A contact formulation based on a volumetric potential: Application to isogeometric simulations of atrioventricular valves

David Kamensky[a,*], Fei Xu[b], Chung-Hao Lee[c], Jinhui Yan[d], Yuri Bazilevs[a], Ming-Chen Hsu[b]

[a]*Department of Structural Engineering, University of California, San Diego, La Jolla, CA 92093, USA*
[b]*Department of Mechanical Engineering, Iowa State University, Ames, IA 50011, USA*
[c]*School of Aerospace and Mechanical Engineering, The University of Oklahoma, Norman, OK 73019, USA*
[d]*Department of Mechanical Engineering, Northwestern University, Evanston, IL 60208, USA*

Abstract

This work formulates frictionless contact between solid bodies in terms of a repulsive potential energy term and illustrates how numerical integration of the resulting forces is computationally similar to the "pinball algorithm" proposed and studied by Belytschko and collaborators in the 1990s. We thereby arrive at a numerical approach that has both the theoretical advantages of a potential-based formulation and the algorithmic simplicity, computational efficiency, and geometrical versatility of pinball contact. The singular nature of the contact potential requires a specialized nonlinear solver and an adaptive time stepping scheme to ensure reliable convergence of implicit dynamic calculations. We illustrate the effectiveness of this numerical method by simulating several benchmark problems and the structural mechanics of the right atrioventricular (tricuspid) heart valve. Atrioventricular valve closure involves contact between every combination of shell surfaces, edges of shells, and cables, but our formulation handles all contact scenarios in a unified manner. We take advantage of this versatility to demonstrate the effects of chordal rupture on tricuspid valve coaptation behavior.

Keywords: Heart valves; Contact problems; Pinball algorithm; Isogeometric analysis

Contents

---

*Corresponding author
Email address: dmkamensky@eng.ucsd.edu (David Kamensky)

1. Introduction

The atrioventricular heart valves (i.e., the mitral and tricuspid valves) are structures in the heart that permit blood flow from the atria into the ventricles during diastole, as the heart relaxes. They then block blood from being pumped back out when the heart contracts, during systole. The semilunar valves (i.e., the aortic and pulmonary valves) play the opposite role, permitting blood to be pumped out of the ventricles, into the aorta and pulmonary artery, during systole, and preventing backflow into the ventricles during diastole.

The semilunar valves function in a relatively simple way: each one consists of three (or occasionally two) cup-like leaflets (cusps), attached to the sides of its respective artery. The convex sides of the leaflets face the ventricles, so that the leaflets fill up in reaction to backflow. When

2

this happens, the convex sides of the leaflets contact one another, blocking flow. The mechanism of the atrioventricular valves is much more complicated. They are generally considered to have several distinct leaflets, but, in reality, these leaflets form contiguous sheets of tissue. These contiguous structures crumple up against themselves during systole to block off the channels between the ventricles and the atria. In as much as 3% of the human population [1], parts of an atrioventricular valve may even prolapse into the atrium. The atrioventricular valves' mechanics are further complicated by a web-like network of tendons—the chordae tendineae, or "heart strings"—which tether the leaflets to the inside of the ventricles.

Computer simulations of the mechanics of heart valves play an increasingly important role in research into treatments for valvular heart disease. The need for and history of such simulations are reviewed in [2, 3]. As one might imagine, the treatment of contact mechanics is a major concern in these computations and proves to be especially difficult in the case of atrioventricular valves. Our previous work on aortic valve simulations [4–7] pointed out that to "avoid stagnation in pathological configurations, we typically select the resolution of the nonlinear algebraic solution by choosing a fixed number of iterations rather than a percentage by which the residual must be reduced" [4, page 1028]. Subsequent autopsies of poorly-converged steps indicated that contact forces were to blame for these "pathological configurations".

Similar disclaimers can be found in other literature on aortic valve analysis. For example, to ensure completion of forward solves in an inverse modeling procedure, Aggarwal and Sacks [8] adhered to the following policy for terminating augmented Lagrangian iteration for contact: "After 10 augment iterations, the solution at current load was assumed to be converged irrespective of the gap and Lagrange multiplier values, and simulation was continued onto the next time step" [8, page 917]. It is clear that simulation of even the relatively simple contact of the aortic valve is fraught with difficulties. This has led some authors to instead apply brute-force explicit solution procedures that do not require solving a nonlinear algebraic problem at each step. However, Morganti et al. [9] reported that converging an explicit computation to a qualitatively-accurate configuration of a closed valve using standard finite elements in the thoroughly-optimized commercial code LS-DYNA [10] required more than 550 hours [9, Table 2] (i.e., more than 20 days). While the cited study indicates that using LS-DYNA's isogeometric analysis capabilities in place of finite elements provided enormous improvements over that figure, the coarsest isogeometric computation still required a time step of 2.3×10^{-7} s [9, Table 2], which is several orders of magnitude smaller than the time step used for fluid–structure interaction (FSI) simulations in [4].

In light of these difficulties with aortic valve simulations, the outlook appears bleak for properly-converged implicit simulation of the more complicated atrioventricular valves, especially if one intends to study pathological configurations, such as prolapse, or incorporate FSI. Some example analyses of normal mitral valve structural mechanics using ABAQUS [11] can be found

in [12–14]. Lee et al. [13, Section 10.4.3] indicate that "implicit dynamic analysis with automatic time stepping" was feasible in some cases, but the same authors note in [14, Section 2.6] that explicit dynamics using a time step of 10^{-6} s and mass scaling (as only the quasi-static limit was of interest) proved to be "a more general computational framework for future extensions, such as in vivo modeling and surgical simulations" [14, page 1287]. It is our experience that, even with the best commercial off-the-shelf software, some degree of case-by-case troubleshooting is frequently needed to solve complex contact problems using implicit methods. In the case of atrioventricular valve analysis, preprocessing and troubleshooting labor may be magnified by the involvement of contact between different combinations of 2D surfaces (leaflets) and 1D curves (representing the chordae tendineae and/or the edges of leaflets). Analysis methods that treat these different contact scenarios as separate cases are not only difficult to implement, but also often require additional pre-processing effort by the end user, to identify contact surfaces and curves, specify types of contact, and set a variety of options for each case (often by trial-and-error).

To approach the challenging problem of atrioventricular valve dynamics in a way that can be directly combined with the fluid–structure interaction (FSI) analysis methods of [4–7], which treat structural mechanics implicitly, we develop a novel contact formulation that builds on the concept of using repulsive forces that derive from a potential, as studied by Sauer and De Lorenzis [15]. We modify this framework to use volume rather than surface potentials, which can be applied to lower-dimensional objects (e.g., shells or cables) by multiplying surface or contour integrals over those objects with factors of thickness or cross-sectional area. This eliminates the need for special cases to handle contacts between objects of different co-dimensions to physical space. We discretize the integrals to compute contact potential energy using direct numerical quadrature (as applied to the case of electrostatic repulsion in [15, Section 5.3]), resulting in an n-body problem to determine contact forces. In general, this would require a Barnes–Hut [16] or fast multipole [17] approximation for computational efficiency. However, we simplify the n-body problem by using a compactly-supported potential with a limited range, so that each quadrature point interacts with only a handful of neighbors.[1] The geometry of the contacting bodies is thus approximated by a union of spherical potential supports, intersections of which are penalized in the potential energy functional for the system. The resulting discrete problem is similar in structure to that solved by the "pinball algorithm" proposed in Belytschko and Neal [20], but we use a different penalty function to apply forces to intersecting pinballs (i.e., supports of potential functions centered at quadrature points). Another related approach to contact was developed earlier by Sauer [21], motivated by coarse-graining of inter-atomic potentials, but, in the present work, we do not attempt

---

[1]Another interesting approach to avoiding nonlocal interactions is to store contact potential energy as elastic potential energy of a fictitious material filling the void between objects [18, 19].

to derive contact forces from any more basic physical principles; our choice of penalty function is motivated by calculations indicating the minimal level of singularity needed to prevent intersection of contacting bodies. We find that this potential-based contact formulation, combined with a specialized globalization strategy for nonlinear solution using Newton's method and an adaptive time stepping technique, provides a reliable way to solve complicated contact problems with shell and cable structures, including atrioventricular valve dynamics.

The remainder of this paper is structured as follows. The formulation of the contact potential energy is detailed in Section 2, and its implementation using numerical quadrature is discussed in Section 3. The generic structural dynamics problem used for exposition in these sections is then specialized to the dynamics of a thin shell coupled to cables in Section 4. The contact formulation is tested in the context of this problem using several benchmarks. The right atrioventricular (i.e., tricuspid) valve is then modeled as an instance of this thin shell and cable problem in Section 5. Both normal and pathological tricuspid valve dynamics are successfully simulated. Section 6 summarizes our findings and discusses several other possible applications and extensions of the developed techniques.

2. Volume potentials for contact

To treat contact between objects of various co-dimensions to physical space, we first consider the case of solid bodies of co-dimension zero, then scale by factors of thickness and cross-sectional area to adapt the theory to thin shells of co-dimension one and cables of co-dimension two. We adopt this approach, in contrast to the surface potential framework of Sauer and De Lorenzis [15], for maximal geometric versatility. While the use of a surface potential allows for analytical simplification in the case of smooth surfaces [15, Appendix A], it requires additional information about the surface geometry. Formulation in terms of a volume potential permits contact to be modeled between objects represented by clouds of quadrature points, requiring no special treatment for contact between structures of arbitrary co-dimensions, with corners, edges, and other features that would complicate traditional contact algorithms. Further, volume potential based contact is compatible with meshfree discretizations that allow new surfaces to emerge spontaneously, through material rupture (e.g., the semi-Lagrangian reproducing kernel particle method (RKPM) [22]). It might be construed as a disadvantage that volume potentials result in a thin layer of body forces, as opposed to a true surface traction, but this force field could easily be integrated in the direction orthogonal to the surface as a post-processing step if surface tractions are desired. Alternatively, for bodies with well-defined surface parameterizations and no possibility of rupture, contact surfaces could be treated like structures of co-dimension one, by introducing a thickness length scale.

For simplicity of exposition, we assume, initially, that contact occurs between two disjoint bodies. The extension to multiple bodies is straightforward. The extension to self-contact is for-

mulated in Section 2.2. The contact potential energy of two bodies with reference configurations Ω_0^1 and Ω_0^2 is given by

$$E_c = \int_{\Omega_0^2} \int_{\Omega_0^1} \phi(r_{12}) \, d\mathbf{X}_1 d\mathbf{X}_2 \,, \tag{1}$$

where $\mathbf{X}_i \in \Omega_0^i$ and

$$r_{12} = \|(\mathbf{X}_2 + \mathbf{y}(\mathbf{X}_2)) - (\mathbf{X}_1 + \mathbf{y}(\mathbf{X}_1))\|_{\ell^2} = \|\mathbf{r}_{12}\|_{\ell^2} \,, \tag{2}$$

in which $\mathbf{y}$ is a displacement field defined over $\Omega_0^1 \cup \Omega_0^2$ and $\|\mathbf{v}\|_{\ell^p} = \left(\sum_{i=1}^{d} |v_i|^p\right)^{\frac{1}{p}}$ for $\mathbf{v} \in \mathbb{R}^d$ and $p \geq 1$. To allow objects of nonzero co-dimension to simultaneously contact other objects from any direction, we must strongly enforce non-penetration of objects. This requires that $E_c \to \infty$ as the distance between $\partial\Omega_t^1$ and $\partial\Omega_t^2$ approaches zero.

Remark 1. This potential energy term models frictionless contact. The absence of frictional effects is clear, as the formulation does not introduce any dissipative mechanisms. For bodies with smooth boundaries, and with typical choices of ϕ, one expects the resulting contact forces to be approximately normal to the boundaries. A detailed explanation of how tangential force contributions cancel is provided by Sauer and De Lorenzis [15, Appendix A] for the case of surface potentials. The cited article also comments briefly on how potential-based contact can be extended to include friction.

Remark 2. Enforcement of a gap between objects may also be useful in conjunction with immersed discretization schemes, in which the motion of a Lagrangian solid is derived from a velocity field represented on an Eulerian background mesh, either directly (e.g., [23, 24]), or through the computation of some coupling traction (e.g., [4, 25]). For example, in the standard material point method [24], a single continuous Eulerian velocity field represents the structure's motion, so sliding contact proves difficult unless material points from different parts of the structure are prevented from coming within a background mesh dependent distance of one another. In the setting of immersed FSI, Kadapa et al. [26] report that stable simulation of fluid–rigid-body interaction using an immersed approach based on Nitsche's method requires that the rigid bodies maintain a background mesh dependent distance between one another. The enforcement of a gap between immersed representations of structural components may seem to be at odds with the oft-touted advantage of immersed discretizations in representing topology changes. However, if gap width is tuned to match the size of elements in the background discretization, then it is, by construction, insufficient to resolve any spurious flow of material. (Thus, in a strange twist of irony, the numerical methods most capable of representing topology change tend to require it the least strictly.)

2.1. Selection of a potential

For computational tractability, we require that potentials have a finite range. To enforce non-penetration, we also clearly need $\phi(r) \to \infty$ as $r \to 0$. However, too weak a singularity might
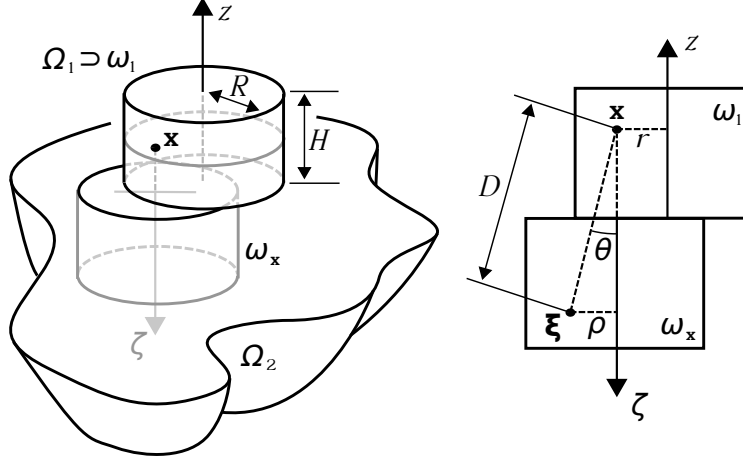
6

Figure 1: Notation for the integral used to show that $p \geq 4$ is sufficient for the z-component of force on ω_1 to diverge.

still permit interpenetration of objects. For instance, consider a finite-thickness slab of electrically-charged material (i.e., $\phi(r) \sim 1/r$): a simple application of Gauss's theorem shows that the resulting electric field perpendicular to the slab never diverges, so that another charged object might tunnel through this slab with a finite amount of work (cf. [15, Section 3, Example 11, Remark 1]). To determine a sufficient strength of singularity for ϕ, let us consider smooth (or, at small scales, essentially flat) surfaces of Ω_1 and Ω_2 in contact. Our goal in designing ϕ is for the total repulsive force in the normal direction acting on $\omega_1 \subset \Omega_1$, as illustrated in Figure 1, to diverge.

Let us assume that, for r less than some $O(1)$ length scale, the potential has the form

$$\phi(r) \sim \frac{1}{r^{p-1}} \qquad \text{(for } r \text{ sufficiently small).} \tag{3}$$

As mentioned above, the obvious inadequacy of electrostatic repulsion for contact modeling leads us to the assumption that $p > 2$. The z-component of the force on ω_1 can be bounded from below as follows:

$$F_z \gtrsim \int_{\omega_1} \left(\int_{\omega_{\mathbf{x}}} -\phi'(D) \cos \theta \, d\boldsymbol{\xi} \right) d\mathbf{x} , \tag{4}$$

where the integration variables $\mathbf{x}$ and $\boldsymbol{\xi}$ have the axial and radial coordinates (r, z) and (ρ, ζ) respectively, as shown in Figure 1, and D and θ depend on the integration variables $\mathbf{x}$ and $\boldsymbol{\xi}$ via

$$D = \sqrt{\rho^2 + (z + \zeta)^2} \tag{5}$$

and

$$\cos \theta = (z + \zeta) / \sqrt{\rho^2 + (z + \zeta)^2} . \tag{6}$$

This is only a lower bound because it neglects additional repulsive forces on $\mathbf{x} \in \Omega_1$ due to $\Omega_2 \setminus \omega_{\mathbf{x}}$.

Integrating first over $\omega_{\mathbf{x}}$, we get

$$\int_{\omega_{\mathbf{x}}} \phi'(D) \cos\theta \, d\boldsymbol{\xi} \sim \int_0^H \int_0^R \frac{2\pi\rho(z+\zeta)}{(D^2)^{\frac{p+1}{2}}} \, d\rho d\zeta \tag{7}$$

$$= \int_0^H \left(-\frac{\left(\frac{p+1}{2} - 1\right)^{-1} \pi(z+\zeta)}{(D^2)^{\frac{p+1}{2}-1}} \Bigg|_0^R \right) d\zeta \tag{8}$$

$$= \int_0^H \left(\frac{C_1(z+\zeta)}{\left((z+\zeta)^2\right)^{\frac{p+1}{2}-1}} + C_2 \right) d\zeta \tag{9}$$

$$= \begin{cases} C_3 \log\left((z+\zeta)^2\right) + C_4 & p = 3 \\ -\dfrac{C_5}{\left((z+\zeta)^2\right)^{\frac{p+1}{2}-2}} + C_6 & p \in (2,\infty) \setminus \{3\} \end{cases} \Bigg|_0^H \tag{10}$$

$$= \begin{cases} -C_3 \log\left(z^2\right) + C_7 & p = 3 \\ C_8 z^{3-p} + C_9 & p \in (2,\infty) \setminus \{3\} \end{cases}, \tag{11}$$

where C_i are (possibly non-constant) finite coefficients and terms whose precise values are not important to the question of whether or not F_z is infinite. Integrating next over ω_1, we obtain

$$F_z \gtrsim \int_{\omega_1} \left\{ \begin{array}{ll} -C_3 \log\left(z^2\right) + C_7 & p = 3 \\ C_8 z^{3-p} + C_9 & p \in (2,\infty) \setminus \{3\} \end{array} \right\} d\mathbf{x} \tag{12}$$

$$= \pi R^2 \int_0^H \left\{ \begin{array}{ll} -C_3 \log\left(z^2\right) + C_7 & p = 3 \\ C_8 z^{3-p} + C_9 & p \in (2,\infty) \setminus \{3\} \end{array} \right\} dz . \tag{13}$$

To ensure that this lower bound diverges, we need to take $p \geq 4$. To have the desired asymptotic behavior in the limit of $r_{12} \to 0$ while maintaining a smooth but compactly-supported force field, we propose to model contact with the following force–separation law:

$$-\phi'(r_{12}) = \begin{cases} \dfrac{k_{\mathrm{c}}}{(r_{12})^p} - c_2 & r_{12} < r_{\mathrm{in}} \\ c_1 \left(r_{12} - r_{\mathrm{out}}\right)^2 & r_{\mathrm{in}} \leq r_{12} < r_{\mathrm{out}} \\ 0 & \text{otherwise} \end{cases} . \tag{14}$$

where r_{in} and r_{out} are length scales governing the support of the potential function, and k_{c} is a dimensional constant. As discussed above, we assume $p \geq 4$, to enforce non-penetration. The scalars c_1 and c_2 are uniquely determined by smoothness and continuity constraints:

$$c_1 = \frac{p k_{\mathrm{c}}}{2 \left(r_{\mathrm{out}} - r_{\mathrm{in}}\right) r_{\mathrm{in}}^{p+1}} \tag{15}$$

8
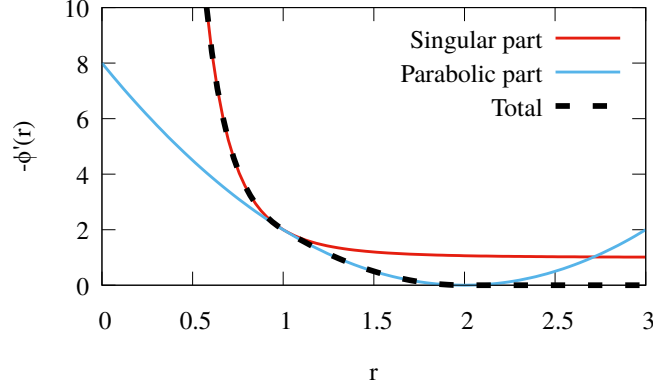
Figure 2: The function (14) with $k_\mathrm{c} = 1$, $p = 4$, $r_\mathrm{in} = 1$, and $r_\mathrm{out} = 2$.

and

$$c_2 = \frac{k_\mathrm{c}}{r_\mathrm{in}^p} - c_1 \left(r_\mathrm{in} - r_\mathrm{out}\right)^2 \ . \tag{16}$$

An instance of this force law is illustrated in Figure 2.

As alluded previously, the extension of this concept to structures of nonzero co-dimension is accomplished by essentially performing one-point quadrature in the degenerated direction(s), to obtain surface, line, or point potentials, scaled by appropriate length, area, or volume factors. The conditions for non-penetration of potentials integrated on lower-dimensional geometries may be less severe. For instance, as shown in [15, Appendix A], the force field due to a less-singular $1/r^2$ potential integrated over a smooth surface is sufficient to produce divergent repulsive forces. The even milder electrostatic $1/r$ potential of an infinite line of charge can easily be shown, using Gauss's theorem, to result in an electric field that diverges as one approaches the wire. However, these results are not robust for surfaces and curves with edges and endpoints, as often encountered in practical examples.

2.2. Extension to self-contact

In many cases of practical interest, there is no clear way to partition a structure into distinct parts that may contact one another; a body must be capable of contacting itself. To accomplish this, we consider all solid parts to be a single body with reference configuration Ω_0 and modify the potential energy of contact to be

$$E_\mathrm{c} = \int_{\Omega_0 \setminus B_{R_\mathrm{self}}(\mathbf{X}_1)} \int_{\Omega_0} \phi\left(r_{12}\right) \, d\mathbf{X}_1 d\mathbf{X}_2 \ , \tag{17}$$

where $B_{R_\mathrm{self}}(\mathbf{X}_1)$ is the Euclidean ball of radius R_self around $\mathbf{X}_1$ and r_{12} is defined as before, but with both $\mathbf{X}_1$ and $\mathbf{X}_2$ in Ω_0. The radius R_self must be selected in a problem-dependent way. It is clearly bounded below by the range of the contact potential, but, in general, must be somewhat larger, to

prevent material compression from being identified as self-contact. If chosen too large, however, it might fail to identify some instances of self-contact.

3. Numerical implementation

In this section, we consider various aspects of implementing the proposed approach in finite element simulations. For simplicity, we consider a generic structural dynamics problem on the time interval $(0, T)$. A specific example will be formulated in Section 5. The generic problem we consider in this section is: Find displacement $\mathbf{y} \in \mathcal{S}_y$ such that for all test functions $\mathbf{z} \in \mathcal{V}_y$

$$B(\mathbf{y}, \mathbf{z}) + D_\mathbf{z} E_\mathrm{c}(\mathbf{y}) = F(\mathbf{z}) \,, \tag{18}$$

where the forms B and F encapsulate the non-contact-related physics of the problem, $\mathcal{S}_y$ and $\mathcal{V}_y$ are displacement trial and test function spaces, and $D_\mathbf{z}$ is a functional derivative in direction $\mathbf{z}$. Restricting $\mathbf{y}$ and $\mathbf{z}$ to $\mathbf{y}^h$ and $\mathbf{z}^h$ in finite-dimensional spaces $\mathcal{S}_y^h$ and $\mathcal{V}_y^h$ results in a finite element semi-discretization of the first term on the left-hand side of (18). We focus in this section on the discretization of the second term, which generates the contact forces.

We propose to compute the double integral in the definition of E_c numerically, using the same collection of quadrature points for each layer of integration. This results in a discrete N_p-body problem:

$$E_\mathrm{c}\left(\mathbf{y}\right) \approx \sum_{i=1}^{N_p} \left(\sum_{j=1}^{N_p} \left\{ \begin{array}{ll} w_i w_j \phi\left(r_{ij}\right) & R_{ij} > R_\mathrm{self} \\ 0 & \text{otherwise} \end{array} \right\} \right) \,, \tag{19}$$

where N_p is the number of quadrature points,

$$r_{ij} = \left\| (\mathbf{X}_i + \mathbf{y}\left(\mathbf{X}_i\right)) - \left(\mathbf{X}_j + \mathbf{y}\left(\mathbf{X}_j\right)\right) \right\|_{\ell^2} \,, \tag{20}$$

$$R_{ij} = \left\| \mathbf{X}_i - \mathbf{X}_j \right\|_{\ell^2} \,, \tag{21}$$

$\mathbf{X}_i$ is the ith quadrature point and w_i is its weight. The weights $\{w_i\}$ are assumed to include any additional length, area, or volume factors needed to convert surface, line, or point quadrature rules into volume quadrature. In this way, there is no difference in treatment between various co-dimensions of structural components when computing contact forces.

To ensure that the singular behavior is recovered as the distance between bodies shrinks all the way to zero, one would need to employ an adaptive quadrature rule; otherwise, a point from Ω_2 could slip in between two points on the surface of Ω_1 with a finite amount of work, making it possible for Ω_2 to tunnel through Ω_1, due to quadrature errors. Alternatively, one can prevent tunneling of objects through one another with a fixed quadrature rule by including an impenetrable
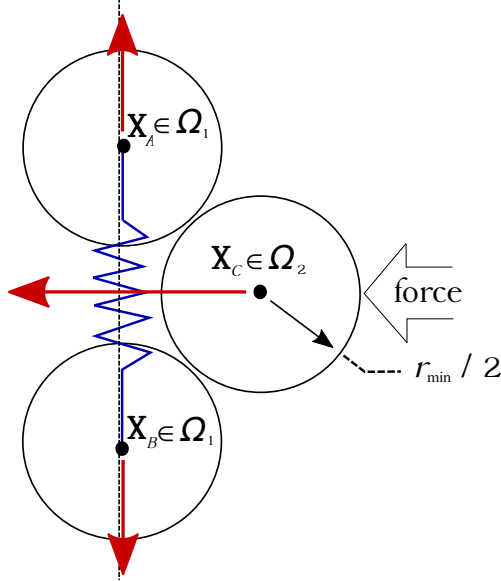
Figure 3: Suppose that quadrature points $\mathbf{X}_A$ and $\mathbf{X}_B$ from body Ω_1 are constrained to slide along the dashed line and coupled by an elastic force, represented by the blue spring. Even with $r_{\min} > 0$ and initially-overlapping regions of infinite penalty for points $\mathbf{X}_A$ and $\mathbf{X}_B$, the contacting point $\mathbf{X}_C$ from Ω_2 can drive them apart if pushed with sufficient force.

core of infinite potential in the force–separation law. Consider the following modified force law:

$$
-\tilde{\phi}'(r_{12}) = \begin{cases} \infty & r < 0 \\ -\phi'(r) & \text{otherwise} \end{cases}, \tag{22}
$$

where $r = r_{12} - r_{\min}$ and ϕ' is as defined in (14). The support of the modified potential is of radius $r_{\min} + r_{\text{out}}$. (A procedure to avoid states of infinite energy during iterative nonlinear minimization of energy is given in Section 3.4.) If the zones of infinite potential about quadrature points in a body overlap, then the body is essentially impenetrable. For elastic bodies, tunneling is still technically possible, as adjacent points may be stretched apart, as illustrated schematically in Figure 3, but this would likely involve stresses well beyond the ultimate strength of most practical materials. We have nevertheless been able to produce this behavior by mashing objects together under displacement control. A similar concern is illustrated in [21, Figure 5.7].

3.1. Searching for interacting pairs of points

Due to the limited range of ϕ, the forces in the N_p-body problem (19) can be evaluated efficiently. Assuming that the spatial number density of points is bounded below, there are only $O(N_p)$ nonzero terms in the double summation (19). We can avoid considering unnecessary pairs of points by sorting all of the quadrature points into the cells of a uniform grid, of cell size greater than or equal to the range $r_{\text{out}} + r_{\min}$ of the potential. Then we only need to consider interactions between

one point and others in the same or adjacent grid cells.

To describe the algorithm used precisely, we introduce some notation. Let the number of grid cells be N_c. These cells are a collection of disjoint subsets of $\mathbb{R}^d$ denoted $\{\omega_i\}_{i=1}^{N_c}$ and cover the entire deformed configuration of the solid. The locations of the quadrature points in the current configuration are $\{\mathbf{x}_i\}_{i=1}^{N_p}$. We assume, as is especially true in the case of structures with nonzero co-dimension, that, while N_c may be very large, the vast majority of these grid cells are unoccupied at any given time. Thus operations of cost $\Theta(N_c)$[2] (such as iterating over all elements of $\{\omega_i\}$) should be avoided. Aside from allocating memory for the grid cells while initializing the data structure, operations to re-sort the points $\{\mathbf{x}_i\}$ into new cells as the structure deforms should only iterate over occupied cells, of which there can be at most N_p. Operations of cost $\Theta(N_p)$ are necessarily acceptable; otherwise, it would be impossible to apply the contact forces, regardless of how they are computed.

We now introduce the necessary data structures for our algorithm:

- L is an array of N_p integers. This will be used to store linked lists of indexes of quadrature points. Because points fall within unique grid cells, these lists can all fit within a single array of size N_p. The storage and traversal of these lists is clarified below.

- S is a stack of integers. This is used to keep track of the indexes of occupied grid cells. S is initialized to be empty.

- O is an array of N_c Boolean variables. $O(i)$[3] indicates whether or not grid cell i is occupied. O is initialized to false once, when it is first allocated, costing $\Theta(N_c)$ operations.

- H is an array of N_c integers. If $O(i)$ is true, then $H(i)$ indicates the head of a linked list of indexes of points from the set $\{\mathbf{x}_j\}$ contained in grid cell ω_i. The array H is initialized to zero once, when it is first allocated, costing $\Theta(N_c)$ operations.

The procedure for sorting elements of $\{\mathbf{x}_i\}$ into grid cells $\{\omega_i\}$ after the structure changes its configuration can be divided into two phases: purge the data structures of information about the previous configuration, then populate the data structures with information about the new configuration. The first phase, purging, is accomplished through the following steps:

1. Set the list L to zero. (Cost: $\Theta(N_p)$, because L is of size N_p.)

2. While S is not empty, pop index i from S, then set $O(i)$ to false and $H(i)$ to zero. (Cost: $O(N_p)$, because at most N_p cells can be occupied by N_p points, assuming each is in a unique cell.)

---

[2]Asymptotic notation in this section follows standard computer science conventions. See, e.g., [27].

[3]Note that, in the present document, this is distinguished from "big $\mathcal{O}$" asymptotic notation by font.

The second phase is accomplished by doing the following for each element $\mathbf{x}_k$ of the set $\{\mathbf{x}_i\}$:

1. Identify the index i such that $\mathbf{x}_k \in \omega_i$. This can be accomplished in $O(1)$ time, because the $\{\omega_j\}$ are elements of a uniform grid.

2. If $O(i)$ is false, set it to true and push i onto the stack S. (If $O(i)$ is already true, do not redundantly push i.) This step also has cost $O(1)$.

3. Update the linked list of points within this cell:

   (a) Store the old head of the linked list: $h_0 \leftarrow H(i)$.

   (b) Set $\mathbf{x}_k$'s index as the new head: $H(i) \leftarrow k$.

   (c) Connect the new head to the linked list: $L(k) \leftarrow h_0$.

   These steps cost $O(1)$ time.

The cost of updating the data structures to reflect a change in the configuration of the structure is therefore $\Theta(N_p)$. Once the data structures have been populated by the above algorithm, the list of points from $\{\mathbf{x}_j\}$ contained in an occupied cell ω_i can be traversed by looking at the point $\mathbf{x}_{i_1}$, where $i_1 = H(i)$, then looking at $\mathbf{x}_{i_2}$, where $i_2 = L(i_1)$, and so on, until $L(i_n) = 0$, indicating the end of the list.[4] If we assume that the points $\{\mathbf{x}_i\}$ are distributed such that the number falling within any cell is $O(1)$, then the cost of computing forces between all points in a cell and any adjacent occupied cells is $O(1)$.

Remark 3. If the memory cost of storing H and O is prohibitive, random subsets of $\{\omega_i\}$ can be mapped to bins of a hash table, through some nontrivial hash function mapping the cell index range $\{1, \ldots, N_c\}$ onto a smaller range $\{1, \ldots, N_H\}$, with $N_H < N_c$. However, increasing the load factor (N_p/N_H) in this way comes at the cost of possibly computing moderately more zero forces between out-of-range pairs of points. We have not found it necessary to use hash functions other than the identity map in our computations, as memory was never a limiting factor.

3.2. Comparison with the pinball algorithm

As mentioned in Section 1, the fully-discrete penalization energy (19) can be construed as a variant of the pinball algorithm proposed by Belytschko and Neal in [20]. The idea of the pinball approach is to approximate the geometry of a structure by a collection of balls, referred to as "pinballs", for the purposes of collision detection. Overlap of pinballs from different contacting bodies is then penalized, to prevent structures from intersecting one another. A refinement of

---

[4]This structure is inspired by the cluster maps used in FAT file systems.

this approach by Belytschko and Yeh, called the "splitting pinball method" [28, 29], tested for collisions using a hierarchical approach, such that, if two pinballs are detected as overlapping, each of them is split into a collection of finer pinballs (which may themselves be split, and so on, up to some user-defined recursion depth), to better approximate the underlying structural geometry.

Despite its simplicity and intuitive appeal, the pinball method remains relatively obscure today. It is, however, implemented in several actively-maintained production codes for explicit dynamics, such as EUROPLEXUS [30] and LS-DYNA [10]. According to the LS-DYNA manual [31], the original pinball method is "not recommended and is included only for back compatibility" [31, page 11-64], but the splitting variant [28, 29], added fairly recently in LS-DYNA R7.1, "is recommended when modeling complex contacts" [31, page 11-65]. Approaches with some similar features to the pinball algorithm have been developed independently for use in meshfree methods, where it is natural to consider the structure geometry as a collection of spheres and use these spheres to detect collisions. An example of such a method would be the kernel contact approach proposed by Guan et al. [22, Section 4].

We may view each quadrature point in our discrete formulation as a pinball of radius (r_{out} + r_{min})/2. The original paper on the pinball method [20] includes a separate definition of the surface normal vector at each pinball, but, in the later splitting approach, "The penalty force is always exerted in the direction between pinball centers, which enables the method to handle edge-to-surface, surface-to-surface and edge-to-edge contact" [29, page 382], which is more similar to what is done in the potential-based formulation of this paper. If we reinterpret the pinball approach as a quadrature scheme for a distributed contact force, the splitting variant is essentially adaptive quadrature.

3.3. Linearization of pairwise interactions

The interaction energy of two quadrature points at material positions $\mathbf{X}_1$ and $\mathbf{X}_2$ (separated by more than R_{self} at time zero) can be written as

$$e_{12} = C\phi\left(r_{12}\right) , \tag{23}$$

where the scalar C accounts for quadrature weights of the two points. The contribution of this pair of points to the variational problem (18) is

$$D_{\mathbf{z}}e_{12} = \left.\frac{d}{d\epsilon}e_{12}(\mathbf{y} + \epsilon\mathbf{z})\right|_{\epsilon=0} = C\phi'\left(r_{12}\right)\hat{\mathbf{r}}_{12} \cdot (\mathbf{z}_2 - \mathbf{z}_1) , \tag{24}$$

14

where $\hat{\mathbf{r}}_{12} = \mathbf{r}_{12}/\|\mathbf{r}_{12}\|_{\ell^2}$ and $\mathbf{z}_i = \mathbf{z}(\mathbf{X}_i)$. The linearization of this virtual work (needed to resolve static or dynamic equilibrium via Newton–Raphson iteration) is the Hessian

$$D_{\Delta\mathbf{y}}(D_{\mathbf{z}}e_{12}) = \left.\frac{d}{d\epsilon}(D_{\mathbf{z}}e_{12})(\mathbf{y} + \epsilon\Delta\mathbf{y})\right|_{\epsilon=0} \tag{25}$$

$$= C\phi''(r_{12})(\Delta\mathbf{y}_2 - \Delta\mathbf{y}_1)(\hat{\mathbf{r}}_{12} \otimes \hat{\mathbf{r}}_{12})(\mathbf{z}_2 - \mathbf{z}_1)$$

$$+ C\frac{\phi'(r_{12})}{r_{12}}(\Delta\mathbf{y}_2 - \Delta\mathbf{y}_1)(\mathbf{I} - \hat{\mathbf{r}}_{12} \otimes \hat{\mathbf{r}}_{12})(\mathbf{z}_2 - \mathbf{z}_1) , \tag{26}$$

where $\mathbf{I}$ is the identity map and $\Delta\mathbf{y}_i = \Delta\mathbf{y}(\mathbf{X}_i)$.

3.4. Solution algorithm

The singular nature of the contact potential warrants special care in the time stepping and non-linear algebraic solution procedures for discretized instances of our generic problem (18). We find that reliably obtaining "fire and forget" robustness in computations using our proposed formulation requires a combination of adaptive time stepping and modifications to Newton's method for nonlinear solution within each time step. For generality, we spell out our solution algorithm for use with the modified potential (22); the potential (14) can be recovered by taking $r_{\min} \to 0$.

Let us suppose that the generic problem (18) represents nonlinear elasticity in the Lagrangian description. Then we write

$$B(\mathbf{y}, \mathbf{z}) = \rho_0\ddot{\mathbf{y}} + A_1(\mathbf{y}, \mathbf{z}) + A_2(\dot{\mathbf{y}}, \mathbf{z}) . \tag{27}$$

If we discretize in time using the generalized-α procedure, (18) yields the following problem within each time step: Given $\mathbf{y}^n$, $\dot{\mathbf{y}}^n$, and $\ddot{\mathbf{y}}^n$, find $\mathbf{y}^{n+1}$ such that, for all $\mathbf{z}$,

$$R(\mathbf{y}^{n+1}, \mathbf{z}) = \int_{\Omega_0} \rho_0\ddot{\mathbf{y}}^{n+\alpha_m} \cdot \mathbf{z}\, d\Omega + A_1\left(\mathbf{y}^{n+\alpha_f}, \mathbf{z}\right) + A_2(\dot{\mathbf{y}}^{n+\alpha_f}, \mathbf{z}) + D_{\mathbf{z}}E_c\left(\mathbf{y}^{n+\alpha_f}\right) - F(\mathbf{z}) = 0 , \tag{28}$$

where ρ_0 is the mass density in the reference configuration Ω_0 and $\ddot{\mathbf{y}}^{n+\alpha_m}$, $\dot{\mathbf{y}}^{n+\alpha_f}$, and $\mathbf{y}^{n+\alpha_f}$ are the "α-level" acceleration, velocity, and displacement (interpolating between the n and $n+1$ level data), which can be expressed in terms of the unknown function $\mathbf{y}^{n+1}$ and the data from the previous time step. The velocity and acceleration $\dot{\mathbf{y}}^{n+1}$ and $\ddot{\mathbf{y}}^{n+1}$ are also defined in terms of $\mathbf{y}^{n+1}$ and n-level quantities, thus providing all data needed for the subsequent step; full formulas are given in [32], using the same notation. In this work, we restrict the selection of generalized-α parameters to the one-parameter family of methods detailed in [32], where the parameter $\rho_\infty \in [0, 1]$ controls the spectral radius of the amplification matrix in the high-frequency limit. If we pose the problem (28) over finite dimensional test and trial spaces, with the test space spanned by the basis $\{\mathbf{N}_1, \ldots, \mathbf{N}_{N_{\text{DOF}}}\}$, we get a nonlinear algebraic problem for N_{DOF} unknown coefficients. For mild instances of the

problem, starting from good initial guesses, this can be resolved using Newton iteration. In our notation, the standard Newton iteration procedure would begin with $\mathbf{y}_0^{n+1}$ and, at the kth iteration would execute: Find $\Delta\mathbf{y}_k$, such that for all $A \in \{1, \ldots, N_{\text{DOF}}\}$,

$$D_{\Delta\mathbf{y}_k} R\left(\mathbf{y}_k^{n+1}, \mathbf{N}_A\right) = -R\left(\mathbf{y}_k^{n+1}, \mathbf{N}_A\right) , \tag{29}$$

then set

$$\mathbf{y}_{k+1}^{n+1} = \mathbf{y}_k^{n+1} + \Delta\mathbf{y}_k . \tag{30}$$

The iteration continues until

$$\frac{\|\mathbf{R}_k\|_{\ell^2}}{\left\|\mathbf{F}_0^k\right\|_{\ell^2}} \leq \epsilon , \tag{31}$$

or until some divergence criterion is encountered, where

$$[\mathbf{R}_k]_A = R\left(\mathbf{y}_k^{n+1}, \mathbf{N}_A\right) , \tag{32}$$

$\mathbf{F}_0^k \in \mathbb{R}^{N_{\text{DOF}}}$ is a reference force vector, and $\epsilon > 0$ is a dimensionless tolerance. The best choice of $\mathbf{F}_0^k$ is problem-dependent; a common choice for static problems or problems with prescribed loads in all time steps is to set its components equal to the external force form F applied to each basis function of the test space.

Newton iteration is only *locally* convergent, meaning that, if $\mathbf{y}_0^{n+1}$ is too far from a fixed point, the iteration may fail to converge. One way to increase the radius of convergence is to use a line search, which modifies (30) to

$$\mathbf{y}_{k+1}^{n+1} = \mathbf{y}_k^{n+1} + \alpha_{\text{relax}}\Delta\mathbf{y}_k , \tag{33}$$

where the scalar α_{relax} is selected to minimize $\|\mathbf{R}_{k+1}\|_{\ell^2}$. Solving for α_{relax} generally requires repeated assembly of the residual vector, to execute a 1D minimization scheme. A cheaper alternative is to compute α_{relax} in a heuristic way, with the goal that $\|\mathbf{R}_{k+1}\|_{\ell^2}$ is never "too big". We have developed a heuristic algorithm for computing α_{relax} that is tailored to our contact potential formulation and pinball-style discretization.

In our nonlinear solution approach, we assume that the initial guess for the Newton iteration is $\mathbf{y}_0^{n+1} = \mathbf{y}^n$. Such a "same-displacement" predictor is not typically considered ideal for dynamic problems, but it prevents the iteration from beginning in a configuration with extreme or infinite contact forces. Then, in each iteration, we compute the relaxation parameter α_{relax} as follows:

1. Initialize $\alpha_{\text{relax}} \leftarrow 1$ and $F \leftarrow 0$.

2. While $F = 0$,

    (a) Identify tentative $n + 1$ and $n + \alpha_f$ level displacement fields using the current value of

α_{relax} and the formula (33).

(b) If one or both of these tentative configurations leads to an infinite potential (i.e., there are quadrature points closer together than $r_{\min}$), then reduce α_{relax} by half: $\alpha_{\text{relax}} \leftarrow \alpha_{\text{relax}}/2$. Otherwise, set $F = 1$.

(c) If there exist a pair of quadrature points in the tentative $n+1$ level configuration that are closer together than $f_{\text{singular}}r_{\text{in}} + r_{\min}$ and the maximum nodal displacement magnitude $\Delta y_{\max}$ of the increment $\alpha_{\text{relax}}\Delta \mathbf{y}_k$ is greater than $f_{\text{max}\Delta y}r_{\text{out}}$, reduce α_{relax} further:

$$\alpha_{\text{relax}} \leftarrow \min \left\{ \alpha_{\text{relax}}, f_{\text{max}\Delta y}r_{\text{out}}/\Delta y_{\max} \right\} . \tag{34}$$

This step prevents iterates from rapidly escaping the modified Newton method's radius of convergence if the iteration wanders into a region of extremely large contact forces. The dimensionless parameters $f_{\text{singular}} \geq 0$ and $f_{\text{max}\Delta y} > 0$ are user-defined.

In principle, if this step alters α_{relax}, we should re-set $F \leftarrow 0$, to check again for the error state of an infinite potential. However, this possibility is extremely unlikely and does not justify the extra computational cost. We omit this precaution in practice, and have never encountered the error state in computations.

With empirically-selected $f_{\text{max}\Delta y}$ and f_{singular}, this heuristic method of limiting the Newton increment size can make the nonlinear iteration much more robust. However, the iteration is still prone to stagnation or divergence when the time step size is too large. To remedy this problem, we adopt an adaptive scheme for time step selection. Much of the existing literature on adaptive time stepping is focused on improving accuracy. Typical criteria for reducing or increasing time step size are based on comparing the solutions obtained from higher and lower order schemes, to estimate the level of error in the current time step. (See, e.g., [33, Algorithm 1].) However, the application of such a criterion presupposes that the algebraic problem at the current time step is sufficiently tractable to obtain the high- and low-order solutions. For contact problems, this may not be a safe assumption. We follow the alternate philosophy of reducing the time step only when it is necessary to permit progress of the simulation and coarsening if the nonlinear problem is re-solved with few iterations. This is based on the assumption that a prescribed maximum time step provides sufficient accuracy, and adaptivity is applied strictly for solubility purposes. One might also attempt to combine accuracy-based and progress-based criteria, but we have not investigated this possibility in detail. Our adaptive time stepping scheme has the following free parameters:

- $\Delta t_{\max}$: The maximum time step.

- I_{coarsen}: A number of nonlinear solver iterations indicating rapid convergence and suggesting that the time step is unnecessarily small.

17

If the nonlinear solver diverges, $\Delta t \leftarrow \Delta t/2$ and the time step is recomputed. If the nonlinear solver converges in a number of iterations less than or equal to I_{coarsen} and the index i that the time step would have had if the entire computation was performed with step size Δt satisfies $\text{mod}(i, 2) = 0$, then $\Delta t \leftarrow \min\{2\Delta t, \Delta t_{\text{max}}\}$. The constraint that i must be even to coarsen ensures that there always exists a sub-sequence of time steps at times $0, \Delta t_{\text{max}}, 2\Delta t_{\text{max}}, \ldots$, which is convenient for post-processing and visualization purposes.

The performance of the adaptive time stepping scheme hinges critically on the choice of divergence criterion in the nonlinear solver. The simplest choice is to enforce a maximum number of steps. However, this may prove wasteful in comparison with monitoring the nonlinear residual's progress and aborting the iteration if it fails to decrease by a sufficient amount. An overly-eager divergence test, on the other hand, may shrink the time step excessively. To balance these two concerns, we introduce two additional free parameters, to define the divergence test in the Newton iteration:

- $0 < f_{\text{progress}} < 1$: The fraction by which the residual must reduce between iterations.

- $I_{\text{check}} > 1$: The nonlinear iteration index after which we check for divergence.

- $I_{\text{max}} > 1$: The absolute maximum number of iterations allowed.

Divergence of the nonlinear iteration at step k is thus signalled by the condition

$$\left\{ (k > 1) \wedge \left(\frac{\|\mathbf{R}_k\|_{\ell^2}}{\|\mathbf{F}_0^k\|_{\ell^2}} > f_{\text{progress}} \frac{\|\mathbf{R}_{k-1}\|_{\ell^2}}{\|\mathbf{F}_0^{k-1}\|_{\ell^2}} \right) \wedge \left(\|\mathbf{R}_k\|_{\ell^2} > f_{\text{progress}} \|\mathbf{R}_{k-1}\|_{\ell^2} \right) \wedge (k \geq I_{\text{check}}) \right\} \vee (k > I_{\text{max}}) \ . \tag{35}$$

The term $\left(\|\mathbf{R}_k\|_{\ell^2} > f_{\text{progress}} \|\mathbf{R}_{k-1}\|_{\ell^2} \right)$ is only necessary if the reference force vector $\mathbf{F}_0^k$ actually depends on k.

Remark 4. It may sound reasonable to simply set $I_{\text{check}} = 2$ and demand monotonic progress. However, in contact problems, we often observe an increase in the nonlinear residual from $k = 1$ to $k = 2$, as new contacts initiate. If $I_{\text{check}} = 2$, then a pathological pattern of refinement may be triggered, in which contact occurs at $k = 2$, the step is flagged as divergent, the time step is halved, one smaller step converges, then, on the next step, contact occurs at $k = 2$, the step is flagged, and so on, resulting in a demonstration of Zeno's paradox. We therefore recommend selecting $I_{\text{check}} \geq 3$.

3.5. Summary of parameters

We have introduced a number of free parameters in the specification of the potential, ϕ, and our solution algorithm. This section lists these parameters and comments on their qualitative effects.

- Parameters of the potential function:

  - p is dimensionless and determines the strength of the singularity in the potential. To ensure non-penetration, it should be chosen ≥ 4. Excessive singularity may make computations less tractable, and we primarily use $p = 4$.

  - k_c has dimensions of $ML^{p-5}T^{-2}$ and scales the overall stiffness of the potential. Larger values may require smaller time steps.

  - r_{in} is the length scale over which the singular part of the potential acts and $r_{out} - r_{in}$ is the distance over which the polynomial part of the potential acts. Reducing r_{out} toward zero will approach a classical contact constraint, but will also make the contact forces stiffer, requiring smaller time steps to resolve impact events.

  - r_{min} is the radius within which the potential goes to infinity. In problems with high contact pressures, it may be helpful to set this so that quadrature points are spaced more closely than r_{min}.

- Parameters of the solution algorithm:

  - $f_{singular} \geq 0$ is dimensionless and determines what fraction of the singular part of the potential function's domain ($r_{min} < r_{12} < r_{in}$) generates sufficiently-rapidly-varying forces to warrant prophylactic relaxation of the nonlinear iteration. Larger values may slow down the nonlinear iteration, but may also prevent divergence and subsequent time step refinement.

  - $f_{max\Delta y} > 0$ is dimensionless and determines the maximum allowable nodal displacement of a nonlinear solution increment, as a fraction of r_{out}, when relaxation is deemed necessary based on the choice of $f_{singular}$. Smaller values slow down the nonlinear iteration, but may prevent divergence and subsequent time step refinement.

  - Δt_{max} is the maximum allowed time step size and should be selected to ensure sufficient time resolution, as the dynamic time stepping algorithm of Section 3.4 is based only on ease or difficulty of nonlinear solution, not time accuracy of the solution.

  - $I_{coarsen}$ is an integer indicating what number of nonlinear iterations suggest that the time step is unnecessarily small. Smaller values of $I_{coarsen}$ make coarsening less likely, and may lead to overly-conservative time step sizes. However, larger values may lead to false positives, in which the time step is coarsened, the nonlinear solution diverges, and the time step is refined again. Such false positives also result in unnecessary computation.

- $0 < f_{\mathrm{progress}} < 1$ indicates a fraction by which the nonlinear residual must be reduced between nonlinear iterations for the solution process to be viewed as making progress. Selecting this too large may permit long, slowly-converging nonlinear solves in situations where it might be more efficient to simply take more time steps. Setting this too small results in excessive temporal refinement.

- $I_{\mathrm{check}} > 1$ is an integer indicating how many iterations to take before testing for divergence or stagnation with f_{progress}. Its selection is discussed in Remark 4.

- I_{max} is an integer indicating the absolute maximum number of iterations allowed in a nonlinear solution process. Given reasonable choices of f_{progress} and I_{check}, this can be set quite large, or even infinite with little impact on performance.

Determination of more precise, quantitative guidelines for parameter selection is deferred to future studies.

4. Specialization to isogeometric analysis of shell and cable structures

In this section, we fill in the details of a particular case of the generic problem (18) considered earlier. To be able to model atrioventricular valves, we need to devise a coupled problem consisting of shell structure and cable subproblems. We specify this problem by defining

$$
\begin{aligned}
B(\mathbf{y}, \mathbf{z}) = &B^{\mathrm{sh}}\left(\mathbf{y}^{\mathrm{sh}}, \mathbf{z}^{\mathrm{sh}}\right) + B^{\mathrm{ca}}\left(\mathbf{y}^{\mathrm{ca}}, \mathbf{z}^{\mathrm{ca}}\right) \\
&+ \sum_{i=1}^{N_\lambda} \boldsymbol{\lambda}_i \cdot \left(\mathbf{z}^{\mathrm{sh}}\left(\mathbf{X}_i^{\lambda,\mathrm{sh}}\right) - \mathbf{z}^{\mathrm{ca}}\left(\mathbf{X}_i^{\lambda,\mathrm{ca}}\right)\right) \\
&+ \sum_{i=1}^{N_\lambda} \delta\boldsymbol{\lambda}_i \cdot \left(\mathbf{y}^{\mathrm{sh}}\left(\mathbf{X}_i^{\lambda,\mathrm{sh}}\right) - \mathbf{y}^{\mathrm{ca}}\left(\mathbf{X}_i^{\lambda,\mathrm{ca}}\right)\right) ,
\end{aligned}
\tag{36}
$$

where we consider the test and trial functions to be tuples

$$
\mathbf{y} = \{\mathbf{y}^{\mathrm{sh}}, \mathbf{y}^{\mathrm{ca}}, \boldsymbol{\lambda}_1, \ldots \boldsymbol{\lambda}_{N_\lambda}\} \in \mathcal{S}_y^{\mathrm{sh}} \times \mathcal{S}_y^{\mathrm{ca}} \times \mathbb{R}^d \times \ldots \times \mathbb{R}^d = \mathcal{S}_y
\tag{37}
$$

and

$$
\mathbf{z} = \{\mathbf{z}^{\mathrm{sh}}, \mathbf{z}^{\mathrm{ca}}, \delta\boldsymbol{\lambda}_1, \ldots \delta\boldsymbol{\lambda}_{N_\lambda}\} \in \mathcal{V}_y^{\mathrm{sh}} \times \mathcal{V}_y^{\mathrm{ca}} \times \mathbb{R}^d \times \ldots \times \mathbb{R}^d = \mathcal{V}_y .
\tag{38}
$$

We likewise define

$$
F(\mathbf{z}) = F^{\mathrm{sh}}\left(\mathbf{z}^{\mathrm{sh}}\right) + F^{\mathrm{ca}}\left(\mathbf{z}^{\mathrm{ca}}\right) .
\tag{39}
$$

In the above notation, symbols superscripted with "sh" are associated with the shell structure subproblem and symbols superscripted "ca" are associated with the cable subproblem. (Point evaluations of "$\mathbf{y}$" and "$\mathbf{z}$" in the contact energy formulation above are understood to employ whichever

solution component is relevant.) Each vector $\boldsymbol{\lambda}_i$ is a Lagrange multiplier to enforce matching displacements of the material point $\mathbf{X}_i^{\lambda,\text{sh}}$ on the shell structure and point $\mathbf{X}_i^{\lambda,\text{ca}}$ on the cable structure, the latter typically being an endpoint of an individual cable. We formulate the constraint in this way to more naturally accommodate non-conforming discretizations, in which the displacement of the shell structure at $\mathbf{X}_i^{\lambda,\text{sh}}$ is not controlled by a single vector-valued unknown (e.g., the displacement of a "node" in traditional finite element analysis, which could simply be shared between the cable and shell subproblems in a conforming discretization). For computational convenience, we approximate these Lagrange multipliers using penalty forces, i.e.,

$$\boldsymbol{\lambda}_i \approx \beta^{\text{sh–ca}} \left(\mathbf{y}^{\text{sh}} \left(\mathbf{X}_i^{\lambda,\text{sh}} \right) - \mathbf{y}^{\text{ca}} \left(\mathbf{X}_i^{\lambda,\text{ca}} \right) \right) , \tag{40}$$

where $\beta^{\text{sh–ca}} > 0$ is a penalty parameter. We thereby replace B with

$$\begin{aligned} B_{\text{pen}}(\mathbf{y}, \mathbf{z}) = & B^{\text{sh}} \left(\mathbf{y}^{\text{sh}}, \mathbf{z}^{\text{sh}} \right) + B^{\text{ca}} \left(\mathbf{y}^{\text{ca}}, \mathbf{z}^{\text{ca}} \right) \\ & + \sum_{i=1}^{N_\lambda} \beta^{\text{sh–ca}} \left(\mathbf{y}^{\text{sh}} \left(\mathbf{X}_i^{\lambda,\text{sh}} \right) - \mathbf{y}^{\text{ca}} \left(\mathbf{X}_i^{\lambda,\text{ca}} \right) \right) \cdot \left(\mathbf{z}^{\text{sh}} \left(\mathbf{X}_i^{\lambda,\text{sh}} \right) - \mathbf{z}^{\text{ca}} \left(\mathbf{X}_i^{\lambda,\text{ca}} \right) \right) . \end{aligned} \tag{41}$$

In general, one might estimate appropriate values for $\beta^{\text{sh–ca}}$ from problem data and mesh element size, using dimensional analysis. In this work, we simply select an effective value of $\beta^{\text{sh–ca}}$ based on numerical experiments, as we are primarily focused on a specific class of problems.

4.1. Shell structure subproblem

The shell subproblem forms B^{sh} and F^{sh} are defined and discretized as in [34]. Breifly, the shell structure is modeled using Kirchhoff–Love thin shell assumptions, with an arbitrary hyperelastic constitutive model. The thin shell problem in terms of displacement demands at least H^2 regularity of the test and trial spaces; to satisfy this requirement in a standard Bubnov–Galerkin discretization, we use smooth spline function spaces to represent the geometry and displacement in an isogeometric fashion. Smooth isogeometric surface discretizations are also known to improve performance in contact problems [35–37]. In particular, we solve for a displacement $\mathbf{y}$ of the current midsurface configuration Γ_t from its reference configuration Γ_0 and let

$$B^{\text{sh}} \left(\mathbf{w}, \mathbf{y}^{\text{sh}} \right) = \int_{\Gamma_0} \mathbf{w} \cdot h_{\text{th}} \rho_0 \ddot{\mathbf{y}}^{\text{sh}} \, d\Gamma + \int_{\Gamma_0} \int_{-h_{\text{th}}/2}^{h_{\text{th}}/2} \delta \mathbf{E} : \mathbf{S} \, d\xi^3 d\Gamma + \int_{\Gamma_0} \mathbf{w} \cdot C^{\text{sh}} h_{\text{th}} \rho_0 \dot{\mathbf{y}}^{\text{sh}} \, d\Gamma \tag{42}$$

and

$$F^{\text{sh}} (\mathbf{w}) = \int_{\Gamma_0} \mathbf{w} \cdot \rho_0 h_{\text{th}} \mathbf{f} \, d\Gamma + \int_{\Gamma_t} \mathbf{w} \cdot \mathbf{h}^{\text{net}} \, d\Gamma , \tag{43}$$

where ρ_0 is the mass density in the reference configuration, h_{th} is the shell thickness, $\mathbf{S}$ is the second Piola–Kirchhoff stress, $\mathbf{E}$ is the Green–Lagrange strain, $\delta\mathbf{E}$ is its functional derivative in the direction $\mathbf{w}$, C^{sh} is a mass damping coefficient, $\mathbf{f}$ is the prescribed body force, and $\mathbf{h}^{\text{net}}$ is the sum of the tractions prescribed on the two sides of the shell structure. The Green–Lagrange strain is determined from $\mathbf{y}$ according to the kinematic assumptions of Kirchhoff–Love shell theory, as detailed in [34]. In the examples of this paper, we determine $\mathbf{S}$ using a variety of hyperelastic constitutive models, which we specify in the sequel.

4.2. Cable subproblem

Our definitions of the forms B^{ca} and F^{ca} are based on the isogeometric bending-stabilized cable formulation of Raknes et al. [38]. The cited work uses kinematic assumptions to express the mechanics of the cable in terms of a 1D middle curve, parameterized by coordinate ξ^1. Let points in the reference configuration S_0 of this curve be parameterized $\mathbf{X}\left(\xi^1\right)$ and points in the deformed configuration S_t be parameterized $\mathbf{x}\left(\xi^1\right)$, such that $\mathbf{x}\left(\xi^1\right) = \phi\left(\mathbf{X}\left(\xi^1\right)\right)$, where $\phi : S_0 \to S_t$ is the motion of the deforming curve. The displacement $\mathbf{y}^{\text{ca}} : S_0 \to \mathbb{R}^d$ is then given by

$$\mathbf{y}^{\text{ca}}\left(\mathbf{X}\left(\xi^1\right)\right) = \mathbf{x}\left(\xi^1\right) - \mathbf{X}\left(\xi^1\right) \ . \tag{44}$$

Covariant and contravariant basis vectors for the reference configuration are

$$\mathbf{G}_1 = \frac{\partial \mathbf{X}}{\partial \xi^1} \qquad \text{and} \qquad \mathbf{G}^1 = \frac{\mathbf{G}_1}{\|\mathbf{G}_1\|_{\ell^2}} \ . \tag{45}$$

Because of the small thickness of the cables in our target applications, we omit bending terms present in formulation of [38] and define

$$B^{\text{ca}}(\mathbf{w}, \mathbf{y}^{\text{ca}}) = \int_{S_0} \mathbf{w} \cdot A_0 \rho_0 \ddot{\mathbf{y}}^{\text{ca}} \, dS + \int_{S_0} \delta\varepsilon \, E^{\text{ca}} A_0 \|\mathbf{G}^1\|_{\ell^2}^4 \varepsilon \, dS + \int_{S_0} \mathbf{w} \cdot C^{\text{ca}} A_0 \rho_0 \dot{\mathbf{y}}^{\text{ca}} \, dS \tag{46}$$

and

$$F^{\text{ca}}(\mathbf{w}) = \int_{S_0} \mathbf{w} \cdot A_0 \rho_0 \mathbf{f}_0 \, dS - \int_{S_t} \mathbf{w} \cdot \mathbf{h} \, dS \ , \tag{47}$$

where

$$\varepsilon = \frac{1}{2} \left(\frac{\partial \mathbf{x}}{\partial \xi^1} \cdot \frac{\partial \mathbf{x}}{\partial \xi^1} - \frac{\partial \mathbf{X}}{\partial \xi^1} \cdot \frac{\partial \mathbf{X}}{\partial \xi^1} \right) \tag{48}$$

is the extensional strain,

$$\delta\varepsilon = \frac{\partial \mathbf{w}}{\partial \xi^1} \cdot \frac{\partial \mathbf{x}}{\partial \xi^1} \tag{49}$$

is its functional derivative in the direction $\mathbf{w}$, ρ_0 is the mass density (per unit volume) in the reference configuration, $\mathbf{f}_0$ is a body force in the reference configuration, $\mathbf{h}$ is a traction (per unit length)

22

in the current configuration, A_0 is the cable's cross-sectional area in the reference configuration, the Young's modulus, E^{ca}, determines the tensile stiffness of the cable, and the mass damping coefficient C^{ca} provides dissipation. The use of a Young's modulus to determine stiffness is based on the assumption of a St. Venant–Kirchhoff material model. The justification for this assumption in the context of modeling chordae tendineae is discussed in Section 5.2.2.

4.3. Quadrature of contact energy

For all computations in this paper, we numerically-integrate the contact force form $D_{\mathbf{z}} E_{\text{c}}(\mathbf{y}^{n+\alpha_f})$ in (18) with the same collection of quadrature points used to integrate the weak forms of the shell and cable subproblems. This choice is not necessary, and, in cases where shell and cable elements are substantially larger than the length scale parameter r_{out} in the contact potential formulation, it may be a poor choice, as it would lead to either unnecessary over-integration of structural mechanics, or inaccurate under-integration of contact forces.

4.4. Numerical testing

We now use several benchmark problems to demonstrate the properties of the proposed contact method. For the examples in this section, we model the shell structure as a St. Venant–Kirchhoff material, i.e., $\mathbf{S} = \mathbb{C}\mathbf{E}$, where $\mathbb{C}$ is the fourth-order isotropic elasticity tensor determined by Young's modulus E and Poisson's ratio ν.

4.4.1. Sphere and plate subjected to gravitational force

The first test problem involves two shell structures: a sphere (of radius 0.125) and a thin circular plate (of radius 0.35), initially positioned as shown in Figure 4. A zero-displacement boundary condition is applied to the edge of the circular plate, and both the sphere and plate are subject to gravitational force and damping until a static equilibrium configuration is reached. The material parameters of both shell structures are $E = 2.1 \times 10^5$ and $\nu = 0.3$. The plate has density $\rho_0 = 1$ and thickness $h_{\text{th}} = 0.003$. The sphere is made heavier and more resistant to bending by setting $\rho_0 = 10$ and $h_{\text{th}} = 0.03$. Gravitational acceleration is set to 980.665. The parameters used to define the contact potential are $p = 4$, $k_{\text{c}} = 1.0$, $r_{\text{in}}/(r_{\text{out}} - r_{\text{min}}) = 0.5$, $r_{\text{min}}/r_{\text{out}} = 0.5$, and $r_{\text{out}} = 0.01$.

The midsurfaces of these shells can be represented exactly using bivariate cubic T-splines [39], allowing an isogeometric discretization. Three T-spline meshes, denoted as M0, M1 and M2, are used in this study. M1 is a global h-refinement of M0, and M2 is an h-refinement of M1. M0, M1 and M2 contain a total of 690, 1565 and 5180 Bézier elements, respectively. The Bézier elements of M1 is shown in Figure 4. The T-spline meshes are generated by the Autodesk T-Splines Plug-in [40] for Rhinoceros [41]. Problems are considered to have reached a static equilibrium configuration when the ℓ^∞ norm of the vector of changes in x_3-direction displacement unknowns between two time steps is smaller than 10^{-8}. A representative numerical solution is shown in
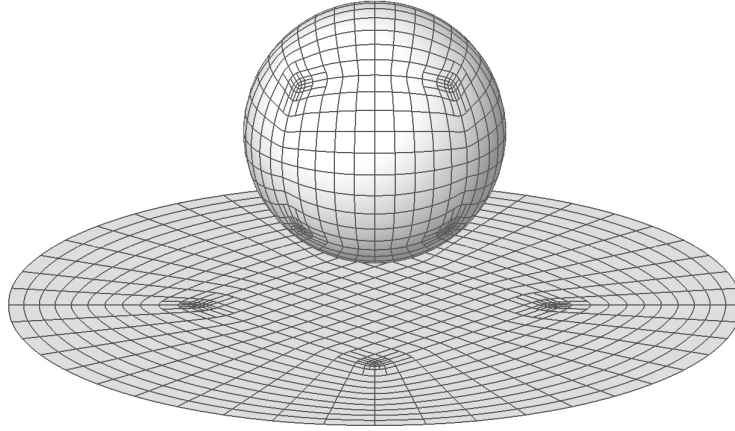
Figure 4: The Bézier elements of the T-spline sphere–plate model. (The sphere is a hollow shell structure.)
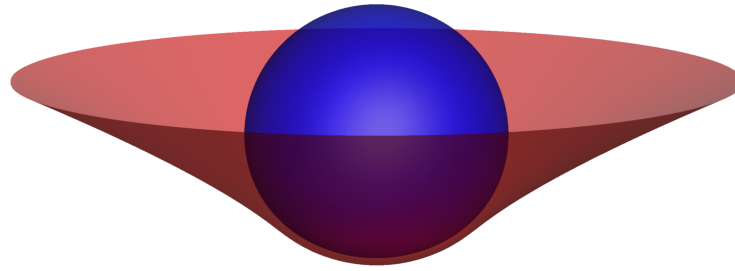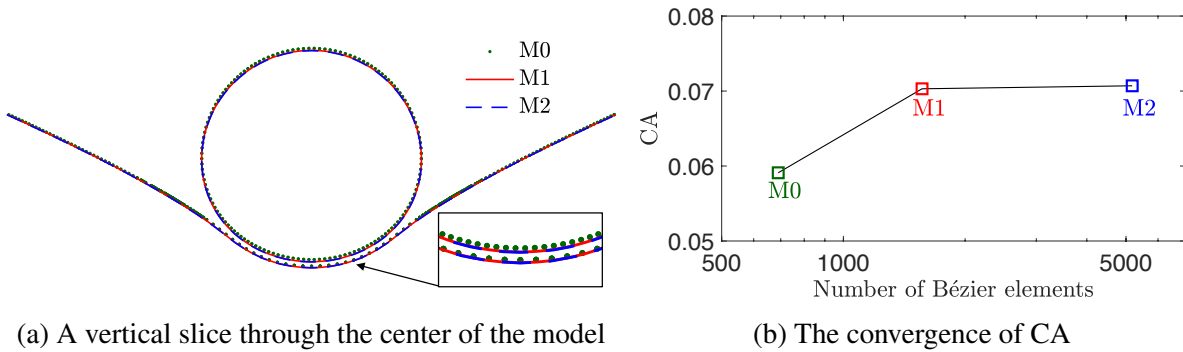

Figure 5: The steady state of the sphere–plate interaction subject to gravitational force.

(a) A vertical slice through the center of the model

(b) The convergence of CA

Figure 6: Results of mesh independence study of the sphere–plate interaction subject to gravitational force.

Figure 5. Figure 6a shows the vertical slice cutting through the center of the model, qualitatively demonstrating convergence of the displacement fields.

To show that quantities of interest associated with contact will converge with mesh refinement, we look at the convergence of the area of the contact region. In our target application of heart valve analysis, area of the contact region, or, the coaptation area (CA), is an important quantitative measure of how effective a valve is at blocking flow [9]. We calculate the CA by integrating the

surface area over all Gauss points with nonzero contact forces.[5] Figure 6b shows CA as a function of the number of Bézier elements in the model. The relative error in CA satisfies the convergence criterion $|CA_1 - CA_2| < 0.01\,CA_2$, where CA_i is the CA on mesh Mi.

4.4.2. Sphere–plate impact

It is clear that a spatial semi-discretization of the problem (18) will conserve energy exactly (assuming that B and F derive from potentials, e.g., B from hyperelasticity and F from gravity). To demonstrate that this property can be recovered to a high degree of accuracy in fully discrete solutions, we use a benchmark proposed by Cirak and West [42] and later studied by Vouga et al. [43]. To study energy conservation, we integrate in time with the implicit midpoint rule, which is symplectic [44, Theorem VI.3.5], and can be obtained as a special case of generalized-α integration, with $\rho_\infty = 1$ (i.e., no high-frequency dissipation).

The problem specification is reproduced from [42]. The problem uses the same T-spline mesh M1 as shown in Figure 4, but the sphere and the plate are both free to move. The sphere and the plate initially travel toward one another with a relative velocity of 100. Both objects are modeled as Kirchhoff–Love thin shells of thickness $h_{\text{th}} = 0.0035$. The initial mass density of the shell structures is $\rho_0 = 0.0785$. The constitutive parameters of the sphere and plate are $E = 2.1 \times 10^5$ and $\nu = 0.3$, as in Section 4.4.1. The parameters used to define the contact potential are $p = 4$, $k_{\text{c}} = 10.0$, $r_{\text{in}}/(r_{\text{out}} - r_{\text{min}}) = 0.7$, $r_{\text{min}}/r_{\text{out}} = 0.5$, and $r_{\text{out}} = 0.003$.

The maximum time step size is set to $\Delta t_{\text{max}} = 10^{-6}$. Adaptive time stepping is never triggered for this problem using that time step, even for very small nonlinear convergence tolerances, so $\Delta t = \Delta t_{\text{max}}$ in all steps. Several snapshots of the displacement solution are rendered in Figure 7. The kinetic, elastic, contact, and total energies are plotted as functions of time in Figure 8. It is clear that total energy is conserved well. The small proportion of total energy stored as contact potential energy is consistent with the contact potential approximating a workless constraint.

Remark 5. This benchmark is quite mild from the perspective of nonlinear solution, and the problem can be converged with even larger time steps. However, using too large a time step results in significant energy errors. Energy errors can be induced at the current time step by shrinking the contact length scale r_{out} or increasing the penalty scaling parameter k_{c} (but conservation is still, of course, recovered under refinement in time). Even when using a symplectic integrator, one must use sufficient resolution in time to obtain the oft-cited long-time (near) energy conservation of such methods.[6] A way to consistently achieve conservation without tuning of the time step or

---

[5]Tacit in this definition of CA is the assumption that r_{out} is small relative to the overall length scale of the structure geometry. If r_{out} were large relative to the structure, the entire surface of the structure would carry contact forces, and convergence of CA would be trivial.

[6]It is possible to conserve energy exactly [44, Section IV.4], but this can actually disrupt the good qualitative
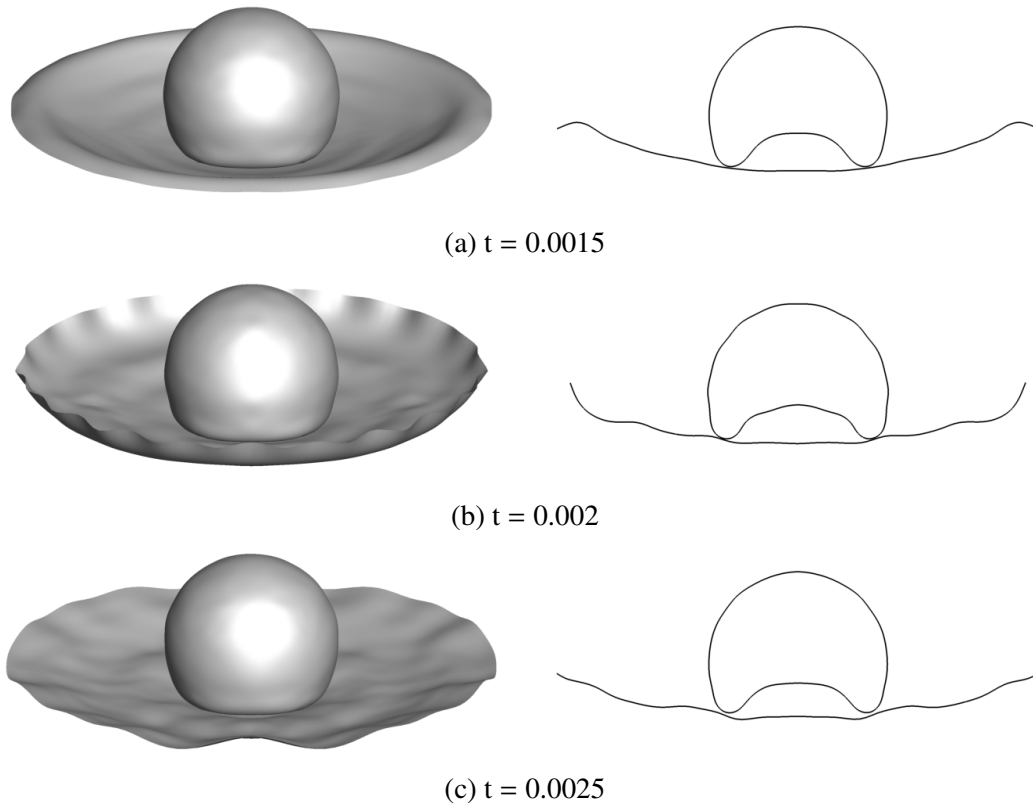
(a) t = 0.0015

(b) t = 0.002

(c) t = 0.0025

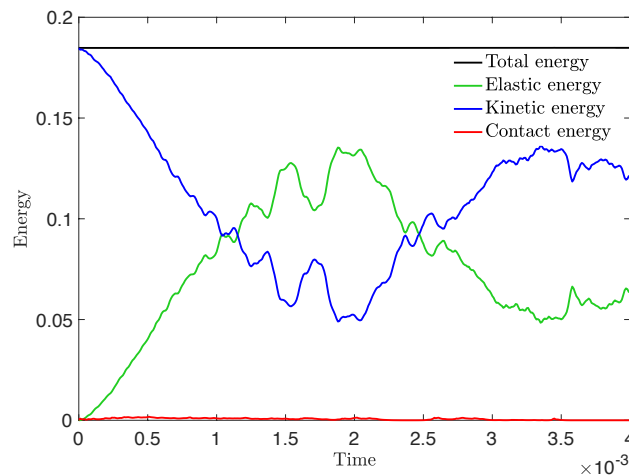Figure 7: Snapshots of the sphere–plate impact problem at several time frames.

Figure 8: The history of the kinetic, elastic, contact, and total energies for the sphere–plate impact problem.

parameters might be to introduce an additional accuracy-based criterion for time step adaptation, as discussed in Section 3.4, but we have not investigated that possibility thoroughly.

---

behavior of symplectic integration that is frequently (and perhaps misleadingly) attributed to good long-time energy conservation [44, Example IV.4.4].
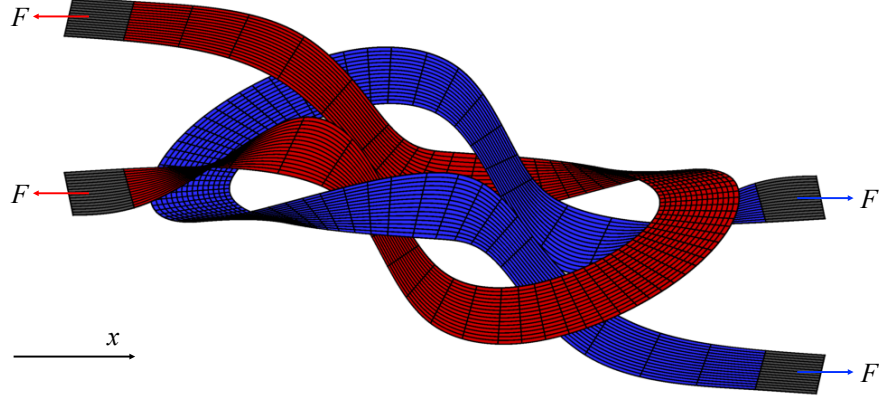
Figure 9: The reef knot problem setup. Black lines indicate the Bézier elements of the reef knot B-spline surface. This is the stress-free configuration of the ribbons.

4.4.3. Reef knot

To demonstrate the ability of the proposed approach to handle two-sided self-contact of thin structures with free edges, we introduce a benchmark inspired by the similar reef knot example of [43, Section 8.4]. Two thin 58 cm × 2 cm rectangular ribbons, which consist of 1980 cubic B-spline elements, are positioned as a loosened reef knot. We constrain the ends of the ribbons to move only in the x-direction. To tighten the knot, we prescribe tractions of $\mathbf{h}^{net} = \pm(10^4 \text{ dyn/cm}^2)\mathbf{e}_1$ on the regions colored gray in Figure 9. The area of each gray region is 3.2395 cm^2, yielding a force (F) with magnitude of 3.2395×10^4 dyn acting on each end of the ribbon. The Young's modulus of the ribbons is $E = 1 \times 10^8$ dyn/cm^2, the Poisson's ratio is $\nu = 0.4$, the thickness is 0.004 cm, and the mass density is 1 g/cm^3. The parameters defining the contact potential are $p = 4$, $k_c = 1.0 \times 10^7$ cm^{p-5}s^{-2}g, $r_{in}/(r_{out} - r_{min}) = 0.7$, $r_{min}/r_{out} = 0.7$, and $r_{out} = 0.1$ cm. A damping traction of $-(10 \text{ cm}^{-2}\text{s}^{-1}\text{g})\dot{\mathbf{y}}$ is applied to the shell structure midsurface throughout the analysis.

Figure 10 shows two snapshot of the reef knot tying simulation at $t = 0.08$ s and $t = 0.4$ s. Intersections of the ribbon midsurfaces with a planar slice illustrate the two-sided self-contact of thin structures with free edges. Adaptive time step size over the history of the simulation is shown in Figure 11. It clearly shows that the adaptive subdivision of time step size is triggered when the contact comes into effect. More nonlinear iterations are needed to address the nonlinearity introduced by the numerical contact. As the simulation approaches steady state, the time step size are coarsen until the Δt_{max} is recovered.

5. Application to isogeometric analysis of atrioventricular valve dynamics

We now demonstrate the proposed contact methodology, and its integration with isogeometric shell and cable formulations, in the context of a specific application: the structural dynamics of

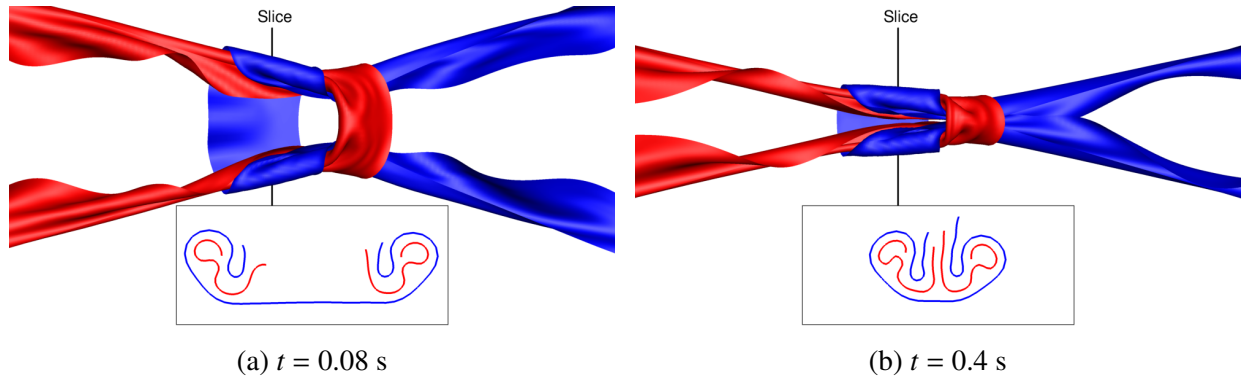(a) $t = 0.08$ s   (b) $t = 0.4$ s
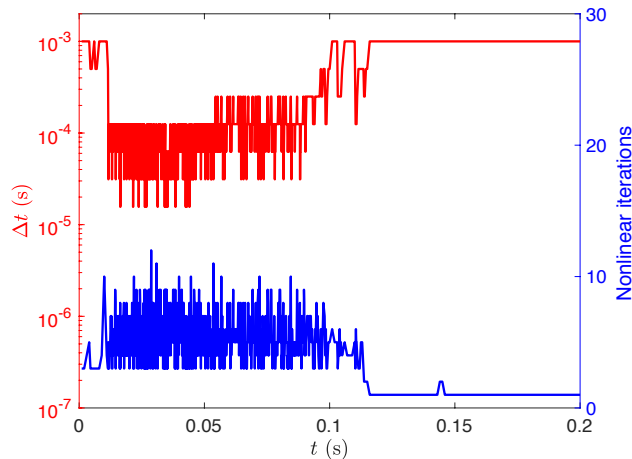
Figure 10: Results with cross-sectional slices.

Figure 11: Time step size and number of nonlinear iterations at each time step of the knot-tying simulation.

the right atrioventricular valve, also referred to as the tricuspid valve. To focus on issues related to contact, we consider a pure structural mechanics problem, in which the surrounding blood flow is modeled by a pressure load on the valve leaflets and mass-proportional damping.

5.1. Geometrical modeling

We construct an idealized tricuspid valve geometry based on a combination of the model shown in [45] and segmented micro-CT images of porcine tricuspid valves from C.-H. Lee's lab. The leaflets and chordae are described geometrically by cubic B-spline surfaces and curves. A number of studies have previously found such spline surfaces and curves to be convenient representations of atrioventricular leaflet and chordae geometries extracted from imaging data [14, 46, 47]. The Bézier element meshes of these spline structures are shown in Figure 12. The cross-sectional area of the chordae is about 0.00171 cm^2 [45, 48], yielding a chordae radius of 0.023 cm. The leaflet thickness is 0.0396 cm [45, Table 1].

To demonstrate the capability of the proposed contact methods, we also consider a modification of the geometry described above. We model the atrioventricular valve pathology of prolapse,
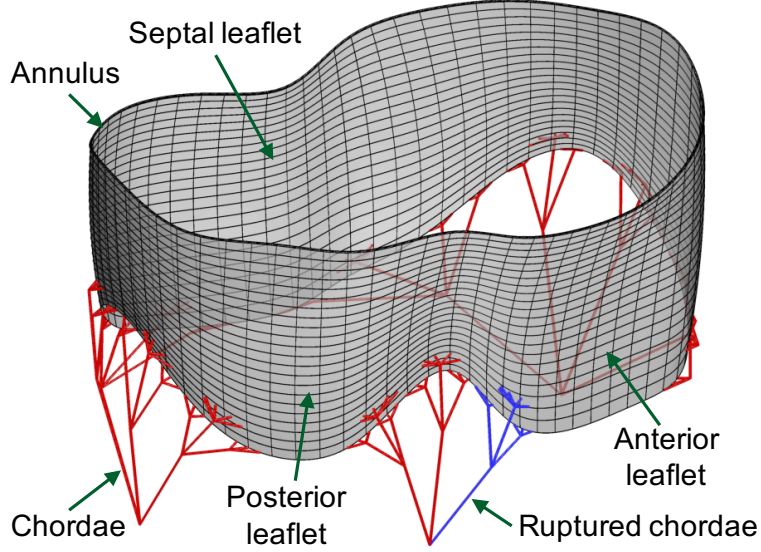
28

Figure 12: The B-spline surface and elements of the tricuspid valve leaflet and the B-spline chordae.

by severing the chordae highlighted in blue in Figure 12. This pathology is common in the left atrioventricular valve [49], but may also occur in the right side of the heart [50]. Prolapse of atrioventricular valves tends to cause regurgitation (i.e., flow of blood back into the atrium) during systole [1].

5.2. Material modeling

Accurate modeling of tricuspid valve structural mechanics requires careful attention to the constitutive modeling of the leaflet tissue. The extreme nonlinearity of a realistic soft tissue model also serves to test the robustness of the solution algorithm developed in Section 3.4.

5.2.1. Leaflets

To model the leaflet tissue, we assume that the material is incompressible and the second Piola–Kirchhoff stress, $\mathbf{S}$, is computed from the Green–Lagrange strain, $\mathbf{E}$, using an isotropic Fung-type material model, in which the extracellular matrix and fiber stiffness contributions are modeled by neo-Hookean and exponential terms, respectively. Specifically,

$$\mathbf{S} = 2\frac{\partial \psi_{el}}{\partial \mathbf{C}} - p\mathbf{C}^{-1} \,, \tag{50}$$

where

$$\psi_{el} = \frac{c_0}{2}\left(I_1 - 3\right) + \frac{c_1}{2}\left(e^{c_2(I_1-3)^2} - 1\right) \,, \tag{51}$$

and

$$\frac{\partial \psi_{el}}{\partial \mathbf{C}} = \frac{1}{2}\left(c_0 + 2c_1c_2(I_1 - 3)e^{c_2(I_1-3)^2}\right)\mathbf{I} \,. \tag{52}$$

In the above, $\mathbf{C} = 2\mathbf{E} + \mathbf{I}$ is the right Cauchy–Green tensor, p is a Lagrange multiplier enforcing incompressibility[7], and c_0, c_1, and c_2 are material parameters. For the tricuspid valve, we use $c_0 = 10$ kPa, $c_1 = 0.209$ kPa, and $c_2 = 9.046$. (c_1 and c_2 are obtained from [45, Table 2].) The leaflet mass density is 1 g/cm^3.

Remark 6. The exponential term in (51) leads to nonlinearities that are especially difficult to resolve. We improve the robustness of our nonlinear solution procedure by assembling the tangent operator $D_{(\cdot)}R(\mathbf{y}_k^{n+1}, \mathbf{N}_A)$ (cf. Section 3.4) with a larger value of c_0. This requires more iterations, but does not change the converged solution and makes the iterative procedure more robust. In the simulation of valve dynamics, we increase the c_0 of the tangent operator to 200 kPa.

Remark 7. Native valve tissue is typically anisotropic, due to collagen fibers being oriented in a preferred direction. Material anisotropy is easily handled by Kirchhoff–Love shell formulations [51–53], but obtaining collagen fiber orientation data, incorporating it into an adequate constitutive model, and calibrating the parameters of that model are challenging research topics outside the scope of the present work, which is primarily focused on contact mechanics.

5.2.2. Chordae

As mentioned in Section 4.2, we restrict the cable material to a simple St. Venant–Kirchhoff model. The St. Venant–Kirchhoff model is typically a poor choice for the stress–strain relation of biological soft tissue about its true reference configuration. However, it can be justified in the case of modeling the tensile response of chordae tendineae, if the reference configuration of the cable is understood to have already been stretched through the chordae's soft "pre-transition" regime, in which most of the collagen fibers remain slackened and contribute no tensile stiffness. The Young's modulus of the cable model is then selected to approximate the stiffer "post-transition" regime, in which most of the collagen fibers are recruited. Pre- and post-transition stiffness values for porcine mitral chordae were calculated by fitting *ex vivo* experimental data in [54, Table 1]. The cited work found that the pre-transition stiffness is several orders of magnitude smaller than the post-transition stiffness. Based on this data, we follow [55] and [56] in neglecting pre-transition stiffness and considering the stress-free reference configuration of the chordae to include tensile strain up to the transition point. For our model of the tricuspid valve, we select a moderately higher effective post-transition stiffness of $E^{\mathrm{ca}} = 4 \times 10^8$ dyn/cm^2, in recognition of the fact that tricuspid chordae are typically stiffer [48] than the mitral chordae considered in the cited studies.

---

[7]For shell analysis, one can use the plane stress condition, $S^{33} = 0$, to analytically determine the Lagrangian multiplier p (see Kiendl et al. [34, Section 5.1] for details).

5.3. Boundary conditions

Contact in the atrioventricular valves is primarily a concern during systole, when these valves are closed. We simulate valve closure by applying a pressure follower load to the ventricular side of the leaflets. The pressure load as a function of time is

$$p = (25 \text{ mmHg}) \begin{cases} t/T & t < T \\ 1 & \text{otherwise} \end{cases}, \tag{53}$$

where the ramp-up time scale T is set to $T = 0.005$ s. The magnitude after time T is representative of normal tricuspid valve pressure gradients during systole [57]. Loss of energy to the surrounding blood is modeled with mass damping, in both the leaflets and the chordae. The damping coefficients C^{sh} and C^{ca} are both 2000 s^{-1}. The shell structure subproblem is subject to a pinned boundary condition at the annulus. The connections of the chordae to the papillary muscles are subject to strongly-enforced homogeneous Dirichlet boundary conditions on displacement.

5.4. Computational setup

The B-spline surface and curves defining the leaflet and chordal geometry are directly used as an isogeometric computational model. The leaflets consist of 1540 cubic B-spline elements and the chordae contain 373 cubic B-spline elements. The pinned boundary condition at the valve annulus is enforced strongly in the discrete model by fixing the first row of control points. Adaptive time stepping is used with a maximum time step size of $\Delta t_{\max} = 10^{-4}$ s. The parameters defining the contact potential are $p = 4$, $k_{\text{c}} = 1.0 \times 10^6$ cm^{p-5}s^{-2}g, $r_{\text{in}}/(r_{\text{out}} - r_{\min}) = 0.7$, $r_{\min}/r_{\text{out}} = 0.2$, and $r_{\text{out}} = 0.05$ cm.

5.5. Results

The closing behavior of the model with intact chordae is illustrated through a series of snapshots in Figure 13. The complex folding of the valve leaflets is clearly outside the scope of geometries like [4, Figure 20] for which our earlier contact approach was designed, but the potential-based method remains effective. While the closed configuration exhibits creases separating three general "leaflets" (to which the tricuspid valve owes its name), the divisions between these parts of the shell structure are not clear-cut, and self-contact is essential. Figure 14 visualizes the pinballs (Gauss points) on the chordae that are involved in contact. The shell–cable contact and the cable–cable contact are clearly shown. The steady-state configuration shown in Figure 13f is compared, in Figures 15 and 16, with the steady-state configuration resulting from removal of chordae tethering the anterior leaflet. The valve with ruptured chordae clearly exhibits prolapse of the anterior leaflet. The large, visible gap between the leaflets in the prolapsed case clearly illustrates that this pathological condition is associated with regurgitation.

31

(a) $t = 0.005$ s   (b) $t = 0.007$ s   (c) $t = 0.01$ s

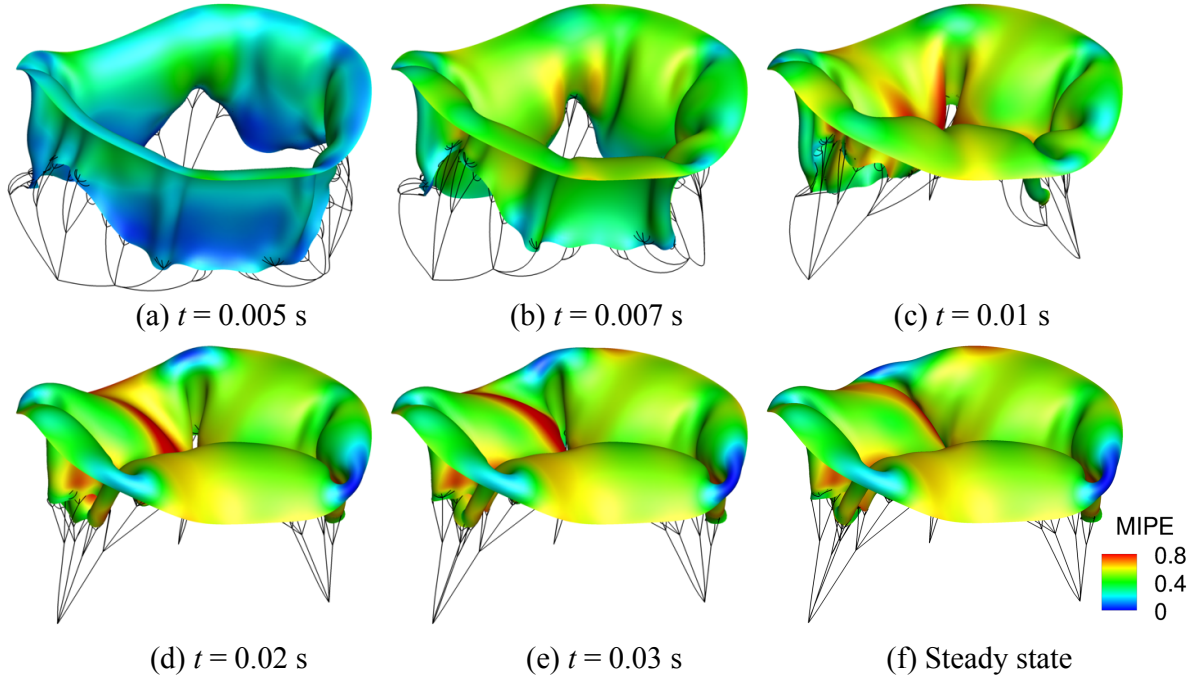(d) $t = 0.02$ s   (e) $t = 0.03$ s   (f) Steady state

Figure 13: A series of snapshots illustrating the closing behavior of the tricuspid valve with intact chordae. Color indicates maximum in-plane eigenvalues of $\mathbf{E}$ (MIPE) on the atrial side of the leaflets.
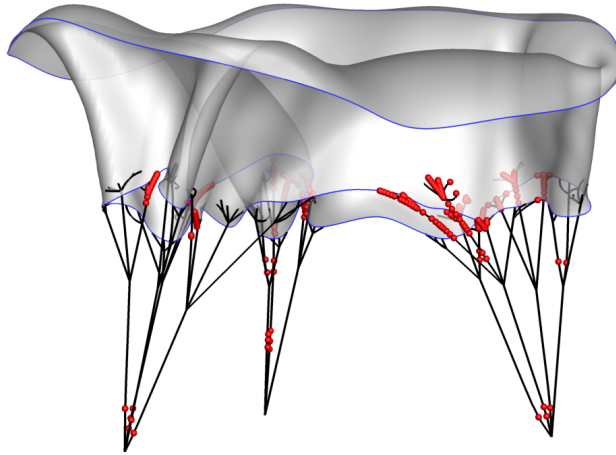

Figure 14: Cable pinballs with nonzero contact forces are highlighted in red, illustrating shell–cable and cable–cable contact. (Recall that pinballs near one another in the reference configuration do not interact.)

Remark 8. The sharp curvature near the free edge of the prolapsed anterior leaflet (visible in Figures 15 and 16) is likely an artifact of our use of a pressure follower load on the ventricular side of the leaflet. This potentially-unphysical solution feature highlights the importance of incorporating fluid–structure interaction into atrioventricular valve prolapse models, as we plan to do in future studies.
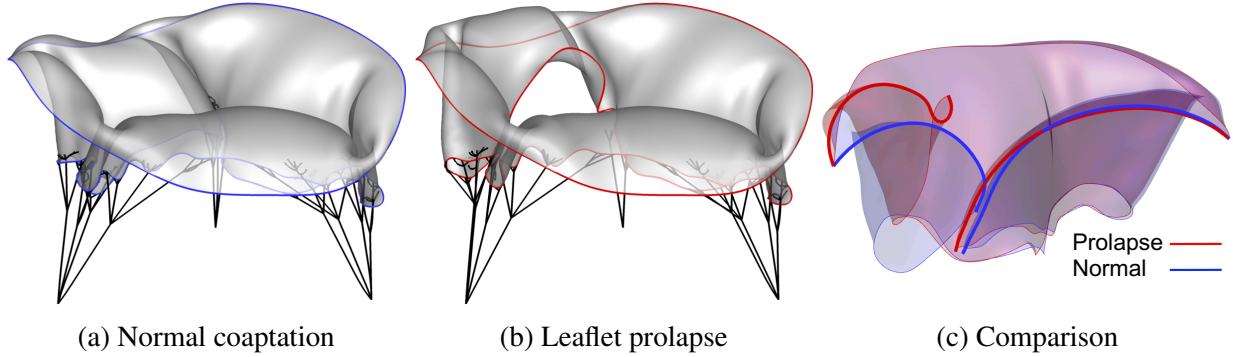
(a) Normal coaptation   (b) Leaflet prolapse   (c) Comparison

Figure 15: Comparison of coaptation in normal and prolapsing valves.

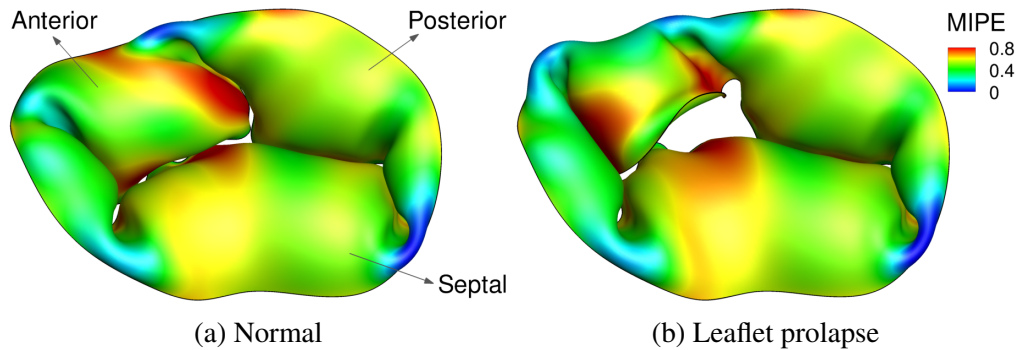(a) Normal   (b) Leaflet prolapse

Figure 16: Comparison of MIPE contours in normal and prolapsing valves.

6. Conclusion

The contact procedures that we developed in [4, Section 5.2] for isogeometric and immerso-geometric analysis of aortic valves are not suitable for analysis of the more complicated atrioventricular valves. To develop a more versatile formulation, we build on the ideas outlined in [15] and find that, for a particular choice of quadrature and a potential function of limited range, one obtains a discrete scheme with some similarities to the pinball algorithm [20]. Singularities in our new formulation require specialized nonlinear solution and time stepping procedures for reliable convergence. The main limitation of the present study is the need for an *ad hoc* selection of parameters defining both the contact potential and the solution algorithm. In future extensions, we plan to develop a more principled procedure for selecting parameters. This will reduce computational cost, by avoiding unnecessary refinement in time and/or damping of the nonlinear iteration.

This will enable future immersogeometric fluid–structure interaction (FSI) simulations of atrioventricular valves using the technologies described in this paper for the structural subproblem. As mentioned in Remark 2, we consider the contact methodology discussed in this paper especially well suited for use in conjunction with immersed numerical approaches. However, additional research is required to determine how best to combine the nonlinear solution and adaptive time stepping algorithms described in Section 3.4 with existing FSI solution procedures.

The versatile formulation developed in this paper could also be applied to other challenging problems. The demonstrated effectiveness of our formulation for shell-against-shell contact looks promising in the context of studying damage resulting from impacts against thin composite structures, which some of us have recently become interested in [58]. The formulation's compatibility with point cloud geometry descriptions also leads us to consider applying it to meshfree discretizations of structures fragmenting after blast loading [59, 60], but the use of an unbounded potential function in explicit computations may lead to difficulties, warranting modifications to the present framework, such as reverting to a non-singular force–separation law.

Acknowledgments

References

[1] E. Hayek, C. N. Gring, and B. P. Griffin. Mitral valve prolapse. *The Lancet*, 365(9458): 507–518, 2005.

[2] K. B. Chandran. Role of computational simulations in heart valve dynamics and design of valvular prostheses. *Cardiovascular Engineering and Technology*, 1(1):18–38, 2010.

[3] A. Kheradvar, E. M. Groves, A. Falahatpisheh, M. K. Mofrad, S. Hamed Alavi, R. Tranquillo, L. P. Dasi, C. A. Simmons, K. J. Grande-Allen, C. J. Goergen, F. Baaijens, S. H. Little, S. Canic, and B. Griffith. Emerging trends in heart valve engineering: Part IV. Computational modeling and experimental studies. *Annals of Biomedical Engineering*, 43(10):2314–2333, 2015.

[4] D. Kamensky, M.-C. Hsu, D. Schillinger, J. A. Evans, A. Aggarwal, Y. Bazilevs, M. S. Sacks, and T. J. R. Hughes. An immersogeometric variational framework for fluid–structure interaction: Application to bioprosthetic heart valves. *Computer Methods in Applied Mechanics and Engineering*, 284:1005–1053, 2015.

[5] M.-C. Hsu, D. Kamensky, Y. Bazilevs, M. S. Sacks, and T. J. R. Hughes. Fluid–structure interaction analysis of bioprosthetic heart valves: significance of arterial wall deformation. *Computational Mechanics*, 54:1055–1071, 2014.

[6] M.-C. Hsu, D. Kamensky, F. Xu, J. Kiendl, C. Wang, M. C. H. Wu, J. Mineroff, A. Reali, Y. Bazilevs, and M. S. Sacks. Dynamic and fluid–structure interaction simulations of bioprosthetic heart valves using parametric design with T-splines and Fung-type material models. *Computational Mechanics*, 55:1211–1225, 2015.

[7] D. Kamensky, M.-C. Hsu, Y. Yu, J. A. Evans, M. S. Sacks, and T. J. R. Hughes. Immersogeometric cardiovascular fluid–structure interaction analysis with divergence-conforming B-splines. *Computer Methods in Applied Mechanics and Engineering*, 314:408–472, 2017.

[8] A. Aggarwal and M. S. Sacks. An inverse modeling approach for semilunar heart valve leaflet mechanics: exploitation of tissue structure. *Biomechanics and Modeling in Mechanobiology*, 15(4):909–932, 2016.

[9] S. Morganti, F. Auricchio, D. J. Benson, F. I. Gambarin, S. Hartmann, T. J. R. Hughes, and A. Reali. Patient-specific isogeometric structural analysis of aortic valve closure. *Computer Methods in Applied Mechanics and Engineering*, 284:508–520, 2015.

[10] LS-DYNA Finite Element Software: Livermore Software Technology Corp. http://www.lstc.com/products/ls-dyna. Accessed 17 July 2016.

[11] Simulia. http://www.3ds.com/products-services/simulia/. Accessed 6 April 2017.

[12] C.-H. Lee, R. Amini, R. C. Gorman, J. H. Gorman, and M. S. Sacks. An inverse modeling approach for stress estimation in mitral valve anterior leaflet valvuloplasty for in-vivo valvular biomaterial assessment. *J Biomech*, 47(9):2055–2063, 2014.

[13] C.-H. Lee, R. Amini, Y. Sakamoto, C. A. Carruthers, A. Aggarwal, R. C. Gorman, J. H. Gorman, and M. S. Sacks. Mitral valves: A computational framework. In S. De, W. Hwang, and E. Kuhl, editors, *Multiscale Modeling in Biomechanics and Mechanobiology*, pages 223–255. Springer London, 2015.

[14] C.-H. Lee, J. P. Rabbah, A. P. Yoganathan, R. C. Gorman, J. H. Gorman, and M. S. Sacks. On the effects of leaflet microstructure and constitutive model on the closing behavior of the mitral valve. *Biomechanics and Modeling in Mechanobiology*, 14(6):1281–1302, 2015.

[15] R. A. Sauer and L. De Lorenzis. A computational contact formulation based on surface potentials. *Computer Methods in Applied Mechanics and Engineering*, 253(0):369–395, 2013.

[16] J. Barnes and P. Hut. A hierarchical $O(N \log N)$ force-calculation algorithm. *Nature*, 324 (6096):446–449, 1986.

[17] V. Rokhlin. Rapid solution of integral equations of classical potential theory. *Journal of Computational Physics*, 60(2):187–207, 1985.

[18] P. Wriggers, J. Schröder, and A. Schwarz. A finite element method for contact using a third medium. *Computational Mechanics*, 52(4):837–847, 2013.

[19] T. Bog, N. Zander, S. Kollmannsberger, and E. Rank. Normal contact with high order finite elements and a fictitious contact material. *Computers & Mathematics with Applications*, 70 (7):1370–1390, 2015.

[20] T. Belytschko and M. O. Neal. Contact-impact by the pinball algorithm with penalty and Lagrangian methods. *International Journal for Numerical Methods in Engineering*, 31(3): 547–572, 1991.

[21] R. A. Sauer. *An Atomic Interaction based Continuum Model for Computational Multiscale Contact Mechanics*. PhD thesis, University of California, Berkeley, Berkeley, California, 2006.

[22] P. C. Guan, S. W. Chi, J. S. Chen, T. R. Slawson, and M. J. Roth. Semi-Lagrangian reproducing kernel particle method for fragment-impact problems. *International Journal of Impact Engineering*, 38(12):1033–1047, 2011.

[23] C. S. Peskin. Flow patterns around heart valves: A numerical method. *Journal of Computational Physics*, 10(2):252–271, 1972.

[24] D. Sulsky, Z. Chen, and H. L. Schreyer. A particle method for history-dependent materials. *Computer Methods in Applied Mechanics and Engineering*, 118(1):179–196, 1994.

[25] F. P. T. Baaijens. A fictitious domain/mortar element method for fluid–structure interaction. *International Journal for Numerical Methods in Fluids*, 35(7):743–761, 2001.

[26] C. Kadapa, W. G. Dettmer, and D. Perić. A stabilised immersed boundary method on hierarchical b-spline grids for fluid–rigid body interaction with solid–solid contact. *Computer Methods in Applied Mechanics and Engineering*, 318:242–269, 2017.

[27] T. H. Cormen, C. Stein, R. L. Rivest, and C. E. Leiserson. *Introduction to Algorithms*. McGraw-Hill Higher Education, 2nd edition, 2001.

[28] T. Belytschko and I. S. Yeh. The splitting pinball method for general contact. In *Proceedings of the 10th International Conference on Computing Methods in Applied Sciences and Engineering on Computing Methods in Applied Sciences and Engineering*, pages 73–87, Commack, NY, USA, 1991. Nova Science Publishers, Inc.

[29] T. Belytschko and I. S. Yeh. The splitting pinball method for contact-impact problems. *Computer Methods in Applied Mechanics and Engineering*, 105(3):375–393, 1993.

[30] F. Casadei, V. Aune, G. Valsamos, and M. Larcher. Generalization of the pinball contact/impact model for use with mesh adaptivity and element erosion in EUROPLEXUS. Technical Report JRC101013, European Commission: Joint Research Centre, 2016.

[31] Livermore Software Technology Corporation. LS-DYNA keyword user's manual R8.0—Volume I. 2015.

[32] Y. Bazilevs, V. M. Calo, T. J. R. Hughes, and Y. Zhang. Isogeometric fluid–structure interaction: theory, algorithms, and computations. *Computational Mechanics*, 43:3–37, 2008.

[33] H. Gomez and T. J. R. Hughes. Provably unconditionally stable, second-order time-accurate, mixed variational methods for phase-field models. *Journal of Computational Physics*, 230 (13):5310–5327, 2011.

[34] J. Kiendl, M.-C. Hsu, M. C. H. Wu, and A. Reali. Isogeometric Kirchhoff–Love shell formulations for general hyperelastic materials. *Computer Methods in Applied Mechanics and Engineering*, 291:280–303, 2015.

[35] İ. Temizer, P. Wriggers, and T. J. R. Hughes. Contact treatment in isogeometric analysis with NURBS. *Computer Methods in Applied Mechanics and Engineering*, 200(9):1100–1112, 2011.

[36] İ. Temizer, P. Wriggers, and T. J. R. Hughes. Three-dimensional mortar-based frictional contact treatment in isogeometric analysis with NURBS. *Computer Methods in Applied Mechanics and Engineering*, 209:115–128, 2012.

[37] L. De Lorenzis, İ. Temizer, P. Wriggers, and G. Zavarise. A large deformation frictional contact formulation using NURBS-based isogeometric analysis. *International Journal for Numerical Methods in Engineering*, 87(13):1278–1300, 2011.

[38] S. B. Raknes, X. Deng, Y. Bazilevs, D. J. Benson, K. M. Mathisen, and T. Kvamsdal. Isogeometric rotation-free bending-stabilized cables: Statics, dynamics, bending strips and coupling with shells. *Computer Methods in Applied Mechanics and Engineering*, 263:127–143, 2013.

[39] T. W. Sederberg, D. L. Cardon, G. T. Finnigan, N. S. North, J. Zheng, and T. Lyche. T-spline simplification and local refinement. *ACM Transactions on Graphics*, 23(3):276–283, 2004.

[40] Autodesk T-Splines Plug-in for Rhino. http://www.tsplines.com/products/tsplines-for-rhino.html. 2015.

[41] Rhinoceros. http://www.rhino3d.com/. [Accessed 19 July 2016].

[42] F. Cirak and M. West. Decomposition contact response (DCR) for explicit finite element dynamics. *International Journal for Numerical Methods in Engineering*, 64(8):1078–1110, 2005.

[43] E. Vouga, D. Harmon, R. Tamstorf, and E. Grinspun. Asynchronous variational contact mechanics. *Computer Methods in Applied Mechanics and Engineering*, 200(2528):2181–2194, 2011.

[44] E. Hairer, G. Wanner, and C. Lubich. *Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2006.

[45] M. Stevanella, E. Votta, M. Lemma, C. Antona, and A. Redaelli. Finite element modelling of the tricuspid valve: A preliminary study. *Medical Engineering & Physics*, 32(10):1213–1223, 2010.

[46] A. Drach, A. H. Khalighi, F. M. ter Huurne, C.-H. Lee, C. Bloodworth, E. L. Pierce, M. O. Jensen, A. P. Yoganathan, and M. S. Sacks. Population-averaged geometric model of mitral valve from patient-specific imaging data. *Journal of Medical Devices*, 9(3):030952–1–030952–3, 2015.

[47] A. H. Khalighi, A. Drach, C. H. Bloodworth, E. L. Pierce, A. P. Yoganathan, R. C. Gorman, J. H. Gorman, and M. S. Sacks. Mitral valve chordae tendineae: Topological and geometrical characterization. *Annals of Biomedical Engineering*, 45(2):378–393, 2017.

[48] K. O. Lim. Mechanical properties and ultrastructure of normal human tricuspid valve chordae tendineae. *The Japanese Journal of Physiology*, 30(3):455–464, 1980.

[49] L. A. Freed, D. Levy, R. A. Levine, M. G. Larson, J. C. Evans, D. L. Fuller, B. Lehman, and E. J. Benjamin. Prevalence and clinical outcome of mitral-valve prolapse. *New England Journal of Medicine*, 341(1):1–7, 1999.

[50] J. S. Raichlen and A. N. Brest. Tricuspid valve prolapse. *Cardiovascular Clinics*, 17(2):97–109, 1987.

[51] A. Buganza Tepole, H. Kabaria, K.-U. Bletzinger, and E. Kuhl. Isogeometric kirchhoff–love shell formulations for biological membranes. *Computer Methods in Applied Mechanics and Engineering*, 293:328–347, 2015.

[52] F. Roohbakhshan and R. A. Sauer. Efficient isogeometric thin shell formulations for soft biological materials. *Biomechanics and Modeling in Mechanobiology*, 16(5):1569–1597, 2017.

[53] M.C.H Wu, R. Zakerzadeh, D. Kamensky, J. Kiendl, M.S. Sacks, and M.-C. Hsu. An anisotropic constitutive model for immersogeometric fluid–structure interaction analysis of bioprosthetic heart valves. *Journal of Biomechanics*, 2017. In review.

[54] K. S. Kunzelman and R. P. Cochran. Stress/strain characteristics of porcine mitral valve tissue: parallel versus perpendicular collagen orientation. *Journal of Cardiac Surgery*, 7(1): 71–78, 1992.

[55] G. Krishnamurthy, D. B. Ennis, A. Itoh, W. Bothe, J. C. Swanson, M. Karlsson, E. Kuhl, D. C. Miller, and N. B. Ingels. Material properties of the ovine mitral valve anterior leaflet in vivo from inverse finite element analysis. *American Journal of Physiology–Heart and Circulatory Physiology*, 295(3):H1141–H1149, 2008.

[56] M. K. Rausch, N. Famaey, T. O. Shultz, W. Bothe, D. C. Miller, and E. Kuhl. Mechanics of the mitral valve: a critical review, an in vivo parameter identification, and the effect of prestrain. *Biomechanics and Modeling in Mechanobiology*, 12(5):1053–1071, 2013.

[57] A. Aessopos, D. Farmakis, H. Taktikou, and D. Loukopoulos. Doppler-determined peak systolic tricuspid pressure gradient in persons with normal pulmonary function and tricuspid regurgitation. *Journal of the American Society of Echocardiography*, 13(7):645–649, 2000.

[58] M. Pigazzini, Y. Bazilevs, D. Benson, H. Kim, and A. Ellison. Isogeometric analysis of damage and residual-strength in aerospace composite structures subjected to low-velocity impact. In *Proceedings of the American Society for Composites: Thirty-First Technical Conference*, 2016.

[59] Y. Bazilevs, K. Kamran, G. Moutsanidis, D. J. Benson, and E. Oñate. A new formulation for air-blast fluid–structure interaction using an immersed approach. Part I: basic methodology and FEM-based simulations. *Computational Mechanics*, 60(1):83–100, 2017.

[60] Y. Bazilevs, G. Moutsanidis, J. Bueno, K. Kamran, D. Kamensky, M. C. Hillman, H. Gomez, and J. S. Chen. A new formulation for air-blast fluid–structure interaction using an immersed

approach: Part II—coupling of IGA and meshfree discretizations. *Computational Mechanics*, 60(1):101–116, 2017.