ORIGINAL PAPER

Internal contact modeling for finite strain topology optimization

Gore Lukas Bluhm[1] · Ole Sigmund[1] · Konstantinos Poulios[1]

Abstract

The present work proposes an extension of the third medium contact method for solving structural topology optimization problems that involve and exploit self-contact. A new regularization of the void region, which acts as the contact medium, makes the method suitable for cases with very large deformations. The proposed contact method is implemented in a second order topology optimization framework, which employs a coupled simultaneous solution of the mechanical, design update, and adjoint problems. All three problems are derived and presented in weak form, and discretized with finite elements of suitable order. The capabilities and accuracy of the developed method are demonstrated in a topology optimization problem for achieving a desired non-linear force–displacement path.

1 Introduction

In structural topology optimization under large deformations, it is not uncommon that beam-like members eventually come into contact. Modeling of self-contact is therefore essential in order to avoid non-physical self-penetration in the mechanical analysis and thereby extend the accuracy and validity range of such optimization solutions. Even more than that, a self-contact aware optimization can lead to radically different designs by exploiting the possibility of unilateral contact between parts of the design in e.g. non-linear mechanism design.

Applying standard methods for modeling of self-contact to situations like this, where the actual contact interfaces are a priori unknown, is in principle possible but also very tedious [1]. Applying the third medium contact method [2,3], instead, is a very natural and simple choice because density based topology optimization already involves a void region which can play the role of the contact medium. Still, there are important technical aspects that require a careful treatment in order to make the combination numerically stable.

The third medium contact method was introduced in [2] as a purely continuum mechanics based alternative to contact domain methods [4]. Its main advantage against standard contact methods, such as augmented Lagrangian and penalty ones [5,6], is that it avoids an explicit treatment of the contact kinematics and the implementation of inequality constraints. However, it comes with the disadvantages of i) parasitic remote forces transferred between the bodies before the actual contact, and ii) not permitting large amounts of sliding between the contacting bodies. Due to these limitations, third medium contact has hitherto not found significant applications as a replacement to the established contact methods based on Lagrange multipliers.

When it comes to topology optimization though, the third medium contact approach is an excellent fit, because density based topology optimization already employs a void phase in the a priori unknown regions where contact may actually occur. By appropriate constitutive modeling, the void phase can easily be used as the contact medium. At the same time, the inherent limitation of the method in capturing a sharp transition from zero to finite contact stresses, turns into an advantage in topology optimization. The parasitic interaction of non-contacting parts of the structure due to small but finite stiffness of the third medium, makes the problem differentiable and hence allows the optimization to drive a design towards a self-contact state, if this is beneficial for the objective being optimized.

✉ Gore Lukas Bluhm
   glubluh@mek.dtu.dk

1   Department of Mechanical Engineering, Technical University of Denmark, Nils Koppels Allé, Building 404, 2800 Kgs. Lyngby, Denmark

The idea of employing third-medium contact methods in topology optimization for modeling internal and external contact has so far been unexplored. Nevertheless, traditional modeling of, mainly external, contact in topology optimization has been pursued in a series of papers. Apart from earlier works in the field, reviewed by Hilding et al. [7], more recent works within small deformations topology optimization under external contact with rigid obstacles, include compliance or volume minimization in a frictionless [8,9] or frictional [10] setting, as well as compliance or contact pressure minimization in a frictional setting [11]. Regarding large deformations topology optimization, Luo et al. [12] performed compliance minimization under frictionless contact with circular rigid obstacles, modeled by means of nonlinear springs. Fernandez et al. [13] used conventional large deformation contact methods to model external contact between the optimized design and a prescribed external deformable solid. Known efforts for accounting for internal contact in compliance minimization problems are limited to a priori fixed internal contact interfaces, either without [14] or with friction [15]. Otherwise, the level set method has also been successfully used for internal contact, though limited to zero initial gap and small sliding [16].

Apart from its simplicity, the major advantage of third medium contact in the context of density based topology optimization is that it completely eliminates the need for prescribing potential internal or external contact interfaces. The main challenge lies in defining a constitutive law for the contact medium that is numerically stable even at extremely severe deformations and large sliding. Apart from the currently employed constitutive laws for the void phase in large deformation topology optimization [17,18], specialized third medium constitutive laws from the contact mechanics literature [2,3] provide a foundation for further developments. The original model from [2] with an explicitly added anisotropic stiffness term, has already been simplified in [3], based on the fact that a highly distorted hyperelastic continuum actually exhibits the desired anisotropy naturally. In the present work, a similar neo-Hookean material is used for the contact medium as the Hencky material used in [3]. In addition, in order to allow for contact situations with considerable sliding and very distorted elements, the present work proposes a new void regularization based on penalization of higher order strains in void elements.

The basic idea can be basically implemented in any density based topology optimization algorithm accounting for large strains [18,19], minimizing a given objective function under a set of given constraints. However, the present work introduces and uses a new second order topology optimization framework, fully defined in the continuous setting prior to discretization. The approach is similar to the continuous adjoint sensitivity analysis by Cho and Jung [20], and is presented in a rather generalized form before being particularized to a specific hyperelastic material law and the objective function considered in the numerical example. The included numerical results were obtained with the displacement field approximated with quadratic quadrilateral elements and the density field approximated with linear quadrilateral elements in a structured mesh but the actual method and its implementation is neither limited to structured meshes nor bound to a specific element type.

The subsequent Sect. 2 presents some important adaptations of the third medium contact method that extend its applicability to larger amounts of sliding. The topology optimization framework for the present work is introduced in Sects. 3 and 4 presents an optimization example that exploits the occurrence of self contact for achieving a desired nonlinear force–displacement response. Important conclusions are summarized in Sect. 5, while Table 1 summarizes notation conventions used throughout the present work.

Table 1 Notation conventions

$A \cdot B = A_i B_i$	Scalar product
$A : B = A_{ij} B_{ij}$	Double contraction
$A \vdots B = A_{ijk} B_{ijk}$	Triple contraction
$\lvert A \rvert$	Determinant of matrix A
$\lVert A \rVert = \sqrt{A : A}$	Frobenius norm of matrix A
$\nabla A = \dfrac{\partial A_i}{\partial X_j}$	Spatial gradient of a vector field A
$\mathbb{H} A = \dfrac{\partial^2 A_i}{\partial X_j \partial X_k}$	Spatial Hessian of a vector field A
$\mathrm{dev}(A)$	3D deviator operator
$\langle x \rangle = \max(0, x)$	Positive part function

2 Mechanical model

The modeling domain Ω consists of a solid structure subdomain Ω_s embedded in a void phase subdomain Ω_v. A vector field u defined in the undeformed domain Ω expresses displacements from the initial to the current deformed configuration. Let also q denote a Lagrange multiplier vector which can either be a single vector or a vector field over some part Γ_D of the domain boundary $\partial\Omega$. In the former case, it is used to enforce a Dirichlet boundary condition in an average sense, while in the latter case, the condition is enforced at every point of Γ_D. Based on these definitions, the mechanical equilibrium of the system can be expressed in weak form as

$$\mathcal{R}(u, q; \delta u, \delta q) = 0 \quad \forall \delta u, \delta q \tag{1}$$

with

$$\mathcal{R}(u, q; \delta u, \delta q) = \int_{\Omega} \Psi_{,u}(u; \delta u) \, d\Omega$$
$$+ \int_{\Gamma_D} \left\{ q \cdot \delta u + (u - u_D) \cdot \delta q \right\} d\Gamma \tag{2}$$

where $\Psi_{,u}(u; \delta u)$ is the virtual variation of the strain energy density function Ψ due to an infinitesimal displacements variation δu. This is the usual virtual work expression that can be rewritten in many equivalent forms of work conjugate pairs of strain variations and stresses. Depending on whether the Dirichlet condition on Γ_D is applied in an average sense or pointwise, u_D is either a single vector or a vector field expressing the prescribed displacement or displacement field.

2.1 Hyperelasticity

Regarding the strain energy function Ψ, there are many possible choices for hyperelastic material laws. Most of these are formulated so that Ψ tends to infinity when the material is compressed towards zero volume. This property, apart from representing the actual physical response of any real material, is also essential for the third medium contact model employed here.

A rather common isotropic neo-Hookean material law according to Simo et al. [21] is adopted in the present work both for the solid and the void domain, defined through the strain energy density function

$$\overline{\Psi}(u) = \frac{K}{2} \left(\ln |F|\right)^2 + \frac{G}{2} \left(|F|^{-2/3} \|F\|^2 - 3\right) \tag{3}$$

with the deformation gradient $F = I + \nabla u$. The initial bulk modulus K and shear modulus G depend on the material domain, being respectively equal to K_s and G_s in Ω_s, and K_v and G_v in Ω_v. These elasticity constants in the void domain Ω_v are very small but finite.

Even if the void stiffness is initially negligible compared to the solid stiffness, the presence of the $\ln |F|$ term in Eq. (3) will eventually lead to an infinite stiffness when the void is compressed to zero volume. This property is crucial for the present application, because it results in a compressed void which is stiffer than the solid phase and can therefore transfer forces between solid members of the structure. In general, every material law with this characteristic can serve as a third medium for modeling contact, in contrast to material laws, like the otherwise frequently used Saint Venant-Kirchhoff model, that result in vanishing stresses under ultimate compression.

The virtual work expression appearing in Eq. (2), results from Gateaux differentiation of Ψ with respect to the dis-

placements field u. For the hyperelastic strain energy density from Eq. (3), this operation results in

$$\overline{\Psi}_{,u}(u; \delta u) = P(\nabla u) : \nabla \delta u \tag{4}$$

with the 1st Piola–Kirchhoff stress tensor

$$P(\nabla u) = K \ln |F| \, F^{-T} + G \, |F|^{-2/3} \operatorname{dev}(FF^T) F^{-T} \tag{5}$$

All equations have been presented in their three dimensional form, however a reduction to plane strain is trivial by defining the three dimensional deformation gradient as

$$F = I + \begin{bmatrix} \nabla u & 0 \\ & 0 \\ 0 \ 0 \ 0 \end{bmatrix}. \tag{6}$$

At the same time, for plane strain problems, P in Eq. (4) needs to be reduced to its in-plane components.

2.2 Void regularization

In general, the constitutive behavior of the void is not essential for the mechanical system as long as stresses in the void are small enough to not affect the deformation of the solid significantly. Nevertheless, a rigorous modeling of the void up to large strains is essential for avoiding numerical instabilities of the overall system. This is especially true when the void is used as the third medium for contact modeling, where many elements will collapse to almost zero volume.

Figure 1 shows a classical benchmark example for dealing with large deformations in the void region [17,18]. In the present work, the upper beam is loaded by prescribing the average vertical displacement within the small region Γ_D through an incremented vertical component u_{Dy} in the vector u_D in Eq. (2). The horizontal displacement of the Γ_D region is left free by defining $q = \{0, q_y\}^T$ and $\delta q = \{0, \delta q_y\}^T$ in Eq. (2), reducing the formulation to a scalar Lagrange multiplier q_y, which is just a single additional degree of freedom in the overall system. The homogeneous Dirichlet condition on Γ_{Dh} can be imposed by restricting the solution space of u.

The plots in the first column of Fig. 2 show simulation results for the reference solution with $K_s = 1$ and $G_s = 6/13$ in stress units, and the void elements completely removed. The second column shows the corresponding results after including void elements with $K_v/K_s = G_v/G_s = 10^{-12}$. All results are obtained with a 20×20 discretization of the total domain Ω with 8-node quadratic quadrilateral elements and numerical integration with nine Gauss points. Although the simulation does not break down in the load range shown in Fig. 2, the deformed void mesh in the second column has many severely deformed elements which are close to
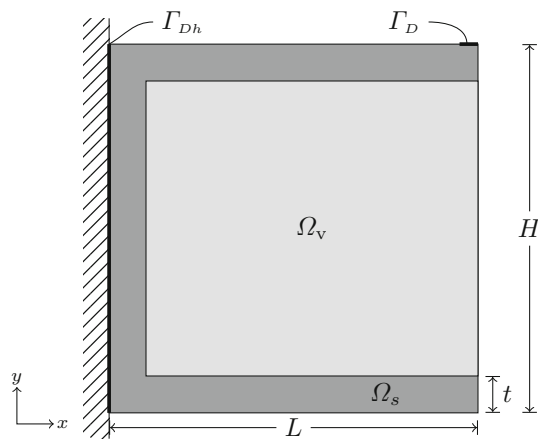
Fig. 1 C-shape structure with $t = 0.1L$ and a length of Γ_D equal to $0.05L$

being inverted. At higher load levels, inverted or severely distorted elements lead to numerical instabilities preventing the employed Newton algorithm from converging.

Different methods have been successfully applied for avoiding this kind of degenerate states of the deformed void [17,18]. The present work proposes a new regularization which gives excellent control over the void mesh deformation and at the same time is compatible with the use of the void as a third medium for contact modeling. The basic idea is to augment the material strain energy function $\overline{\Psi}$ with an energy associated with higher order strains within each finite element according to an energy density proportional to the function

$$\tilde{\Phi} = \frac{1}{2}\mathbb{H}u : \mathbb{H}u. \tag{7}$$

where $\mathbb{H}u$ is the Hessian of the displacement field, i.e. the second order strain tensor. The colorplot shown in the second column of Fig. 2 illustrates large values of $\tilde{\Phi}$ occurring in regions of severe mesh distortion.

In order to penalize these deformation modes, one could consider an augmented strain energy function $\overline{\Psi} + k_r \mathcal{I} \tilde{\Phi}$, with the regularization scaling constant k_r and the void indicator function $\mathcal{I}$, which would lead to the virtual work expression

$$\overline{\Psi}_{,u}(u;\ \delta u) + k_r \mathcal{I} \mathbb{H}u : \mathbb{H}\delta u. \tag{8}$$

This is a rather non-intrusive approach because it would only penalize bending and warping deformation modes in void elements, while it would not penalize homogeneous deformation states even at very severe homogeneous compression and shearing.

In order to reduce the impact of the regularization term further, an ad hoc scaling of the higher order term in Eq. (8) is proposed, leading to the weak form
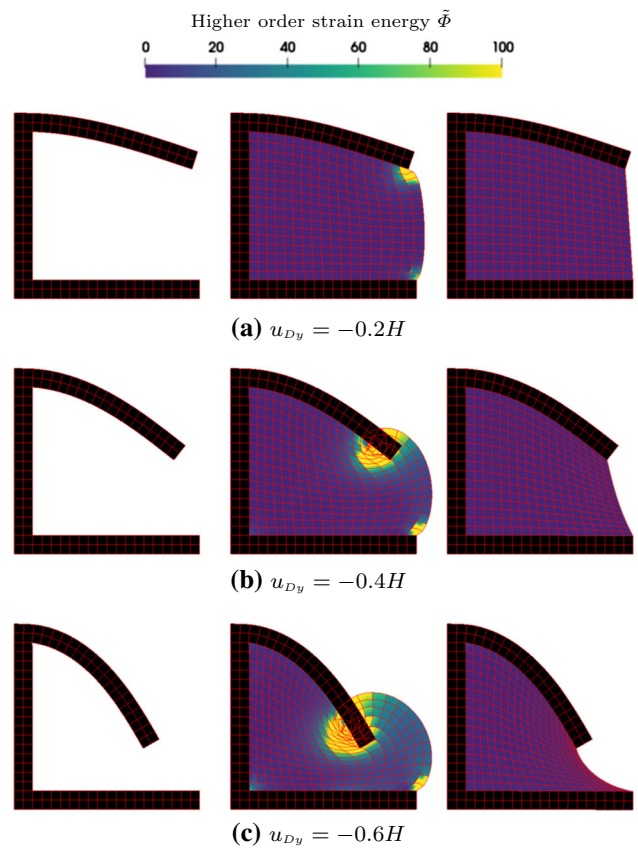
Fig. 2 Deformed C-shape structure for $H = L = 1$, with no void material (left), without void regularization (middle) and with void regularization (right). Solid structure is shown in black, void is colored according to the higher order strain energy $\tilde{\Phi}$ defined in Eq. (7)

$$\Psi_{,u}(u;\ \delta u) = \overline{\Psi}_{,u}(u;\ \delta u) + k_r \mathcal{I} e^{-5|F|} \mathbb{H}u : \mathbb{H}\delta u \tag{9}$$

to be used, with a slight exploitation in notation, in Eq. (2). The added negative exponential dependency on the determinant of the deformation gradient is derived from numerical experiments. It leads to a decaying intensity of the regularization for elements that are not severely compressed, or even less for elements that are stretched.

The penalization constant k_r has force units. In order to work with a dimensionless quantity with more general validity, the following scaling relationship is proposed

$$k_r = \bar{k}_r L^2 K_s. \tag{10}$$

where L is used as a characteristic length of the structure and K_s is the elastic bulk modulus of the solid. The dimensionless constant $\bar{k}_r$ has to be chosen just large enough to avoid extreme distortion of void elements. Convergence studies with regard to the parameter $\bar{k}_r$ have led to the value $\bar{k}_r = 10^{-6}$ used in all numerical examples throughout the present work. Note that the added Hessian term serves its purpose of regularizing the deformed element shape, even

for finite element spaces for u which are not C^1 continuous across element boundaries, as a compatible discretization of this term would require.

C-shape structure example

The last column in Fig. 2 shows simulation results for the square C-shape structure with the void regularization according to Eq. (9) and otherwise the same void material parameters as the results of the second column. Before adding the regularization, the void material was squeezed out at the open right boundary of the domain, resulting in a severely bent mesh, especially near the tip of the loaded beam. With the proposed regularization, all void region elements remain completely regular, which is also reflected in the much lower bending and warping energy plotted in the void region. When looking at the bottom beam at the highest load, a barely visible negative vertical deflection for the case with the regularized void, reveals a stronger parasitic interaction between the loaded and the unloaded beam. This weak but still detectable effect is consequence of an added bending stiffness due to the regularization.

The last load step, shown in Fig. 2c, indicates how more prone to numerical instabilities the model without void regularization is, compared to the model with void regularization. There exist other void regularization schemes in the literature which are as effective in this benchmark example. For this reason, the additional potential of the proposed regularization is further illustrated in a modified version of this example, that involves contact between the two beams.

Half height C-shape structure example

In the next example, shown in Fig. 3, the height of the C-shape structure is reduced to half, i.e. $H = 0.5L$, so that upon vertical loading of the upper beam, the two horizontal beams will eventually come into contact. The effectiveness of the proposed void regularization in modeling self-contact is demonstrated by imposing a displacement of the upper beam up to $u_{Dy} = -H$. The structure is discretized with 40 times 20 elements along the length and height directions, respectively.

In fact, this is a rather challenging contact example, because contact will occur at the very corner of the upper beam against the top surface of the lower beam. The first column of Fig. 3 shows the deformed structure at different loading states, with no void material and conventional contact modeling with Lagrange multipliers [22], as a reference solution. The second column shows results with the void between the two horizontal beams modeled with the void regularization according to Eq. (9). The void mesh remains very regular, as was the case in the previous example, but despite the absence of any inverted elements, the free edge

at the right boundary of the domain folds into a configuration that leads to overlapping. The fact that the expected contact point is at the boundary of the void domain prevents the void from acting as a contact medium between the two beams.

This is easy to overcome simply by extending the void domain with an additional column of elements at the right boundary of the domain, as shown in the last column of Fig. 3. In fact, when applying the third medium approach for contact, it is essential, that every potential contact region is entirely embedded in the contact medium. With a void domain that sufficiently embeds the contact region in this case, the regularized void acts as a contact medium transferring forces and permitting penetration between the two beams even for the significant amount of sliding observed at the maximum imposed displacement $u_{Dy} = -H$. At a moderate displacement of the upper beam with $u_{Dy} = -0.2H$, both simulations with the regularized void exhibit no visible deviations from the reference solution. At the intermediate loading state with $u_{Dy} = -0.5H$, a barely visible deflection of the lower beam indicates a small parasitic interaction between the two beams that have not yet come into contact. At the post contact deformation state with $u_{Dy} = -H$, the result seen in the right hand side column for the regularized void with the additional buffer layer of void is in good agreement with the reference solution.

A more quantitative comparison is provided in Fig. 4 that illustrates the obtained load displacement curves for the three cases from Fig. 3. The two curves for the simulations with void material coincide until the onset of contact and exhibit a just slightly larger reaction force compared to the reference case. Beyond the occurrence of contact, the case with the appropriately extended void domain follows the reference solution with a very small deviation even at a significant amount of sliding. The inset detail view of the transition point at the onset of contact helps estimating the loss of accuracy in predicting a sharp transition due to the inherent limitation of the third medium contact approach, compared to conventional contact modeling.

2.3 Material interpolation

So far, only structures have been treated that involve solid and void domains clearly delimited by a discontinuous interface. However, density based topology optimization relies on the concept of a continuous material density varying from void to solid in a differentiable manner. A graded interface between the two regions is essential for establishing the sensitivity of an objective function with respect to changes of the design. With this motivation, a material interpolation scheme is introduced below for representing a transition from solid to void regions, and the effectiveness of the proposed third medium contact is also evaluated in this context.
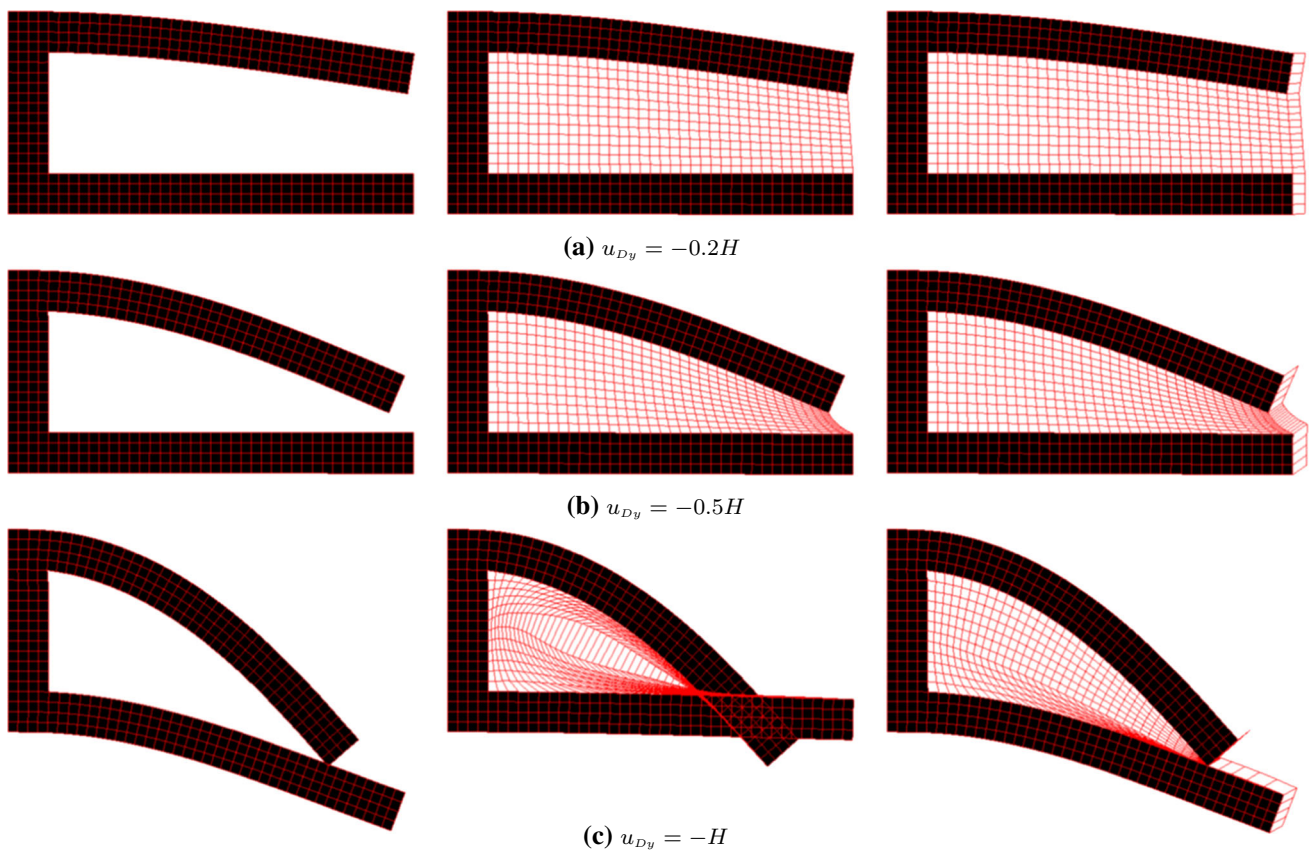
(a) $u_{D_y} = -0.2H$

(b) $u_{D_y} = -0.5H$

(c) $u_{D_y} = -H$

Fig. 3 Deformations of the C-shape structure with $H = 0.5L$, with no void material solved using a conventional contact formulation based on Lagrange multipliers (left), with void regularization (middle) and with void regularization and an additional layer of void material at the open right hand side boundary (right)
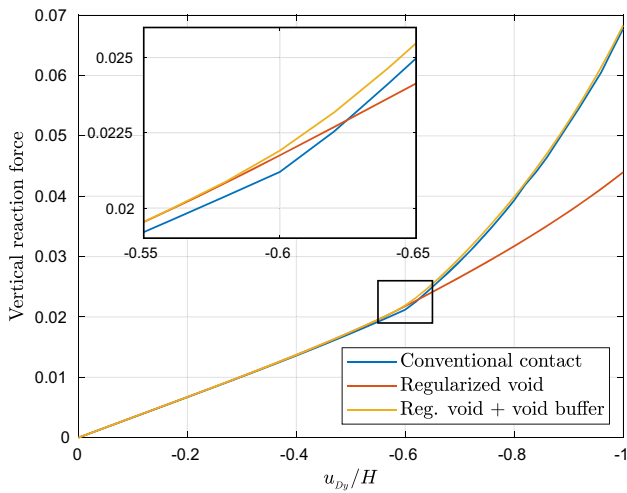

Fig. 4 Load–displacement responses at Γ_D for the three simulations of the C-shape structure from Fig. 3

The design is described by a continuous design variable field χ which represents a purely mathematical unconstrained quantity, i.e. $\chi \in (-\infty, \infty)$. To express material

consumption, χ is mapped to a physical material density ρ through the function

$$\rho(\chi) = \frac{1}{1 + e^{-\chi}}, \tag{11}$$

which results in a density field within the asymptotic limits of 0 and 1, corresponding to a perfect void and a perfect solid material, respectively, as illustrated in Fig. 5a. According to this mapping, the conventional solid to void boundary at $\rho = 0.5$ corresponds to $\chi = 0$, while for values of χ in the order of $+10$ and -10, respectively, the deviation from a perfect solid and void material, has no practical significance.

Although the density field is continuous, the aim is nevertheless to achieve manufacturable and discrete designs where ρ is either at its upper or lower bound, while intermediate densities are limited to the interface between solid and void domains. Usually this is done by penalizing the stiffness of intermediate densities through a convex relationship between material usage ρ and material performance, i.e. stiffness for structural topology optimization. The present work adopts the RAMP material interpolation function by Stolpe and Svanberg [23] due to its finite initial slope. Expressed in
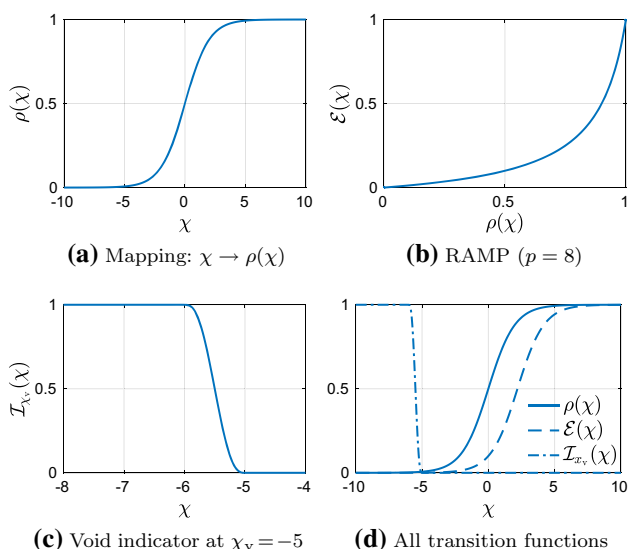
(a) Mapping: $\chi \rightarrow \rho(\chi)$

(b) RAMP $(p = 8)$

(c) Void indicator at $\chi_v = -5$

(d) All transition functions

Fig. 5 **a** Mapping from the design variable to the physical density, Eq. (11). **b** RAMP material stiffness interpolation function, Eq. (12). **c** Void indicator function, Eq. (13), for activating void regularization. **d** Comparison of material density, material stiffness and void regularization transition functions

terms of the mathematical design variable χ, this interpolation is represented by the function

$$\mathcal{E}(\chi) = \mathcal{E}_0 + (1 - \mathcal{E}_0) \; \frac{\rho(\chi)}{1 + p(1 - \rho(\chi))}, \tag{12}$$

where p is the penalization parameter and $\mathcal{E}_0$ is a small positive value ensuring minimum stiffness even if the density is very close to zero. The RAMP interpolation function from Eq. (12) is applied as scaling to the solid material elastic properties resulting in the initial bulk modulus $K = \mathcal{E}(\chi)K_s$ and the initial shear modulus $G = \mathcal{E}(\chi)G_s$. A linear interpolation is obtained for $p = 0$, while increasingly positive values for p result in an increasing penalization of intermediate densities. Figure 5b illustrates the RAMP function for the penalization parameter $p = 8$, used throughout the present work, while the compound function $\mathcal{E}(\chi)$ for the same value of p is shown in Fig. 5d.

In the context of a continuous material transition from void to solid, discussed in this subsection, the void indicator function $\mathcal{I}$, which acts as a switch for the void regularization introduced in Eq. (9), needs also to be redefined in a differentiable manner. In order to apply the regularization term only to regions where the design variable χ is below a specified threshold χ_v, the following smooth negative step function between $\chi_v - 1$ and χ_v, adopted from [24], can be used.

$$\mathcal{I}_{\chi_v}(\chi) = 6\langle\chi_v - \chi\rangle^5_{[0,1]} - 15\langle\chi_v - \chi\rangle^4_{[0,1]} \\ + 10\langle\chi_v - \chi\rangle^3_{[0,1]} \tag{13}$$

with

$$\langle\chi\rangle_{[0,1]} = \begin{cases} 0 & \text{if } \chi < 0 \\ 1 & \text{if } \chi > 1 \\ \chi & \text{if } 0 \leq \chi \leq 1 \end{cases} \tag{14}$$

Figure 5c illustrates this smooth indicator function for the threshold value $\chi_v = -5$, used throughout this work. The function yields 0 for $\chi \geq -5$ and 1 for $\chi \leq -6$, which means that the void regularization is completely absent in solid and intermediate density regions, as also seen in Fig. 5d.

After having introduced a dependence of the material elasticity parameters on the design variable χ and after replacing $\mathcal{I}$ in Eq. (9) with the smooth void indicator function from Eq. (13), the mechanical equilibrium residual according to Eq. (2) becomes

$$\mathcal{R}(\chi, u, q; \delta u, \delta q) = \\ \int_{\Omega} \left\{ \mathcal{E}(\chi) P_s(\nabla u) : \nabla \delta u \\ + k_r \, \mathcal{I}_{\chi_v}(\chi) \, e^{-5|F(\nabla u)|} \, \mathbb{H}u \, \vdots \, \mathbb{H}\delta u \right\} d\Omega \\ + \int_{\Gamma_D} \left\{ q \cdot \delta u + (u - u_D) \cdot \delta q \right\} d\Gamma, \tag{15}$$

where P_s is the 1st Piola–Kirchhoff stress tensor from Eq. (5) for the completely solid material, i.e. $K = K_s$ and $G = G_s$.

Half height C-shape structure with graded interface

Introducing a graded transition from void to solid will unavoidably affect the structure's mechanical behavior compared to a discontinuous transition, especially when studying the contact between solid members. In order to evaluate the effectiveness of the third medium contact approach with the proposed void regularization, in the context of a graded void-solid interface, the last example from the previous subsection is reevaluated. The domain discretization is now refined by a factor of two, resulting in 82×40 elements, and the transition from solid to void, i.e. from $\chi = -10$ to $\chi = 10$, is spread over two elements, with the transition being linear in terms of χ. In order to match the previous example, the minimum stiffness factor in Eq. (12) is set to $\mathcal{E}_0 = 10^{-12}$. Figure 6 shows the resulting material density field and the deformed structure at different loading states.

Figure 7 provides a more quantitative comparison between the model with the discontinuous solid to void transition from Fig. 3 and the model with the graded transition from Fig. 6. The upper diagram of Fig. 7 shows the reaction force as a function of the imposed vertical displacement to the upper beam for the two investigated cases, while the lower diagram shows the corresponding curve slopes.
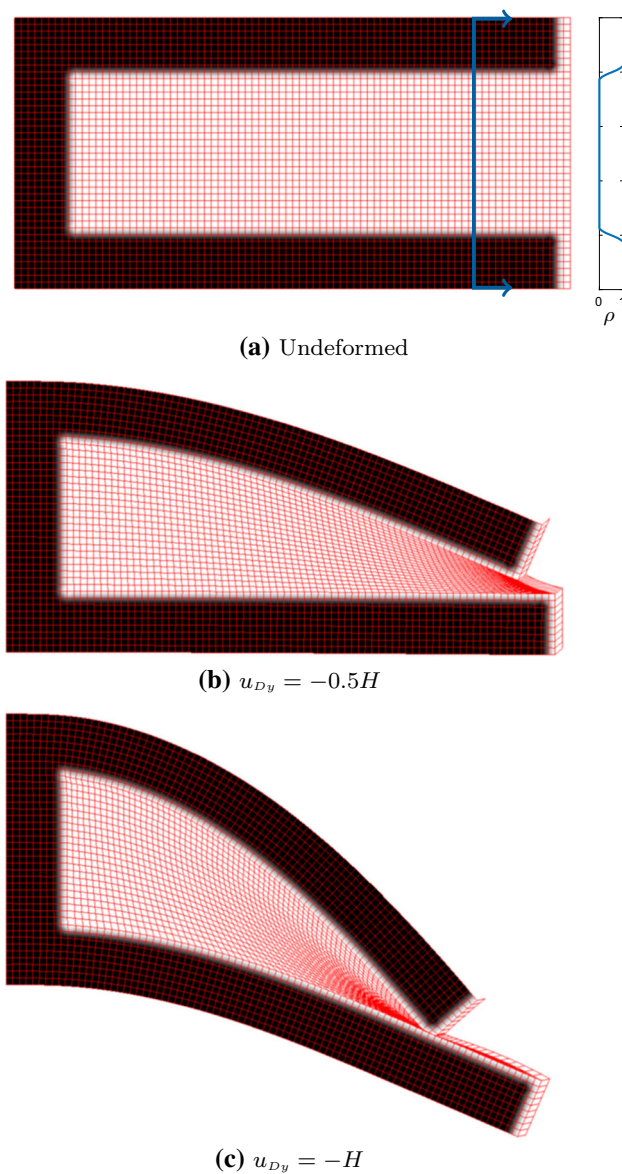
(a) Undeformed

(b) $u_{Dy} = -0.5H$

(c) $u_{Dy} = -H$

Fig. 6 C-shape structure with $H = 0.5L$ and graded interface in its underfomed state (**a**) and two deformed states at different load levels (**b**, **c**)

Fig. 7 Load and stiffness response as function of prescribed displacement for the discontinuous void-solid transition (Fig. 3) and the graded void-solid interface (Fig. 6)

Apart from that, additional effects emerge in the context of contact modeling. Most significantly due to the finite thickness of the graded interface between solid and void, contact now occurs prematurely as seen in Fig. 6c and the lower diagram of Fig. 7. Moreover, the presence of a layer of intermediate density material between the contacting solids adds some compliance to the contact, resulting in a less sharp transition from non-contact to contact, as observed e.g. in the force displacement curves shown.

Despite the presence of the mentioned effects, it is to be concluded that the overall contact behavior is captured adequately for topology optimization applications. Especially the inaccuracy of remote interaction of solid members, can actually be an advantage in the context of optimization, because it provides information about potential contact interfaces in the surroundings that should be avoided or exploited, depending on the goal of the optimization and makes the contact problem differentiable.

The observed deviations between the two cases are due to different effects. The interpretation of the original discontinuous interface as the graded interface shown in Fig. 6a, in combination with the RAMP interpolation applied to the intermediate density material, leads to a reduced effective beam thickness due to the lower stiffness of the intermediate density material at the interface. For this reason, the graded interface model is in general more compliant than the one with the discontinuous interface. It is an accepted effect in topology optimization, that designs with intermediate densities are in general more compliant than crisp solid and void designs.
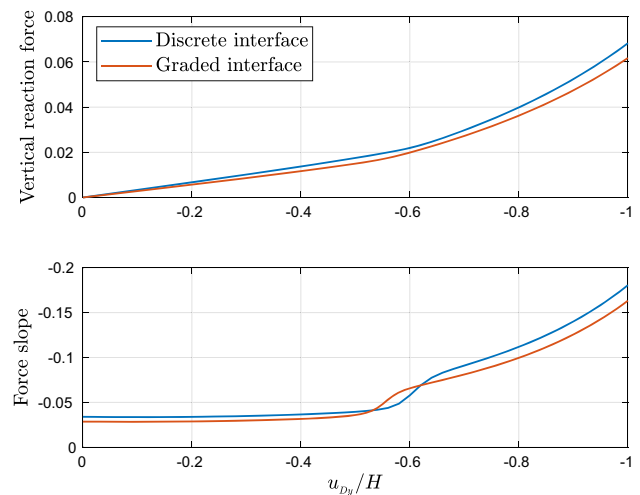
3 Topology optimization formulation

The method for third medium contact with the void regularization presented above, can in theory be implemented in every topology optimization framework for finite strains. This section presents its implementation in a new generalized framework, which is based on the design parametrization from Sect. 2.3, and a discretization-agnostic formulation. The proposed optimization scheme is entirely defined in the continuous setting in weak form and is second order consistent.

3.1 Optimization framework

Consider an optimization problem defined as the minimization of an objective function C, i.e.

$$\min_{\chi} \ C(\chi, u, q) \tag{16}$$

subjected to the mechanical equilibrium constraint

$$\mathcal{R}(\chi, u, q; \ \delta u, \delta q) = 0 \quad \forall \delta u, \delta q \tag{17}$$

corresponding to Eq. (1) but with the residual function $\mathcal{R}$ also including the dependence on the design variable χ introduced in Eq. (15).

A simple but still rather general form for the objective function C is the following additive split in three single variable functions

$$C(\chi, u, q) = C_{\chi}(\chi) + C_u(u) + C_q(q). \tag{18}$$

Minimization of the objective function under the aforestated mechanical equilibrium constraint, is equivalent to minimization of the augmented objective function

$$C^*(\chi, u, q) = C(\chi, u, q) + \mathcal{R}(\chi, u, q; \ \lambda_u, \lambda_q) \tag{19}$$

for any λ_u and λ_q respectively in the spaces of δu and δq. The optimality condition is based on the variation of the objective function C which can be evaluated as

$$
\begin{aligned}
\delta C = \delta C^* \\
= \ & C_{\chi,\chi}(\chi; \ \delta\chi) + \mathcal{R}_{,\chi}(\chi, u, q; \ \lambda_u, \lambda_q; \delta\chi) \\
& + C_{u,u}(u; \ \delta u) + \mathcal{R}_{,u}(\chi, u, q; \ \lambda_u, \lambda_q; \delta u) \\
& + C_{q,q}(q; \ \delta q) + \mathcal{R}_{,q}(\chi, u, q; \ \lambda_u, \lambda_q; \delta q).
\end{aligned}
\tag{20}
$$

Applying the adjoint sensitivity analysis method, the variations δu and δq are eliminated from Eq. (20) by finding multipliers λ_u and λ_q that satisfy the adjoint equations

$$
\begin{aligned}
\mathcal{R}_{,u}(\chi, u, q; \ \lambda_u, \lambda_q; \delta u) + C_{u,u}(u; \ \delta u) = 0 \ \forall \delta u \\
\mathcal{R}_{,q}(\chi, u, q; \ \lambda_u, \lambda_q; \delta q) + C_{q,q}(q; \ \delta q) = 0 \ \forall \delta q.
\end{aligned}
\tag{21}
$$

The variation of the objective is thereby reduced to

$$\delta C = C_{\chi,\chi}(\chi; \ \delta\chi) + \mathcal{R}_{,\chi}(\chi, u, q; \ \lambda_u, \lambda_q; \delta\chi). \tag{22}$$

In total, mechanical equilibrium, adjoint analysis and optimality are expressed by the following system of coupled nonlinear equations

$$
\begin{aligned}
\mathcal{R}(\chi, u, q; \ \delta u, \delta q) = 0 \quad \forall \delta u, \delta q \\
\mathcal{R}_{,u}(\chi, u, q; \ \lambda_u, \lambda_q; \delta u) + C_{u,u}(u; \ \delta u) = 0 \ \forall \delta u \\
\mathcal{R}_{,q}(\chi, u, q; \ \lambda_u, \lambda_q; \delta q) + C_{q,q}(q; \ \delta q) = 0 \ \forall \delta q \\
C_{\chi,\chi}(\chi; \ \delta\chi) + \mathcal{R}_{,\chi}(\chi, u, q; \ \lambda_u, \lambda_q; \delta\chi) = 0 \ \forall \delta\chi
\end{aligned}
\tag{23}
$$

with regard to the unknowns χ, u, q, λ_u and λ_q. Assuming that all functions involved in Eq. (23) are differentiable with respect to the five unknowns, a consistent linearization of this system of equations is trivial to obtain in order to apply Newton's method. Application of Newton's method to the fully coupled Eq. (23) corresponds to a second order optimization scheme and it will lead to quadratic converge, simultaneously towards the optimal design and mechanical equilibrium, if the initial guess is sufficiently close to the solution.

Due to the extremely nonlinear nature of the design problem though, the initial guess for χ will never be close enough to the final design for Newton's algorithm to converge. A reasonable strategy for dealing with this without losing the second order information on the system is to add a damping on the last equation in Eq. (23) and solve the system as an ODE with an implicit time discretization, where each time step is solved with Newton's method. Somewhat similar to reference [25], but also introducing damping of spatial gradients of χ, the optimality condition $\delta C = 0$, appearing in the last row of Eq. (23), is now substituted with

$$\delta C + \int_{\Omega} \left\{ \dot{\chi}\delta\chi + l_t^2 \nabla\dot{\chi} \cdot \nabla\delta\chi \right\} d\Omega = 0 \ \forall \delta\chi \tag{24}$$

where time derivatives are with regard to a pseudo-time t, and l_t is a characteristic length-scale for the spatial evolution of the design field χ. Assuming that the design field χ_{old} at pseudo-time $t - \Delta t$ is known, the design field χ at pseudo-time t can be found by solving the system

$$
\begin{aligned}
\mathcal{R}(\chi, u, q; \ \delta u, \delta q) = 0 \quad \forall \delta u, \delta q \\
\mathcal{R}_{,u}(\chi, u, q; \ \lambda_u, \lambda_q; \delta u) + C_{u,u}(u; \ \delta u) = 0 \ \forall \delta u \\
\mathcal{R}_{,q}(\chi, u, q; \ \lambda_u, \lambda_q; \delta q) + C_{q,q}(q; \ \delta q) = 0 \ \forall \delta q \\
C_{\chi,\chi}(\chi; \ \delta\chi) + \mathcal{R}_{,\chi}(\chi, u, q; \ \lambda_u, \lambda_q; \delta\chi) \\
+ \int_{\Omega} \left\{ \frac{\chi - \chi_{\text{old}}}{\Delta t}\delta\chi \right. \\
\left. + l_t^2 \frac{\nabla\chi - \nabla\chi_{\text{old}}}{\Delta t} \cdot \nabla\delta\chi \right\} d\Omega = 0 \ \forall \delta\chi
\end{aligned}
\tag{25}
$$

obtained from substituting a backward Euler approximated version of Eq. (24) into Eq. (23).

It is easy to verify that the steady state solution of Eq. (25) corresponds to an exact solution of Eq. (23). Hence, the introduced damping constitutes a consistent regularization in the sense that the exact solution is recovered as the time step Δt
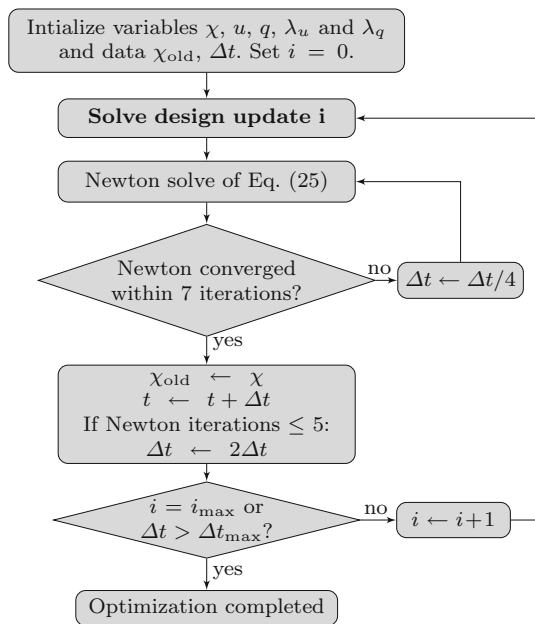
Fig. 8 Flow chart of the adaptive time step solution algorithm

approaches infinity. At the same time, the convex term added to the original optimality equation ensures that, for Δt small enough, the last row of Eq. (25) will not lead to a diverging Newton loop. In that sense, Eq. (25) provides a very convenient way to start the optimization with small pseudo-time increments Δt, that ensure a stable design evolution, and end it with exceedingly large time steps Δt, that recover the quadratic convergence rate of the original Eq. (23) towards its exact optimality solution.

Solution algorithm

Due to its importance for the overall efficiency of the proposed topology optimization framework, the exact adaptive time stepping algorithm for Eq. (25) is documented schematically in Fig. 8. The basic idea is to use the convergence behavior of the Newton loop on Eq. (25) as a measure of the nonlinearity of the system and adapt Δt accordingly. In practice this leads to small time steps when rapid design changes occur and large time steps when design changes are small.

3.2 Application to load–displacement path control

In order to demonstrate the capabilities of the proposed contact modeling approach in topology optimization, the next section deals with a load–displacement path optimization problem. For this reason, the general framework introduced in the previous subsection needs to be applied to this specific optimization problem.

Assuming a set of prescribed displacements $u_D^{\langle j\rangle}$ for $j = 1, 2, \ldots$ in the mechanical residual Eq. (2), the goal is to

find a design field χ that minimizes the deviation between the respective reaction tractions $q^{\langle j\rangle}$ obtained, and a set of corresponding prescribed average reactions $q_*^{\langle j\rangle}$. This can be achieved by minimizing the squared deviation between the actual and desired tractions through an objective function in the form

$$C_q(q) = \sum_j \int_{\Gamma_D} w^{\langle j\rangle} \left\| q^{\langle j\rangle} - q_*^{\langle j\rangle} \right\|^2 d\Gamma, \qquad (26)$$

where $w^{\langle j\rangle}$ are individual weighting factors for each control point. To maintain the compact notation of the previous subsection, all variables $q^{\langle j\rangle}$ are combined in a set $q = \{q^{\langle 1\rangle}, q^{\langle 2\rangle}, \ldots\}$. The corresponding set of displacement fields $u^{\langle j\rangle}$ are combined in a set $u = \{u^{\langle 1\rangle}, u^{\langle 2\rangle}, \ldots\}$.

With this notation, the first row of Eq. (25) can express the mechanical equilibrium at all control points $j = 1, 2, \ldots$, which are otherwise assumed to be independent from each other, i.e. excluding path dependent mechanical responses. Corresponding adjoint problems need to be defined for each control point j. In order to express these in the second and third row of Eq. (25), sets of the involved adjoint variables are defined as $\lambda_u = \{\lambda_u^{\langle 1\rangle}, \lambda_u^{\langle 2\rangle}, \ldots\}$ and $\lambda_q = \{\lambda_q^{\langle 1\rangle}, \lambda_q^{\langle 2\rangle}, \ldots\}$.

So far, there has been no dependence between mechanical equilibrium or adjoint problems for different control points j. This is not true anymore when considering the damped optimality equation in the last row of Eq. (25), where the term

$$\begin{aligned}\mathcal{R}_{,\chi}(\chi, u, q; \lambda_u, \lambda_q; \delta\chi) \\ = \sum_j \mathcal{R}_{,\chi}(\chi, u^{\langle j\rangle}, q^{\langle j\rangle}; \lambda_u^{\langle j\rangle}, \lambda_q^{\langle j\rangle}; \delta\chi)\end{aligned} \qquad (27)$$

couples all subproblems for the various control points involved in the load–displacement path optimization. As a consequence, in order to maintain a second order optimization scheme, Eq. (25) needs to be solved as a coupled system of equations involving all variables at the different control points simultaneously.

Regarding the contribution $C_u(u)$ in the objective function in Eq. (18), there is no need for this kind of contribution in this specific application. Hence

$$C_u(u) = 0. \qquad (28)$$

On the contrary, the design field contribution $C_\chi(\chi)$ in Eq. (18) is essential both for controlling the width of the graded void-solid interface and for enforcing material utilization constraints. By observing Fig. 5a, the width of the void-solid interface for a linearly varying χ from large negative to large positive values is approximately equal to $2/(\rho'(0) \|\nabla\chi\|)$. Enforcing a minimum width of the diffuse

interface equal to l_i can therefore be obtained by penalizing positive values of the quantity

$$\left\langle \rho'(0) \| \nabla \chi \| - 2/l_i \right\rangle, \tag{29}$$

for example by means of a p-norm over the entire design domain. This kind of penalization is essential not only for providing a length-scale for the void-solid interface but also for obtaining a converged design from Eq. (25), with the otherwise unbounded design variable χ.

Based on a dimensionless version of Eq. (29) as well as on an enforced upper bound $\bar{\rho}_{\mathrm{max}}$ to the average density $\bar{\rho}$ over the design domain, i.e. $\bar{\rho} \leq \bar{\rho}_{\mathrm{max}}$, an extended function C_χ is defined as

$$
\begin{aligned}
C_\chi & (\chi, \bar{\rho}, \mu_{\mathrm{avg}}, \mu_{\mathrm{max}}) \\
&= \int_\Omega \left\{ k_i \left\langle \rho'(0) \| \nabla \chi \| l_i/2 - 1 \right\rangle^n \right. \\
&\quad + (\rho(\chi) - \bar{\rho}) \mu_{\mathrm{avg}} \\
&\quad \left. + \frac{k_\rho}{2} \langle \frac{\bar{\rho}}{\bar{\rho}_{\mathrm{max}}} - 1 - \frac{\mu_{\mathrm{max}}}{k_\rho} \rangle^2 - \frac{\mu_{\mathrm{max}}^2}{2k_\rho} \right\} d\Omega.
\end{aligned} \tag{30}
$$

In this expression, k_i is a weighting factor and n is a p-norm exponent for the enforcement of the minimum interface width l_i. The equality of the scalar variable $\bar{\rho}$ to the average of the field ρ over Ω is enforced through the Lagrangian with the Lagrange multiplier μ_{avg}. The inequality constraint $\bar{\rho} \leq \bar{\rho}_{\mathrm{max}}$ is enforced through the augmented Lagrangian with the Lagrange multiplier μ_{max} and the augmentation constant k_ρ. The additional equations necessary for solving for the three new scalar unknowns $\bar{\rho}$, μ_{avg} and μ_{max} are provided by saddle point conditions of the respective Lagrangians, i.e.

$$
\begin{aligned}
C_{\chi, \bar{\rho}}(\bar{\rho}, \mu_{\mathrm{avg}}, \mu_{\mathrm{max}}) &= 0 \\
C_{\chi, \mu_{\mathrm{avg}}}(\chi, \bar{\rho}) &= 0 \\
C_{\chi, \mu_{\mathrm{max}}}(\bar{\rho}, \mu_{\mathrm{max}}) &= 0
\end{aligned} \tag{31}
$$

These three equations need to be appended to Eq. (25), linearized and solved simultaneously with all other problem variables.

With the adaptations presented in the present subsection so far, the application of the framework to load–displacement path control is essentially completed. Nevertheless, for more clarity, the most complex terms involved in the solution of Eq. (25) are presented in more detail. With $C_{u,u}$ vanishing due to Eq. (28), the adjoint equation in the second row of Eq. (25) actually reads

$$
\begin{aligned}
\int_\Omega &\left\{ \mathcal{E}(\chi) \left(\frac{\partial P_{\mathrm{s}}^{\langle j \rangle}}{\partial \nabla u^{\langle j \rangle}} : \nabla \delta u \right) : \nabla \lambda_u^{\langle j \rangle} \right. \\
&+ k_r \mathcal{I}_{\chi_{\mathrm{v}}}(\chi) e^{-5|F_j|} \\
&\left. \left(\mathbb{H} \delta u - 5 |F_j| \left(F_j^{-1} : \nabla \delta u \right) \mathbb{H} u^{\langle j \rangle} \right) : \mathbb{H} \lambda_u^{\langle j \rangle} \right\} d\Omega \\
&+ \int_{\Gamma_D} \delta u \cdot \lambda_q^{\langle j \rangle} \, d\Gamma = 0 \quad \forall j, \delta u
\end{aligned} \tag{32}
$$

where $F_j = F(\nabla u^{\langle j \rangle})$.

It should be noted that applying Newton's method on Eq. (25) requires linearization of Eq. (32) with respect to all involved unknowns, which means that the derivative of the 1st Piola–Kirchhoff stress P with respect to ∇u, appearing in this equation, needs to be differentiated one additional time with respect to ∇u. This is easy for the adopted hyperelastic material law with the resulting expression for P shown in Eq. (5), but this is not in general the case for any constitutive law. The term in Eq. (32) due to the proposed void regularization, weighted with k_r, is also straightforward to linearize with respect to ∇u.

For this specific application, the adjoint equation in the third row of Eq. (25) becomes

$$
\begin{aligned}
\int_{\Gamma_D} &\left\{ \delta q \cdot \lambda_u^{\langle j \rangle} \right. \\
&\left. + 2 w^{\langle j \rangle} \left(q^{\langle j \rangle} - q_*^{\langle j \rangle} \right) \cdot \delta q \right\} d\Gamma = 0 \quad \forall j, \delta q
\end{aligned} \tag{33}
$$

Finally, the damped optimality equation for a design update from a previous design χ_{old}, appearing last in Eq. (25), is particularized for this application as

$$
\begin{aligned}
\int_\Omega &\left\{ n \left\langle \rho'(0) \| \nabla \chi \| \frac{l_i}{2} - 1 \right\rangle^{n-1} \rho'(0) \frac{\nabla \chi}{\| \nabla \chi \|} \cdot \nabla \delta \chi \right. \\
&+ \rho'(\chi) \mu_{\mathrm{avg}} \delta \chi \\
&+ \sum_j \left(\mathcal{E}'(\chi) P_{\mathrm{s}}(\nabla u^{\langle j \rangle}) : \nabla \lambda_u^{\langle j \rangle} \right. \\
&\quad \left. + k_r \mathcal{I}_{\chi_{\mathrm{v}}}'(\chi) e^{-5|F(\nabla u^{\langle j \rangle})|} \mathbb{H} u^{\langle j \rangle} : \mathbb{H} \lambda_u^{\langle j \rangle} \right) \delta \chi \\
&+ \frac{\chi - \chi_{\mathrm{old}}}{\Delta t} \delta \chi \\
&\left. + l_t^2 \frac{\nabla \chi - \nabla \chi_{\mathrm{old}}}{\Delta t} \cdot \nabla \delta \chi \right\} d\Omega = 0 \quad \forall \delta \chi.
\end{aligned} \tag{34}
$$

Despite some lengthy expressions, all involved terms are differentiable also in this last equation, so that application of Newton's method with a consistent tangent matrix is possible. The model is implemented in the automated finite element framework GetFEM [26], which performs necessary symbolic linearizations and the numerical assembly of the discretized system of equations.
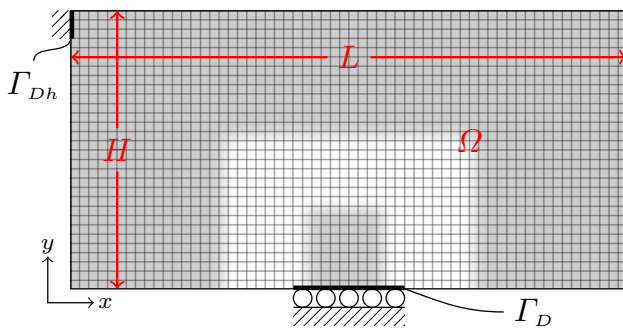
Fig. 9 Problem domain for designing a stiff coupling with initial clearance. Homogeneous Dirichlet condition region Γ_{Dh} has a length $0.1H$ and the nonhomogeneous Dirichlet condition region Γ_D spans from $0.4L$ to $0.6L$ along the bottom edge. Dimensions: $L = 52$, $H = 26$

Table 2 Model parameters for the optimization example

Domain dimension	$L \times H$	52×26	mm^2
Mesh size		60×30	–
Solid bulk modulus	K_s	$5/3$	MPa
Solid shear modulus	G_s	$5/14$	MPa
Void/solid stiffness contrast	$\mathcal{E}_0$	10^{-12}	–
Void regularization scaling	$\bar{k}_r$	10^{-6}	–
Max. material vol. fraction	$\bar{\rho}_{\max}$	0.35	–
Vol. constr. augm. param.	k_ρ	10^{-3}	–
Control point weights	$w^{\langle 1 \rangle}/w^{\langle 2 \rangle}$	$10^4/10^3$	MPa^{-2}
Target tractions	$q_{*x}^{\langle 1 \rangle}/q_{*x}^{\langle 2 \rangle}$	$0 / 0.1$	MPa
Transient length scale	l_t	2	mm
Minimum interface width	l_i	2	mm
Interf. width p-norm weight	k_i	10^{-1}	–
Interf. width p-norm exponent	n	6	–

4 Results and discussion

The internal contact enabled topology optimization formulation for load–displacement path control, described in the previous section, is applied now to design a structure that switches from a completely decoupled to a high stiffness connection, within a given displacement interval. Figure 9 shows the problem domain, along with the initial material density distribution and a discretization with 60×30 square elements. A small region Γ_{Dh}, near the top of the left side of the domain, is fixed, while a prescribed horizontal displacement is imposed on a region Γ_D in the middle of the bottom side of the domain. The path control problem to be solved consists in minimizing the reaction force on Γ_D for a horizontal displacement $u_{Dx}^{\langle 1 \rangle} = 0.1L$ and at the same time maximize the reaction force for $u_{Dx}^{\langle 2 \rangle} = 0.15L$.

Before applying the method, a suitable discretization needs to be specified for all unknown fields. The displacement fields $u^{\langle j \rangle}$ are approximated with 8-node quadratic elements. All degrees of freedom within the Γ_{Dh} region and all vertical degrees of freedom within the Γ_D region are removed from the finite element space. Moreover, all horizontal degrees of freedom within Γ_D are reduced to a single master degree of freedom, enforcing a rigid body translation of this boundary. The finite element space resulting after these restrictions is used for $u^{\langle j \rangle}$, the corresponding adjoint fields $\lambda_u^{\langle j \rangle}$ and variations δu. Linear 4-node elements are used for approximating the design field χ and the corresponding variations $\delta \chi$ without any further restrictions. Since χ is unbounded, it does not suffer by the usual limitation to linear elements when approximating the density field ρ directly, in order to ensure density values in the [0, 1] interval. Nevertheless, numerical experiments with quadratic elements for χ showed an only relatively small accuracy gain. For this choice of finite element spaces for the different fields, all terms involved in the the problem domain Ω are assembled with nine Gauss points in each quadrilateral element. Terms
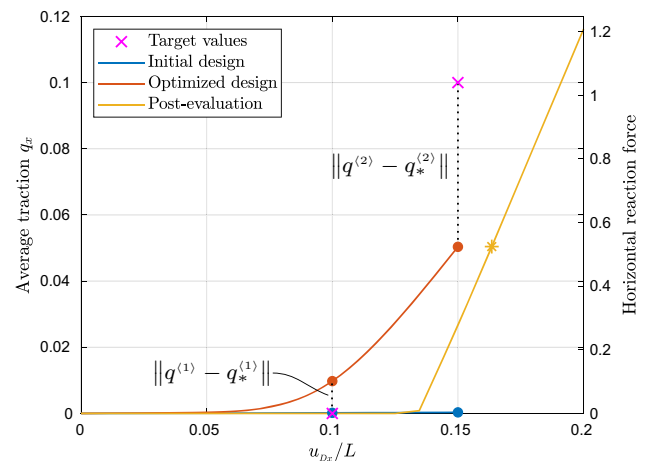

Fig. 10 Load–displacement response of the initial and optimized design, together with a post-evaluation analysis with a body fitted mesh. The colored dots mark traction values at the control points and the asterisk indicates the state shown in Fig. 13c

defined on the boundary Γ_D are assembled with the corresponding three Gauss points per element edge.

Since vertical displacements on Γ_D are already fixed by restricting the relevant finite element space, the unknown multipliers $q^{\langle j \rangle}$ on Γ_D are reduced to just their x component, written in the form $q^{\langle j \rangle} = \{q_x^{\langle j \rangle}, 0\}^T$. The same format applies to the corresponding adjoint variables $\lambda_q^{\langle j \rangle}$, variations δq, and target reactions $q_*^{\langle j \rangle}$, which also reduce to scalars.

Based on this discretization, the optimization procedure from the previous section is applied with a volume constraint $\bar{\rho}_{\max}$ of 35% and two control points $j = \{1, 2\}$, illustrated in Fig. 10. In order to minimize the initial stiffness of the coupling, a target average traction $q_{*x}^{\langle 1 \rangle} = 0$ is set for the first control point at $u_{Dx}^{\langle 1 \rangle} = 0.1L$. Requiring a stiff coupling upon further displacement of the moving part is represented
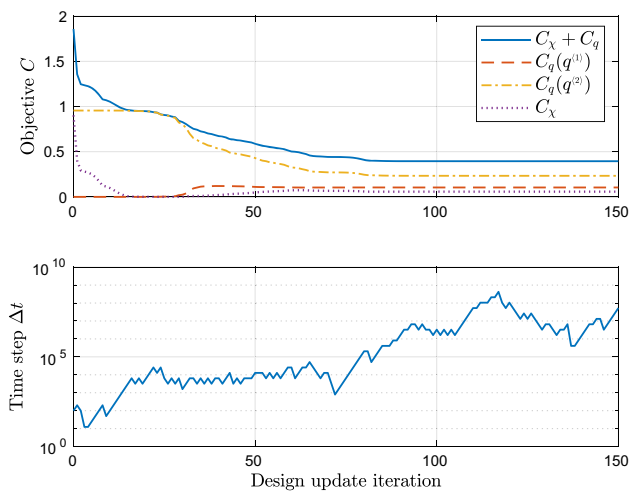
Fig. 11 Evolution of objective value and design update time step. The total objective value $C = C_\chi + C_q$ and the contributions of the two control points to C_q are shown separately
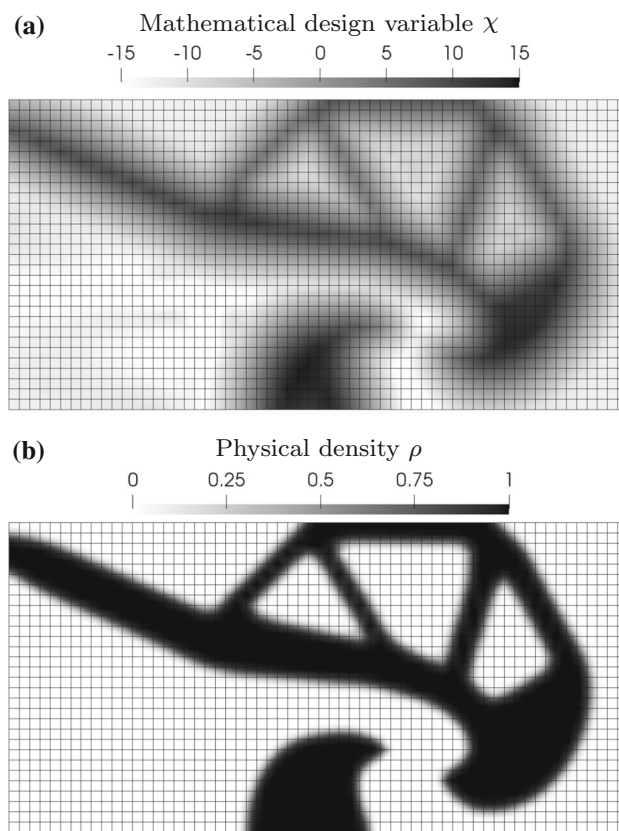

Fig. 12 Optimization results after 150 iterations. The mathematical design variable field χ, shown in **a**, is interpolated with the bi-linear finite elements chosen for this field, while the physical density field ρ, shown in **b** and actually used in the mechanical simulation, is obtained through Eq. (11) without any further post-processing.

by a high target value $q_{*x}^{\langle 2 \rangle} = 0.1$ for the second control point at $u_{Dx}^{\langle 2 \rangle} = 0.15L$. The corresponding weightings $w^{\langle j \rangle}$ for the objective function C_q according to Eq. (26), are provided in Table 2. With this kind of objective C_q, a design is expected where contact onset occurs between the two control points. All further parameters for this optimization example are reported in Table 2.

The initial design, shown in Fig. 9, consists of two disconnected regions with material density $\rho = 0.2$, separated by a stripe of void material with $\rho = 10^{-5}$. Other non-uniform starting guesses like this one result in similar albeit shifted designs, whereas completely uniform initial designs lack of adequate initial non-linearity to guide the optimization towards a contact dominated structure.

After 150 design evolution iterations through the objective history shown in Fig. 11, the optimization procedure converges to the design shown in Fig. 12b, in terms of the physical density ρ. The corresponding field of the underlying mathematical design variable χ is shown in Fig. 12a. The performance of the initial and the optimized designs with respect to the set objective can be evaluated based on the corresponding load–displacement curves, shown in Fig. 10. The initial design is very compliant satisfying basically only the zero reaction force required at the first control point but being far from the objective at the second control point. The optimized design exhibits a small but non-zero reaction force at the first control point and a large reaction force at the second control point, which is nevertheless still below the target traction $q_{*x}^{\langle 2 \rangle}$. This solution represents a compromise between the two competing requirements at the two control points. Unless a design exists that can satisfy both target values, a compromise is to be expected. The larger, by an order of magnitude, weighting $w^{\langle 1 \rangle}$ for the first control point, compared

to $w^{\langle 2 \rangle}$, results in a smaller deviation from the reaction-free objective at the first point compared to the deviation at the second control point.

As expected, the transition from the approximately decoupled to the stiff response is rather smooth and thus there is no distinct contact point. This is a consequence of the remaining stiffness in the void between the contact surfaces, and especially in the graded interface, as already has been shown for the C-shape example from Figs. 6 and 7. This parasitic interaction, however, is what makes the contact problem differentiable, and therefore more suitable for optimization.

After the first approximately 20 iterations, spent on satisfying the minimum interface width l_i constraint (contained in C_χ), which is not respected in the initial guess, the objective function C_q for the two control points dominates the total objective function C. At that point, the objective of the second control point starts to decrease at the cost of an increasing objective for the first control point. During the whole optimization procedure, the pseudo-time step for each design update increases by several orders of magnitude. Note that the larger the value of Δt, the closer Eq. (25) is to the
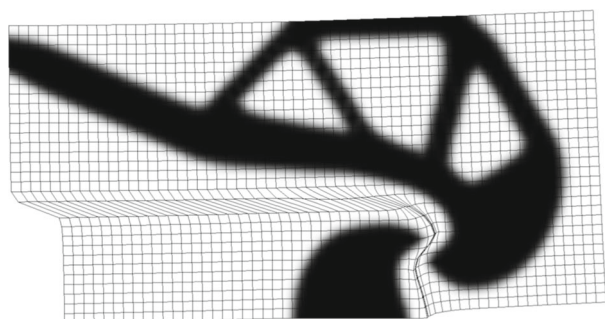
(a) Deformed structure at control point 1, $u_{Dx} = 0.1L$.
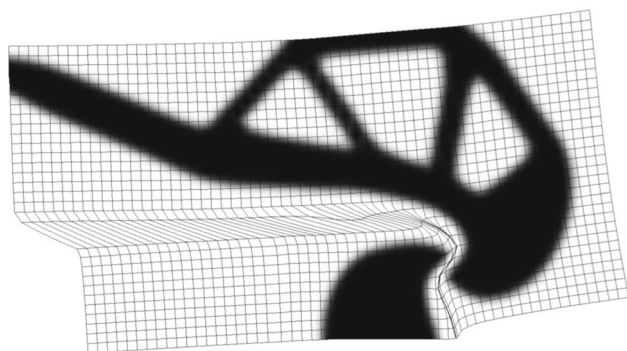
(b) Deformed structure at control point 2, $u_{Dx} = 0.15L$.
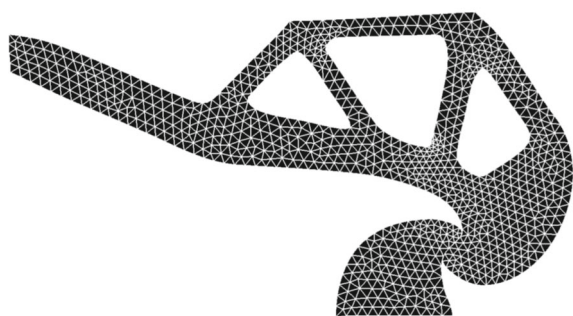
(c) Deformed post-evaluation structure at $u_{Dx} \approx 0.17L$.

Fig. 13 Deformations of the optimized structure at the two control points (**a**, **b**). Deformations from post-evaluation with a body fitted mesh (**c**), based on a discrete solid/void interface at $\rho = 0.5$, at the same load as for the optimized structure at control point 2 (marked with an asterisk in Fig. 10)

exact (local) optimum from Eq. (23). The ultimately obtained design contains a truss like structure, typical for compliance minimization, and a hook like coupling, which involves third medium contact. Figure 13b, c show the optimized structure in its deformed state at the first and second control point, respectively. It can be seen both in Figs. 10 and 13b, that third medium contact interaction actually occurs prior to the first control point. Nevertheless, a very clear transition can be observed, in general, from a practically zero initial stiffness to an ultimately stiff post-contact response.

It should be noted that, as the load is increased, the slope of the contact interface in the obtained hook geometry changes

due to the deformation of the structure. Ultimately, at a sufficiently high load, this geometrical change will lead to sliding between the two parts and disengagement. The large deformation aware optimization procedure, employed in the present work, intrinsically accounts for this effect, compensating for such load-dependent geometrical changes to prevent a disengagement, at least until the second control point. This is obviously a design characteristic which can only be obtained through a large deformation consistent modeling of both solid and contact mechanics, as applied here.

Post-evaluation with body fitted mesh

The design obtained exhibits a graded solid-void interface of constant width that is perfectly consistent with the prescribed value l_i. Nevertheless, the finite size of the interface is the major source of inaccuracy in the modeling of contact. For this reason a post-evaluation and comparison with a body fitted mesh is highly relevant.

In order to investigate the behavior of a corresponding discrete void and solid structure, which is actually manufacturable, a body fitted mesh has been generated with GMSH [27], based on the interpretation of the contour line $\rho = 0.5$, i.e. $\chi = 0$, as the external surface of the solid. This body fitted mesh is shown in Fig. 13c, in its undeformed state, and in Fig. 13c, in its deformed state under the same load as the graded density structure at the second control point. Triangular 6-node elements and the Lagrange multiplier based contact method from [22] were used for the simulations with this body fitted mesh. The corresponding load–displacement curve is shown in Fig. 10 alongside the one from the optimized graded density design.

The post-evaluation curve reveals how delayed the onset of contact actually occurs, being closer to the second control point rather than the first one. This discrepancy is due to the presence of the graded density layer between the contacting solids in the graded density structure, seen in Figures 13a and 13b. On the other hand, the post-evaluation confirms a high stiffness of the coupling after onset of contact, which is actually even higher than predicted by the graded density representation, again due to a compliant layer of intermediate density material between the solids in the latter case. Both effects observed are mainly caused by the remaining stiffness inside the graded solid-void interface, and as such, they are expected to decrease with finer discretization and smaller minimum interface width l_i.

All in all, the example presented clearly demonstrates that the proposed method is capable of producing valuable structural designs that exploit internal contact. These can either be used as input to shape optimization methods or be further optimized within the same framework in a multigrid scheme with subsequent mesh refinements and decreased interface width l_i.
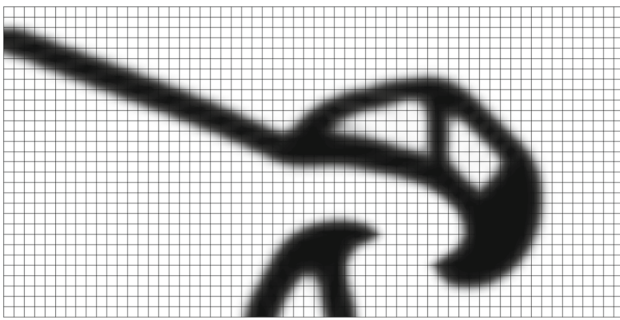
Fig. 14 Optimized design (converged ρ field after 150 iterations) for a reduced volume fraction $\bar{\rho}_{max} = 0.2$ and an accordingly reduced target reaction $q_*^{\langle 2 \rangle} = 0.05\,\text{MPa}$ for the second control point.
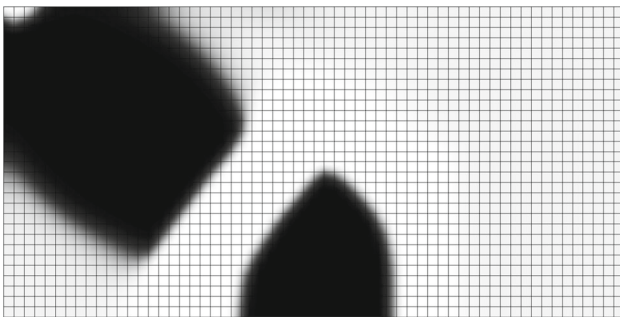

Fig. 15 Optimized design (converged ρ field after 300 iterations) for reversed loading direction, i.e. $u_{Dx}^{\langle 1 \rangle} = -0.1\,\text{L}$, $u_{Dx}^{\langle 2 \rangle} = -0.2\,\text{L}$ and $q_*^{\langle 2 \rangle} = -0.1\,\text{MPa}$, and the original volume fraction $\bar{\rho}_{max} = 0.35$

Additional examples

To illustrate the generality of the proposed method, two variations of the previous example are presented in Figs. 14 and 15. For the first variation, the same properties as in Table 2 are used, except for a stricter upper limit on the material usage of $\bar{\rho}_{max} = 0.2$ and a lower target traction value at the second control point $q_*^{\langle 2 \rangle} = 0.05\,\text{MPa}$, to account for the reduced volume fraction. The optimization converged towards the design shown in Fig. 14 with similar but thinner features as in the original example.

For the second variation, the loading is reversed by prescribing negative displacements $u_{Dx}^{\langle 1 \rangle} = -0.1\,\text{L}$ and $u_{Dx}^{\langle 2 \rangle} = -0.2\,\text{L}$, and desired traction $q_*^{\langle 2 \rangle} = -0.1\,\text{MPa}$, while maintaining all further parameters from Table 2. Figure 15 shows the optimized design for this variation, after 150 iterations, consisting of two solid blocks, separated by a straight contact interface with orientation perpendicular to the load path between Γ_{Dh} and Γ_D.

Computational performance

For the strongly nonlinear optimization examples of the present work, the computational cost of a single design iter-

ation does not only depend on the number of degrees of freedom of the system, but also on the number of required Newton iterations. The occurrence of contact does not affect the computational cost for a single Newton iteration. Nevertheless, the increased nonlinearity due to contact is likely to lead to more frequent time step adaptations according to the scheme of Fig. 8. This will affect the average number of Newton iterations performed per design update. In quantitative terms, the optimization shown in Fig. 11 for the first example of this section, with 46,294 degrees of freedom in total, required 26 s of cpu time per design update on average, on a single Xeon E5-2660 v3 processor. The implementation of the example was done in Python with all computational intensive operations, like vector and matrix assemblies performed in the underlying C++ library GetFEM [26], and linear system solutions performed in the Fortran direct solver MUMPS, [28].

5 Conclusion

A method for indirect internal contact in finite strain topology optimization has been presented, based on the third medium contact approach. Void regions, which are inherently present in density based topology optimization, were exploited as the contact medium, circumventing the need for defining contact surfaces and detecting contact. To this end, the present work has introduced a new void material regularization which allowed to approximate frictionless contact between structures with discrete or graded solid-void interface, accommodating considerable amounts of sliding.

In general, the usefulness of including a simple and stable contact formulation in topology optimization is twofold. It enables the optimization procedure to utilize contact as a design feature, as the example shown in the present work, but it also allows to account for collisions, in cases where contact should be avoided. Hereby it opens up for numerous potential applications in the design of compliant mechanisms, microstructures of periodic material with extremal mechanical properties and energy absorbing structures.

The present work has also assessed the limitations of the proposed approach. The main source of discrepancies between the proposed third medium contact and graded density design method, and a body fitted mesh based post evaluation, was found to be related to the finite width of the solid-void interface layer. Possible mitigations have also been discussed.

Besides the contact modeling contribution, this work has also introduced a new general framework for finite strain topology optimization. The proposed framework is expressed entirely in weak form in a discretization agnostic manner. It is based on the idea of a simultaneous solving of the fully coupled nonlinear equations for mechanical equilibrium, adjoint

analysis, and design optimality, providing a second order consistent optimization scheme. Moreover, it includes a new method for imposing a length scale for the graded solid-void interface, which is very essential for third medium contact modeling, as explained above. The potential of this framework has been demonstrated by the optimization examples included, and more in depth studies shall follow in future work.

References

1. Kumar P, Saxena A, Sauer RA (2019) Computational synthesis of large deformation compliant mechanisms undergoing self and mutual contact. J Mech Des Trans ASME 141(1):012302. https://doi.org/10.1115/1.4041054
2. Wriggers P, Schrder J, Schwarz A (2013) A finite element method for contact using a third medium. Comput Mech 52(4):837–847. https://doi.org/10.1007/s00466-013-0848-5
3. Bog T, Zander N, Kollmannsberger S, Rank E (2015) Normal contact with high order finite elements and a fictitious contact material. Comput Math Appl 70(7):1370–1390. https://doi.org/10.1016/j.camwa.2015.04.020
4. Oliver J, Hartmann S, Cante J, Weyler R, Hernndez J (2009) A contact domain method for large deformation frictional contact problems. Part 1: theoretical basis. Comput Methods Appl Mech Eng 198(33):2591–2606. https://doi.org/10.1016/j.cma.2009.03.006
5. Simo J, Laursen T (1992) An augmented Lagrangian treatment of contact problems involving friction. Comput Struct 42(1):97–116. https://doi.org/10.1016/0045-7949(92)90540-G
6. Gitterle M, Popp A, Gee MW, Wall WA (2010) Finite deformation frictional mortar contact using a semi-smooth newton method with consistent linearization. Int J Numer Methods Eng 84(5):543–571. https://doi.org/10.1002/nme.2907
7. Hilding D, Klarbring A, Petersson J (1999) Optimization of structures in unilateral contact. Appl Mech Rev 52(4):139–160. https://doi.org/10.1115/1.3098931
8. Strmberg N (2010) Topology optimization of structures with manufacturing and unilateral contact constraints by minimizing an adjustable compliance-volume product. Struct Multidiscip Optim 42(3):341–350. https://doi.org/10.1007/s00158-010-0502-1
9. Strömberg N, Klarbring A (2010) Topology optimization of structures in unilateral contact. Struct Multidiscip Optim 41(1):57–64. https://doi.org/10.1007/s00158-009-0407-z
10. Strömberg N (2013) The influence of sliding friction on optimal topologies. Springer, Berlin, pp 327–336. https://doi.org/10.1007/978-3-642-33968-4_20
11. Kristiansen H, Poulios K, Aage N (2020) Topology optimization for compliance and contact pressure distribution in structural problems with friction. Comput Methods Appl Mech Eng 364:112915. https://doi.org/10.1016/j.cma.2020.112915
12. Luo Y, Li M, Kang Z (2016) Topology optimization of hyperelastic structures with frictionless contact supports. Int J Solids Struct 81:373–382. https://doi.org/10.1016/j.ijsolstr.2015.12.018
13. Fernandez F, Puso MA, Solberg J, Tortorelli DA (2020) Topology optimization of multiple deformable bodies in contact with large deformations. Comput Methods Appl Mech Eng 371:113288. https://doi.org/10.1016/j.cma.2020.113288
14. Desmorat B (2007) Structural rigidity optimization with frictionless unilateral contact. Int J Solids Struct 44(3):1132–1144. https://doi.org/10.1016/j.ijsolstr.2006.06.010
15. Niu C, Zhang W, Gao T (2020) Topology optimization of elastic contact problems with friction using efficient adjoint sensitivity analysis with load increment reduction. Comput Struct 238:106296. https://doi.org/10.1016/j.compstruc.2020.106296
16. Lawry M, Maute K (2015) Level set topology optimization of problems with sliding contact interfaces. Struct Multidiscip Optim 52(6):1107–1119. https://doi.org/10.1007/s00158-015-1301-5
17. Yoon GH, Kim YY (2005) Element connectivity parameterization for topology optimization of geometrically nonlinear structures. Int J Solids Struct 42(7):1983–2009. https://doi.org/10.1016/j.ijsolstr.2004.09.005
18. Wang F, Lazarov BS, Sigmund O, Jensen JS (2014) Interpolation scheme for fictitious domain techniques and topology optimization of finite strain elastic problems. Comput Methods Appl Mech Eng 276:453–472. https://doi.org/10.1016/j.cma.2014.03.021
19. Wallin M, Ristinmaa M (2015) Finite strain topology optimization based on phase-field regularization. Struct Multidiscip Optim 51(2):305–317. https://doi.org/10.1007/s00158-014-1141-8
20. Cho S, Jung HS (2003) Design sensitivity analysis and topology optimization of displacement-loaded non-linear structures. Comput Methods Appl Mech Eng 192(22–23):2539–2553. https://doi.org/10.1016/S0045-7825(03)00274-3
21. Simo JC, Taylor RL, Pister KS (1985) Variational and projection methods for the volume constraint in finite deformation elastoplasticity. Comput Methods Appl Mech Eng 51(1-3):177–208, 177–208. https://doi.org/10.1016/0045-7825(85)90033-7
22. Poulios K, Renard Y (2015) An unconstrained integral approximation of large sliding frictional contact between deformable solids. Comput Struct 153:75–90. https://doi.org/10.1016/j.compstruc.2015.02.027
23. Stolpe M, Svanberg K (2001) An alternative interpolation scheme for minimum compliance topology optimization. Struct Multidiscip Optim 22(2):116–124. https://doi.org/10.1007/s001580100129
24. Ebert DS, Musgrave FK, Peachey D, Perlin K, Worley S, Mark WR, Hart JC (2003) Texturing and modeling: a procedural approach, 3rd edn. Elsevier Inc, Amsterdam
25. Klarbring A, Torstenfelt B (2009) Ode approach to topology optimization
26. Renard Y, Poulios K (2020) Getfem: Automated FE modeling of multiphysics problems based on a generic weak form language. ACM Trans Math Softw 47(1). https://doi.org/10.1145/3412849
27. Geuzaine C, Remacle JF (2009) Gmsh: a 3-D finite element mesh generator with built-in pre- and post-processing facilities. Int J Numer Methods Eng 79(11):1309–1331. https://doi.org/10.1002/nme.2579
28. Amestoy P, Duff I, L'Excellent JY (2000) Multifrontal parallel distributed symmetric and unsymmetric solvers. Comput Methods Appl Mech Eng 14(2):501–520. https://doi.org/10.1016/S0045-7825(99)