMPA. We then fixed $S = 40$ and applied the CMPA to obtain constrained local mountain passes of E as shown in Table 4. They are again ordered according to the increasing value of stored energy E. Their graphs and renderings on a cylinder are shown in Figures 5 and 6. For endpoints w_1, w_2 of the path in the CMPA we used the constrained local minimizers 1.1–1.7.
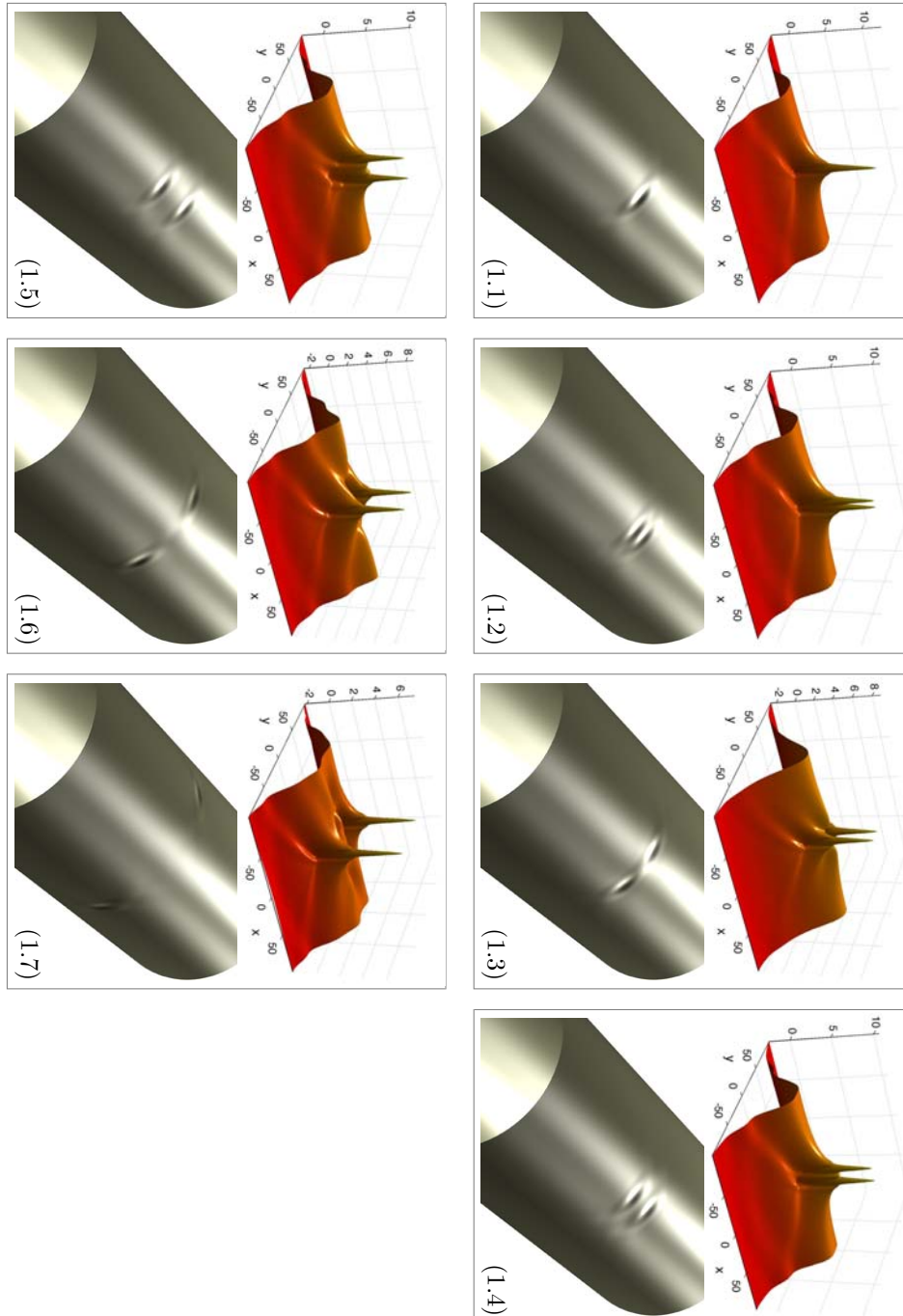
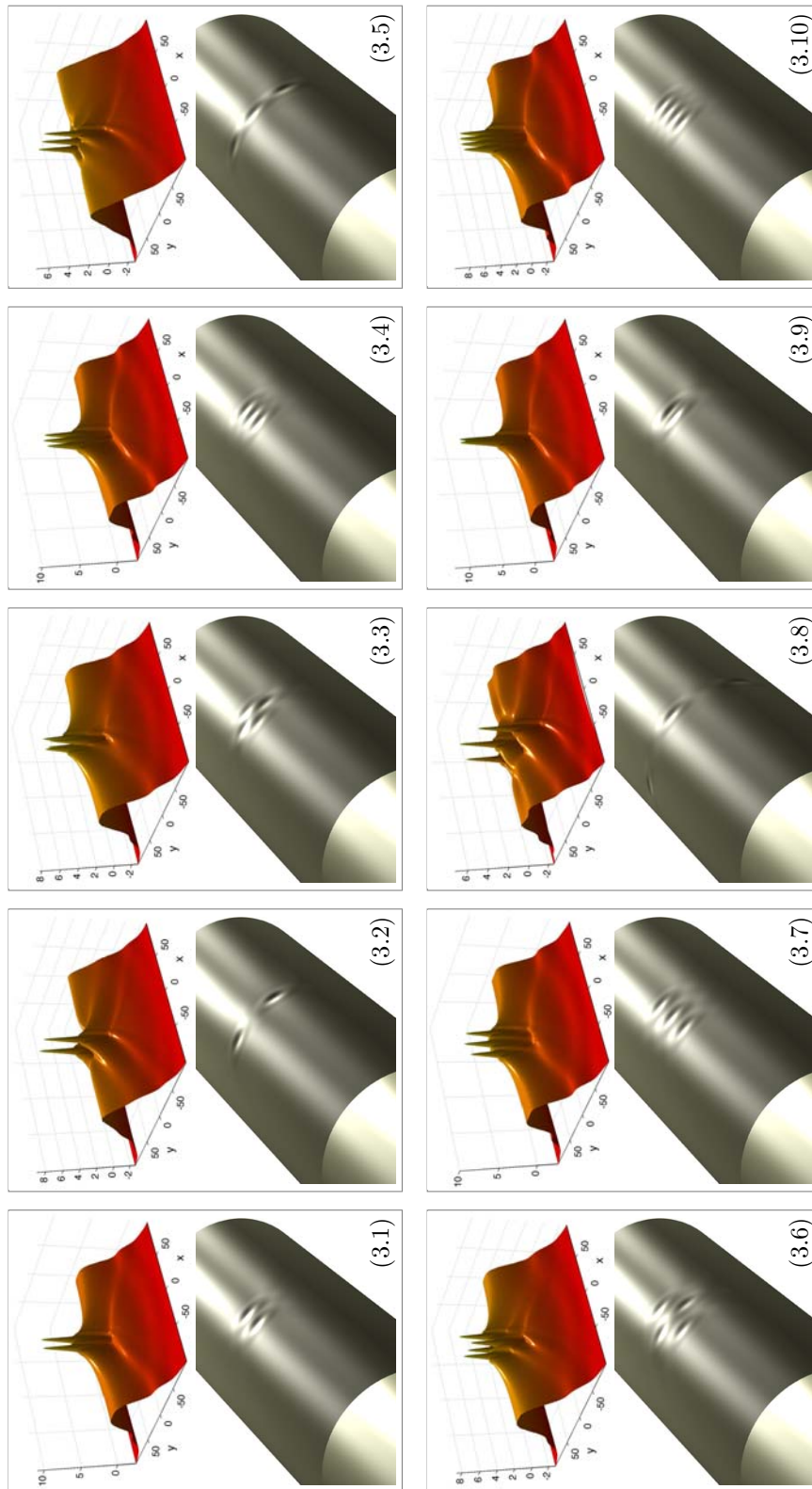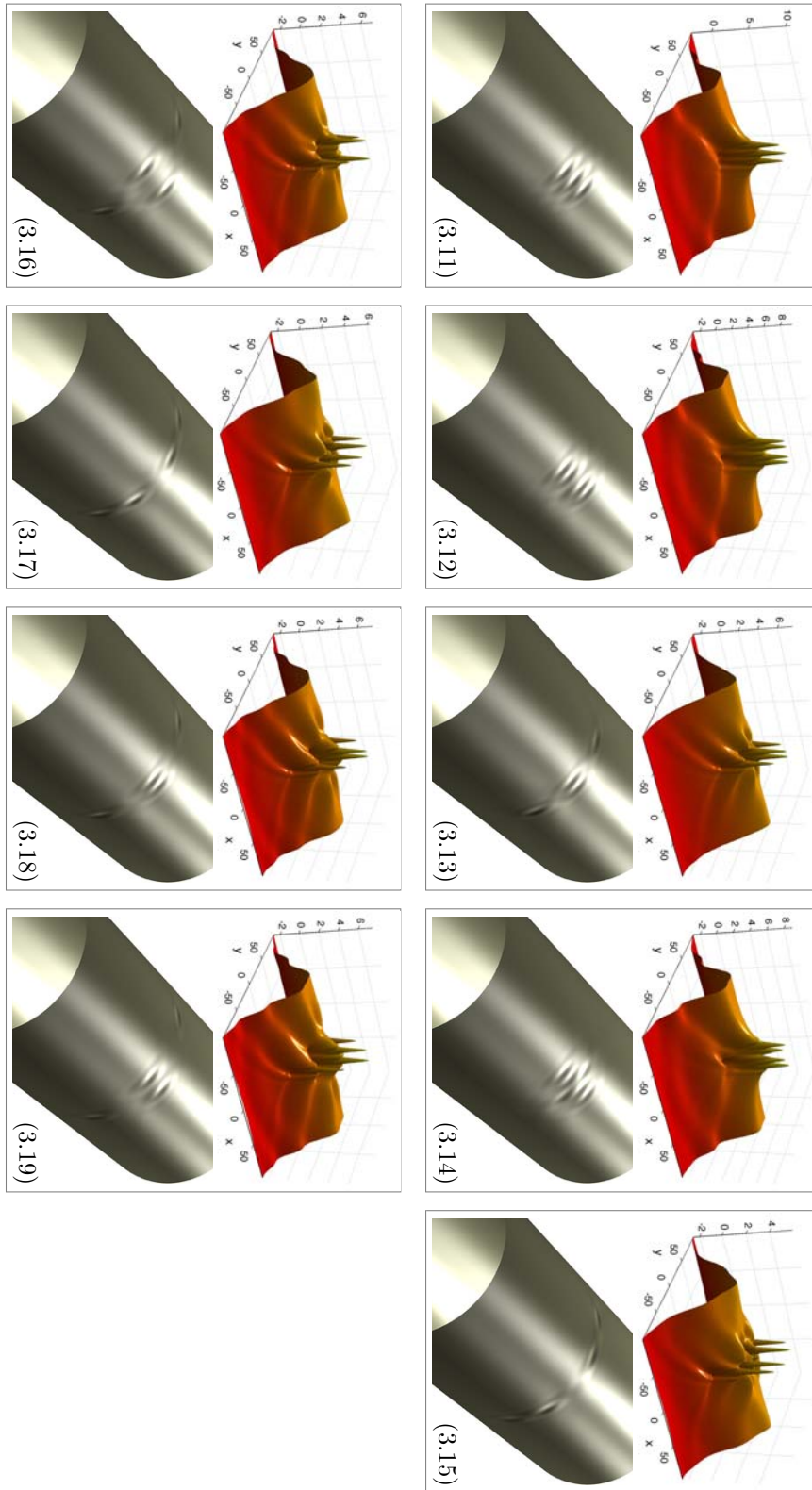FIG. 4. *Numerical solutions found using the CSDM with axial end shortening $S = 40$. More details are in Table 2.*

FIG. 6. Numerical solutions found using the CMPA/Newton with S = 40. More details are in Table 4.

There are 21 possible pairs (w_1, w_2) to be used but only 19 solutions in Table 4. The algorithm did not converge for the following three pairs: (1.1, 1.3), (1.5, 1.6), and (1.4, 1.6). This is most likely due to the complicated nature of the energy landscape between these endpoints. On the other hand, two choices of pairs, denoted by $*$ and $\dagger$ in the table, yielded two solutions each. When the path is deformed it sometimes comes close to another critical point of E which is not a constrained mountain pass. In that case the algorithm slows down and one can apply Newton's method to such a point. It is a matter of luck whether Newton's method converges. The CMPA then runs further and might converge to another point, this time a constrained mountain-pass point. And finally, two choices of (w_1, w_2) yielded the same solution 3.3.

5. Remarks on the numerics.

5.1. Bias in the discretization of ∂_{xy}. In this section we examine the influence of the discretization of the mixed derivative ∂_{xy} on the numerical solution. We recall that the mixed derivative ∂_{xy} can be discretized using left-/right-sided finite differences (2.3), (2.4) or using the fast Fourier transform (section 2.5). For comparison we use the single-dimple solution on $\Omega = (-100, 100)^2$ at load $\lambda = 1.4$ obtained by the MPA.

Let $\Delta x = \Delta y = 0.5$. Table 5 gives a list of numerical experiments together with the values of shortening and energy. Figure 7 shows a profile of the numerical solutions in the circumferential direction at $x = 0$.

TABLE 5

Single-dimple numerical solution obtained by the MPA with and without the symmetry assumption (1.9) and with various kinds of discretization of ∂_{xy}.

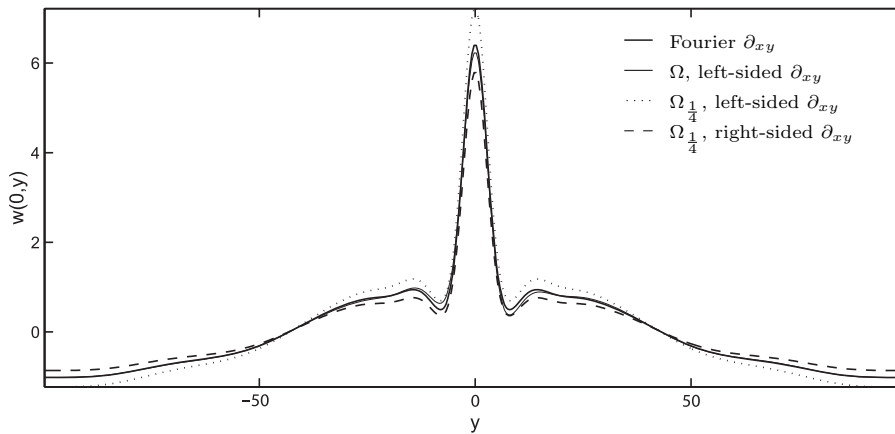| Domain	Discretization of ∂_{xy}	λ	S	E	F_λ	Figure 7
Ω or $\Omega_{\frac{1}{4}}$	Fourier	1.4	14.93529	24.71825	3.808850	————
Ω	left-/right-sided	1.4	14.93617	24.70828	3.797636	————
$\Omega_{\frac{1}{4}}$	left-sided	1.4	17.73822	29.42997	4.596460	··········
$\Omega_{\frac{1}{4}}$	right-sided	1.4	12.81205	21.16342	3.226549	- - - - -

FIG. 7. *Profile of the single-dimple numerical solution w_{MP} at $x = 0$ obtained by the MPA with and without the symmetry assumption (1.9) and with various kinds of discretization of ∂_{xy}.*

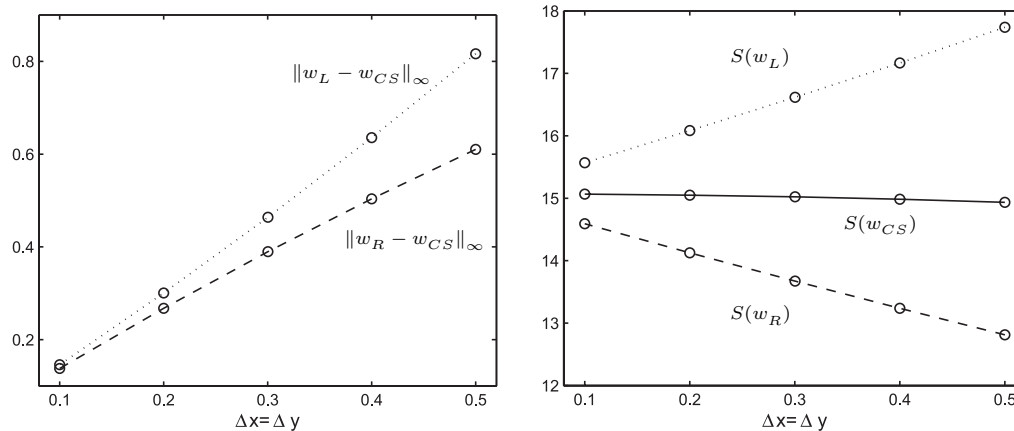Fig. 8. *Influence of the size of the space step Δx, Δy on the numerical solution w_{MP} obtained by the MPA for three different kinds of discretization of ∂_{xy}. Let w_L, w_R denote the numerical solutions obtained using the left- and right-sided discretization of ∂_{xy}, respectively, with w_{CS} using the fast cosine/sine transform. Left: comparison of the solutions in the maximum norm; right: the value of shortening S.*

On the full domain Ω with no assumption on symmetry of solutions, the discretization of ∂_{xy} using the left-/right-sided finite differences (2.3) provides a numerical solution that is slightly asymmetric (Figure 7, thin solid line). The Fourier transform provides a symmetric solution (Figure 7, thick solid line). The same numerical solution can be obtained on $\Omega_{\frac{1}{4}}$ under the symmetry assumption (1.9) with ∂_{xy} discretized using the fast cosine/sine transform.

On $\Omega_{\frac{1}{4}}$ the symmetry of numerical solutions is guaranteed by assumption (1.9). The use of left-/right-sided discretization of ∂_{xy} does, however, have an influence on the shape of the numerical solution, as Figure 7 shows (the dotted and the dashed line).

5.2. Convergence. We now turn to the influence of the size of the space step Δx, Δy on the numerical solution. We run the MPA on $\Omega_{\frac{1}{4}}$ under the symmetry assumption (1.9) with ∂_{xy} discretized using (a) the fast cosine/sine transform, (b) left-sided finite differences, (c) right-sided finite differences. We consider $\Delta x = \Delta y = 0.5, 0.4, 0.3, 0.2, 0.1$, i.e., we take 200, 250, 333, 500, and 1000 points in both axis directions, respectively. Figure 8 illustrates convergence as $\Delta x, \Delta y \to 0$ of the numerical solutions obtained by various types of discretization of ∂_{xy}.

5.3. Dependence on the size of the domain. As observed in [7], the localized nature of the solutions suggests that they should be independent of domain size, in the sense that for a sequence of domains of increasing size the solutions converge (for instance, uniformly on compact subsets). Such a convergence would also imply convergence of the associated energy levels. Similarly, we would expect that the aspect ratio of the domain is of little importance in the limit of large domains.

We tested these hypotheses by computing the single-dimple solution on domains of different sizes and aspect ratios. In all the computations the space step $\Delta x = \Delta y = 0.5$ is fixed. In order to use the continuation method of section 3.3, we discretized ∂_{xy} using the left-sided finite differences. We also assumed symmetry of solutions given by (1.9) and worked on $\Omega_{\frac{1}{4}}$. We recall the notation of computational domains, $\Omega = (-a, a) \times (-b, b)$, $\Omega_{\frac{1}{4}} = (-a, 0) \times (-b, 0)$.
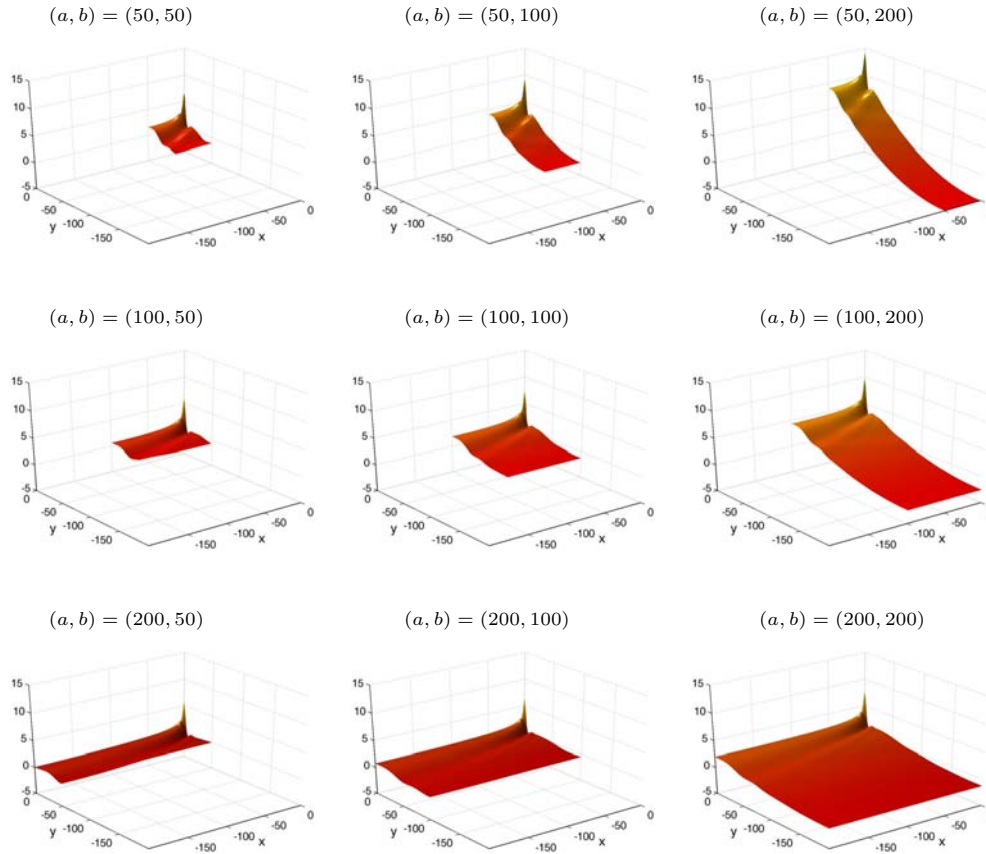
FIG. 9. *The single-dimple mountain-pass solution with* $\lambda = 1.4$ *computed under the assumption of symmetry* (1.9) *with left-sided discretization of* ∂_{xy} *for various domain sizes.*

Figure 9 shows the results for load $\lambda = 1.4$. First we notice that the central dimple has almost the same shape in all the cases shown. But there seems to be a difference in the slope of the "flat" part leading to this dimple. On domains with small a (short cylinder) the derivative in the circumferential y-direction in this part is larger than on domains with larger a (longer cylinder). The circumferential length b seems to be less important for the shape of the solution: for example, the cases $(200, 50)$ and $(200, 100)$ look like restrictions of the case $(200, 200)$ to smaller domains.

We take a closer look at domains of sizes $(a, b) = (100, 100)$, $(100, 200)$, $(200, 100)$, and $(200, 200)$ and the corresponding solutions $w_{100,100}$, $w_{100,200}$, $w_{200,100}$, and $w_{200,200}$ shown in the figure. We compare the first three with the last one, respectively. It does not make sense to compare the values of w itself since the energy functional F_λ depends on derivatives of w only. We choose to compare w_{xx} and w_{yy}. Table 6 gives the infinity norm of the relative differences. Figure 10 shows graphs of the difference $w_{100,100} - w_{200,200}$ and of the second derivatives $(w_{200,200} - w_{100,100})_{xx}$, $(w_{200,200} - w_{100,100})_{yy}$ on the subdomain $(-100, 0)^2$.

We conclude that solutions on different domains compare well; the maximal difference in the second derivatives of w is three orders of magnitude smaller than the supremum norm of the same derivative. We also observe that varying the length

TABLE 6
Comparison of the second derivatives of solutions from Figure 9 computed on domains with $(a,b) = (100, 100)$, $(100, 200)$, $(200, 100)$, and $(200, 200)$.

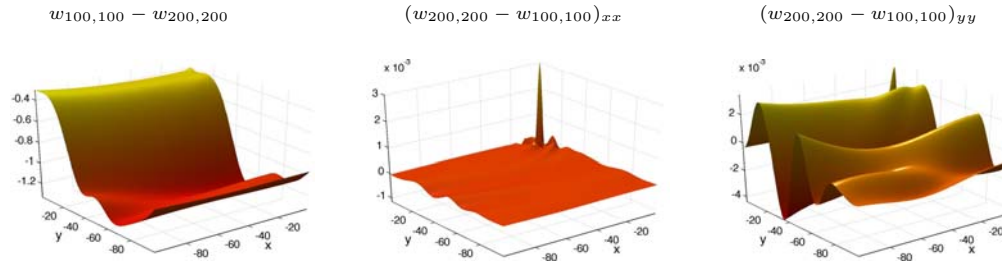|	$\dfrac{\|(w-w_{200,200})_{xx}\|_\infty}{\|(w_{200,200})_{xx}\|_\infty}$	$\dfrac{\|(w-w_{200,200})_{yy}\|_\infty}{\|(w_{200,200})_{yy}\|_\infty}$
$w = w_{100,100}$	$2.835 \cdot 10^{-3}$	$5.313 \cdot 10^{-3}$
$w = w_{100,200}$	$1.943 \cdot 10^{-3}$	$4.917 \cdot 10^{-3}$
$w = w_{200,100}$	$1.827 \cdot 10^{-4}$	$9.638 \cdot 10^{-4}$

FIG. 10. *Comparison of solutions $w_{100,100}$, $w_{200,200}$ from Figure 9 obtained on square domains with $a = b = 100$ and $a = b = 200$, respectively, and their second derivatives. For a reference, we note that $\|(w_{200,200})_{xx}\|_\infty = 1.064522$ and $\|(w_{200,200})_{yy}\|_\infty = 0.8242491$.*
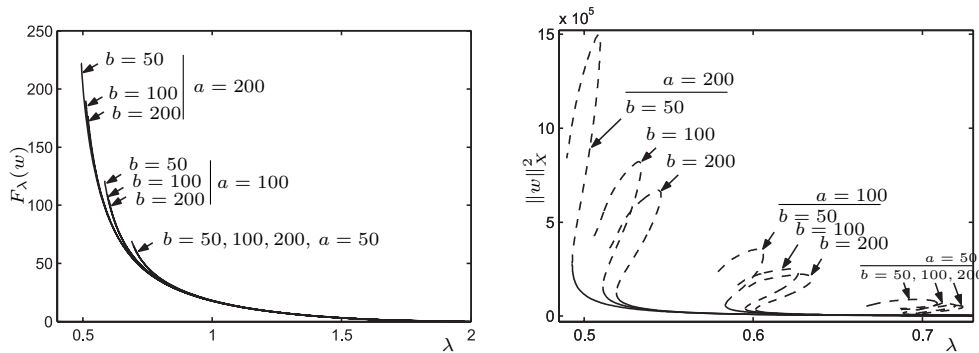

FIG. 11. *Continuation of the single-dimple solution found as a numerical mountain pass for $\lambda = 1.4$ on domains of various sizes for a range of values λ. Left: $F_\lambda(w)$ as a function of λ; right: $\|w\|_X$ as a function of λ.*

parameter a while keeping the circumference parameter b fixed causes larger changes in the numerical solution than varying the cylinder circumference while keeping the length fixed.

Another way of studying the influence of the domain size on the numerical solution is comparing solution branches obtained by continuation as described in section 3.3. We start with the mountain-pass solution for $\lambda = 1.4$ shown in Figure 9 and continue it for both $\lambda > 1.4$ and $\lambda < 1.4$. The results are presented in Figure 11. We observe that the branches corresponding to the considered domains do not differ much for the range of λ between approximately 0.71 and 2. Below $\lambda \approx 0.71$ the size of the domain, particularly the length of the cylinder described by a, has a strong influence. The graph on the right shows that the larger (longer) the domain Ω, the smaller the value
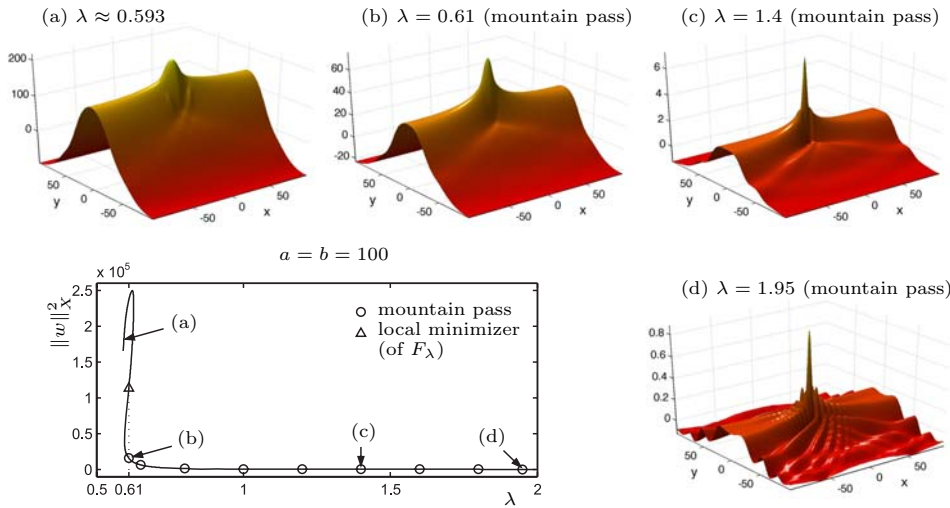
FIG. 12. *A detailed look at the continuation of the single-dimple solution on the domain with* $a = b = 100$.

of λ at which the norm $\|w\|_X$ starts to rapidly increase for decreasing λ. The graph on the left shows the energy $F_\lambda(w)$ along a solution branch. The data shown here correspond to the ones in the graph on the right marked by a solid line. The dashed line in the right graph shows also some data after the first limit point is passed.

Figure 12 shows how the graph of $w(x, y)$ changes as a solution branch is followed. Here we chose a square domain with $a = b = 100$ and plotted the solution for four values of λ (note that Figures 12 (c), 9 (100,100), and 7 (dotted line) show the same numerical solution). We observe that with decreasing λ the height of the central dimple increases, the dimple becomes wider, and the ripples (present at λ close to 2) disappear. In Figure 12 (a) we observe that new dimples are being formed next to the central dimple.

It should be noted that although we started the continuation at $\lambda = 1.4$ at a mountain-pass point, not all the points along a continuation branch are mountain passes. Since it is not feasible to use the MPA to verify this for each point, we chose just a few. Still on the example of the domain with $a = b = 100$ in Figure 12, the circles on the continuation branch mark those points that have also been found by the MPA (for $\lambda = 0.61, 0.65, 0.8, 1.0, 1.2, 1.4, 1.6, 1.8, 1.95$). As described in section 1.2.2, in order to start the MPA a point w_2 is needed such that $F_\lambda(w_2) < 0$. The analysis in [7] shows that for a given λ such a point exists provided the domain Ω is large enough, and in practice it is found by the SDM of section 1.2.1. This was indeed the case for all the chosen values of λ except for $\lambda = 0.61$. In this case, starting from some w_0 with a large norm, the SDM provides a trajectory $w(t)$ such that $F_\lambda(w(t)) > 0$ for all $t > 0$. In fact, the steepest descent method converges to a local minimizer w_M with $F_\lambda(w_M) \approx 76.1$. This is hence no mountain pass but, nevertheless, lies on the same continuation branch and is marked by a triangle in the figure. Despite $F_\lambda(w_M) > 0$ we can still try to run the MPA with $w_2 = w_M$. It converges and yields w_{MP} with $F_\lambda(w_{MP}) \approx 94.8$ (marked by a circle at $\lambda = 0.61$ and shown in graph (b)).

The comparison of solutions computed on different domains and their respective energies suggests that for each λ we are indeed dealing with a single, localized function

defined on $\mathbb{R}^2$, of which our computed solutions are finite-domain adaptations. Based on this suggestion and the above discussion of the mountain-pass solutions we could, for example, conclude that the mountain-pass energy

$$V(\lambda, \Omega) := \inf_{w_2}\big\{ F_\lambda\big(w_{\mathrm{MP}}(\lambda, \Omega, w_2)\big) : F_\lambda(w_2) < 0 \big\}$$

is a finite-domain approximation of a function $V(\lambda)$, whose graph almost coincides with that of $V(\lambda, \Omega)$ for λ not too small (cf. Figure 11 (left)).

6. Discussion.

6.1. Variational numerical methods. We have seen that given a complex energy surface many solutions may be found using these variational techniques. For example, for a fixed end shortening of $S = 40$, Figures 4, 5, and 6 are all solutions. Which of these solutions is of greatest relevance depends on the question that is being asked.

In the context of the cylinder (and similar structural applications) the mountain-pass solution from the unbuckled state ($w_1 = 0$) with minimal energy is of physical interest. Often the experimental buckling load may be at 20–30% of the linear prediction from a bifurcation analysis (in our scaling this corresponds to $\lambda = 2$). This uncertainty in the buckling load is a drawback for design and so "knockdown" factors have been introduced, based on experimental data. It was argued in [7] that the energy of the mountain-pass solution w_{MP} in fact provides a lower bound on the energy required to buckle the cylinder, and so these solutions provide bounds on the (observed) buckling load of the cylinder.

This example illustrates an important aspect of the (constrained) mountain-pass algorithm: its explicit nonlocality. The algorithm produces a saddle point which has an additional property: it is the separating point (and level) between the basins of attraction of the endpoints w_1 and w_2.

Another technique to investigate a complex energy surface is to perform a simulated annealing computation, essentially to solve the SDM (or the CSDM) problem with additive stochastic forcing. The aim in these techniques is often to find a global minimizer (if it exists) where there are a large number of local minimizers. Here by either the MPA or the CMPA we find the solution between two such minima and so get an estimate on the surplus energy needed to change between local minima.

6.2. Numerical issues. The numerical issues that we encountered are of two types. First there are the requirements that are related to the specific problem of the Von Kármán–Donnell equations, such as the discretization of the mixed derivative and the bracket, and the fact that the solutions are symmetric and highly localized.

For other difficulties it is less clear. For smaller values of λ each of the variational methods converged remarkably slowly. Newton's method provides a way of improving the convergence, but the question of whether this slow convergence is typical for a whole class of variational problems is relevant. It would be interesting to connect the rate of convergence of, for instance, the SDM to certain easily measurable features of the energy landscape.

Appendix. This appendix shows the transformation between the Von Kármán–Donnell equations in the physical coordinates and their scaled form (1.1)–(1.2). For the full derivation of the Von Kármán–Donnell equations we refer to [7].

In the physical coordinates the Von Kármán–Donnell equations, the stored energy, and the shortening assume the following form:

(A.1) $\quad \dfrac{t^2}{12(1-\nu^2)}\tilde{\Delta}^2\tilde{w} + \dfrac{P}{2\pi REt}\tilde{w}_{\tilde{x}\tilde{x}} - \dfrac{1}{R}\tilde{\phi}_{\tilde{x}\tilde{x}} - 2\left[\tilde{w},\tilde{\phi}\,\right] = 0,$

(A.2) $\quad \tilde{\Delta}^2\tilde{\phi} + \dfrac{1}{R}\tilde{w}_{\tilde{x}\tilde{x}} + [\tilde{w},\tilde{w}\,] = 0,$

(A.3) $\quad \tilde{E}(\tilde{w}) = \dfrac{t^3 E}{24(1-\nu^2)}\displaystyle\int_{\tilde{\Omega}}(\tilde{\Delta}\tilde{w})^2 + \dfrac{tE}{2}\int_{\tilde{\Omega}}(\tilde{\Delta}\tilde{\phi})^2, \qquad \tilde{S}(\tilde{w}) = \dfrac{1}{4\pi R}\int_{\tilde{\Omega}}(\tilde{w}_{\tilde{x}})^2,$

where $(\tilde{x},\tilde{y}) \in \tilde{\Omega} := (0,L)\times(0,2\pi R)$, $\tilde{x}$ is the axial coordinate, $\tilde{y}$ the circumferential coordinate, L the length of the cylinder, R its radius, t the thickness of the shell, E Young's modulus, and ν Poisson's ratio. The cylinder is subject to a compressive axial force P. The function $\tilde{w}$ is an inward radial displacement measured from the unbuckled state, and $\tilde{\phi}$ is the Airy stress function. The bracket $[\cdot,\cdot]$ is defined as in (1.3), only the derivatives are taken with respect to the physical coordinates $\tilde{x},\tilde{y}$. We note that $\tilde{E}$ does not include the energy of the prebuckling deformation.

In order to eliminate most of the physical parameters, we rescale the equations using the following transformation:

$$\begin{aligned} \tilde{x} &= (2\pi R\sqrt{k})\,(x+a), & a &= L/(4\pi R\sqrt{k})\,, & x &\in (-a,a)\,, \\ \tilde{y} &= (2\pi R\sqrt{k})\,(y+b), & b &= 1/(2\sqrt{k})\,, & y &\in (-b,b)\,, \end{aligned}$$

where $k = t/(8\pi^2 R\sqrt{3(1-\nu^2)})$. We introduce new unknown functions,

$$\tilde{w}(\tilde{x},\tilde{y}) = (4\pi^2 Rk)\,w(x,y)\,, \qquad \tilde{\phi}(\tilde{x},\tilde{y}) = (4\pi^2 Rk)^2\,\phi(x,y),$$

and a load parameter $\lambda = P\sqrt{3(1-\nu^2)}/(\pi Et^2)$. We then obtain (1.1)–(1.2), which depend on three parameters only: λ, a, and b. The functionals in physical coordinates defined in (A.3) and their scaled counterparts defined in (1.6) are related through

$$\tilde{E}(\tilde{w}) = tER^2(4\pi^2 k)^3 E(w)\,, \qquad \tilde{S}(\tilde{w}) = 8\pi^3 Rk^2 S(w)\,.$$

Although the scaled energy and Young's modulus in the above equations are both denoted by E, Young's modulus does not appear in the scaled version of the problem. This avoids possible confusion.

The numerical results in this paper are presented mostly as graphs of function $w(x,y)$ over the domain $\Omega = (-a,a)\times(-b,b)$, and as a rendering of $\tilde{w}$ on a cylinder. For this rendering we used a sample cylinder of thickness $t = 0.247\,\text{mm}$ and Poisson's ratio $\nu = 0.3$. Hence, for example, a computational domain $\Omega = (-100,100)^2$ corresponds to a physical cylinder of length $47.58\,\text{cm}$ and radius $7.57\,\text{cm}$.

REFERENCES

[1] C. CANUTO, M. Y. HUSSAINI, A. QUATERONI, AND T. A ZANG, *Spectral Methods in Fluid Dynamics*, Springer, New York, 1987.

[2] Y. S. CHOI AND P. J. MCKENNA, *A mountain pass method for the numerical solution of semilinear elliptic problems*, Nonlinear Anal., 20 (1993), pp. 417–437.

[3] G. FRIESECKE AND J. B. MCLEOD, *Dynamic stability of non-minimizing phase mixtures*, Proc. Royal Soc. London A, 453 (1997), pp. 2427–2436.

[4] J. HORÁK, *Constrained mountain pass algorithm for the numerical solution of semilinear elliptic problems*, Numer. Math., 98 (2004), pp. 251–276.

[5] J. HORÁK AND P. J. MCKENNA, *Traveling waves in nonlinearly supported beams and plates*, in Nonlinear Equations: Methods, Models, and Applications (Bergamo, 2001), Progr. Nonlinear Differential Equations Appl. 54, Birkhäuser, Basel, 2003, pp. 197–215.

[6] J. HORÁK AND W. REICHEL, *Analytical and numerical results for the Fučík spectrum of the Laplacian*, J. Comput. Appl. Math., 161 (2003), pp. 313–338.

[7] J. HORÁK, G. J. LORD, AND M. A. PELETIER, *Cylinder buckling: The mountain pass as an organizing center*, SIAM J. Appl. Math., 66 (2006), pp. 1793–1824.

[8] H. B. KELLER, *Numerical solution of bifurcation and nonlinear eigenvalue problems*, in Applications of Bifurcation Theory (Proceedings of the Advanced Seminar, University of Wisconsin, Madison, WI, 1976), Publ. Math. Res. Center 38, Academic Press, New York, 1977, pp. 359–384.

[9] V. MAZ′YA, S. NAZAROV, AND B. PLAMENEVSKIJ, *Asymptotic Theory of Elliptic Boundary Value Problems in Singularly Perturbed Domains. Vol. I*, Oper. Theory Adv. Appl. 111, Birkhäuser Verlag, Basel, 2000. Translated from the German by Georg Heinig and Christian Posthoff.

[10] R. D. RICHTMEYER AND K. W. MORTON, *Difference Methods for Initial-Value Problems*, 2nd ed., Krieger, Malabar, FL, 1994.

[11] G. STRANG, *Introduction to Applied Mathematics*, Wellesley-Cambridge Press, Wellesley, MA, 1986.

[12] P. J. SWART AND P. J. HOLMES, *Energy minimization and the formation of microstructure in dynamic anti-plane shear*, Arch. Rational Mech. Anal., 121 (1992), pp. 37–85.