

CORINNA KLAPPROTH

The Contact-Stabilized Newmark Method – Consistency Error of a Spatiotemporal Discretization¹

¹Supported by the DFG Research Center MATHEON, “Mathematics for key technologies: Modelling, simulation, and optimization of real-world processes”, Berlin

The Contact-Stabilized Newmark Method – Consistency Error of a Spatiotemporal Discretization[†]

Corinna Klapproth

Abstract

The paper considers an improved variant of the contact-stabilized Newmark method by Deuffhard et al., which provides a spatiotemporal numerical integration of dynamical contact problems between viscoelastic bodies in the frame of the Signorini condition. Up to now, the question of consistency in the case of contact constraints has been discussed for time integrators in function space under the assumption of bounded total variation of the solution. Here, interest focuses on the consistency error of the Newmark scheme in physical energy norm after discretization both in time and in space. The resulting estimate for the local discretization error allows to prove global convergence of the Newmark scheme under an additional assumption on the active contact boundaries.

AMS MSC 2000: 35L85, 65M15, 74H15, 74M15

Keywords: dynamical contact problems, contact-stabilized Newmark method, consistency

1 Introduction

Dynamical contact problems play an important role in different application areas such as structural mechanics and biomechanics. Typically, they are modelled via Signorini’s classical contact conditions, which lead to nonlinear variational inequalities that are highly nonsmooth at contact interfaces between bodies.

In view of the numerical treatment of dynamical contact problems, a reliable integrator should mainly achieve two properties: a tight energy conservation or at least dissipativity and the avoidance of numerical instabilities during phases of active contact. Since the most wide-spread algorithms in

[†]Supported by the DFG Research Center MATHEON, “Mathematics for key technologies: Modelling, simulation, and optimization of real-world processes”, Berlin

engineering do not satisfy these demands, a large variety of alternative integrators has been designed in the last years. The fundamental difficulty is to adequately cope with the interaction of space and time discretization, which has turned out to be the main cause for the spurious oscillations in many approaches. In this paper, an improved variant of the contact-stabilized Newmark method suggested by Deuffhard, Krause, and Ertel [3] is considered. Compared to the original Newmark scheme with contact-stabilization, the method is not only energy dissipative and free of any artificial oscillations, but also produces velocities equal to zero at active contact boundaries.

In the unconstrained situation, the Newmark method is well-known to have consistency and convergence order two with respect to the timestep size (see, e.g., the textbook [9]). In the constrained situation, the intrinsic discontinuities at contact interfaces make the question of consistency and convergence much more difficult. As a result, a consistency result under contact constraints necessitates a novel regularity assumption on the solution, which has only recently been addressed in [16] for the first time. The authors presented a local discretization error analysis within the physical energy norm studied in [14], where the solution and its derivatives are required to be of bounded total variation. Furthermore, a novel proof technique for convergence has been developed in [13], which allows to show that the global discretization error of the scheme tends to zero with order $1/2$.

However, the consistency and convergence theory presented in these publications has been performed within the method of time layers, i.e. for the corresponding contact-stabilized Newmark scheme in function space. In this framework, the effect of the contact-stabilization completely vanishes. Hence, its benefits in space-and-time discretization are not reflected by the results in [16, 13]. For this reason, the present paper fills this gap by analyzing the consistency and convergence error of the improved contact-stabilized Newmark method after discretization both in time and in space.

The paper will start with the mathematical formulation of the underlying dynamical contact problem in Section 2. In Section 3, an improved variant of the contact-stabilized Newmark method by Deuffhard et al. will be presented and its conservation and stabilization features will be discussed. After a review of existing consistency results within the method of time layers, the central Section 4 will contain the novel consistency result for the improved contact-stabilized Newmark method with discretization both in time and in space. The section will be concluded by a discussion of both the consistency and the convergence behavior of the spatiotemporal integrator.

2 Notation and Background

The paper at hand is concerned with a contact-stabilized Newmark method for dynamical contact problems between two bodies. In this initial section,

the classical problem formulation based on linearized Signorini's contact conditions is written down in order to set notations and definitions. In view of existing perturbation and consistency results [15, 16], the model utilizes linear viscoelastic materials fulfilling the Kelvin-Voigt constitutive law.

Notation. Let the two bodies be identified with the union of two domains which are understood to be bounded subsets in \mathbb{R}^d with $d = 2, 3$. Each of the boundaries are assumed to be Lipschitz and decomposed into three disjoint parts: Γ_D , the Dirichlet boundary, Γ_N , the Neumann boundary, and Γ_C , the possible contact boundary. The actual contact boundary is not known in advance, but is assumed to be contained in a subset of Γ_C . The Dirichlet boundary conditions give rise to $\mathbf{H}_D^1 := \{\mathbf{v} \in \mathbf{H}^1 \mid \mathbf{v}|_{\Gamma_D} = 0\}$.

For given Banach space \mathbf{V} and time interval $t_0 < T < \infty$, let $C([t_0, T], \mathbf{V})$ be the continuous functions $\mathbf{v} : [t_0, T] \rightarrow \mathbf{V}$. The space $\mathbf{L}_2(t_0, T; \mathbf{V})$ consists of all measurable functions $\mathbf{v} : (t_0, T) \rightarrow \mathbf{V}$ for which $\|\mathbf{v}\|_{\mathbf{L}_2(t_0, T; \mathbf{V})}^2 := \int_{t_0}^T \|\mathbf{v}(t)\|_{\mathbf{V}}^2 dt < \infty$ holds. The space \mathbf{L}_2 is identified with its dual space, and this yields the evolution triple $\mathbf{H}^1 \subset \mathbf{L}_2 \subset (\mathbf{H}^1)^*$ where $(\mathbf{H}^1)^*$ denotes the dual space of \mathbf{H}^1 . With reference to this evolution triple, the Sobolev space $\mathbf{W}^{1,2}(t_0, T; \mathbf{H}^1, \mathbf{L}_2)$ means the set of all functions $\mathbf{v} \in \mathbf{L}_2(t_0, T; \mathbf{H}^1)$ that have generalized derivatives $\dot{\mathbf{v}} \in \mathbf{L}_2(t_0, T; (\mathbf{H}^1)^*)$, see, e.g., [28].

For the sake of clear arrangement, the abbreviation $\bar{\mathbf{v}} = (\mathbf{v}, \dot{\mathbf{v}})$ will be used for a function and its first time derivative.

Non-penetration condition. At the contact interface Γ_C , the two bodies may come into contact but must not penetrate each other. Assuming a bijective mapping $\phi : \Gamma_C^S \rightarrow \Gamma_C^M$ between the two possible contact surfaces to be given, linearized non-penetration can be defined with respect to ϕ by [6]

$$[\mathbf{u} \cdot \boldsymbol{\nu}]_\phi(x, t) = \mathbf{u}^S(x, t) \cdot \boldsymbol{\nu}_\phi(x) - \mathbf{u}^M(\phi(x), t) \cdot \boldsymbol{\nu}_\phi(x) \leq g(x), \quad x \in \Gamma_C^S. \quad (1)$$

This condition is given with respect to the initial gap

$$\Gamma_C^S \ni x \mapsto g(x) = |x - \phi(x)| \in \mathbb{R} \quad (2)$$

between the two bodies in the reference configuration and the normalized vector

$$\boldsymbol{\nu}_\phi = \begin{cases} \frac{\phi(x) - x}{|\phi(x) - x|}, & \text{if } x \neq \phi(x), \\ \boldsymbol{\mu}^S(x) = -\boldsymbol{\mu}^M(x), & \text{if } x = \phi(x). \end{cases} \quad (3)$$

Variational problem formulation. For the weak formulation of the dynamical contact problem, the convex set of all admissible displacements is denoted by

$$\mathcal{K} = \{\mathbf{v} \in \mathbf{H}_D^1 \mid [\mathbf{v} \cdot \boldsymbol{\nu}]_\phi \leq g\}. \quad (4)$$

The materials under consideration are assumed to be linearly viscoelastic, i.e. the stresses satisfy the Kelvin-Voigt constitutive relation. Both elasticity and viscoelasticity tensors should be sufficiently smooth, symmetric, and uniformly positive definite. In this case, the internal forces can be written as a bilinear form a in \mathbf{H}^1 for the linearly elastic part, respectively b for the viscous part. Both bilinear forms are bounded in \mathbf{H}^1 and give rise to seminorms $\|\cdot\|_a^2 = a(\cdot, \cdot)$ and $\|\cdot\|_b^2 = b(\cdot, \cdot)$.

The external forces are represented by a linear functional f_{ext} on \mathbf{H}_D^1 which accounts for the volume forces and the tractions on the Neumann boundary. The sum of internal elastic and external forces can be represented by

$$\langle \mathbf{F}(\mathbf{w}), \mathbf{v} \rangle_{(\mathbf{H}^1)^* \times \mathbf{H}^1} = a(\mathbf{w}, \mathbf{v}) - f_{\text{ext}}(\mathbf{v}), \quad \mathbf{v}, \mathbf{w} \in \mathbf{H}^1, \quad (5)$$

and the viscoelastic forces can be written as

$$\langle \mathbf{G}(\mathbf{w}), \mathbf{v} \rangle_{(\mathbf{H}^1)^* \times \mathbf{H}^1} = b(\mathbf{w}, \mathbf{v}), \quad \mathbf{v}, \mathbf{w} \in \mathbf{H}^1. \quad (6)$$

In the weak formulation [5, 12], the dynamical contact problem can be written as the following variational inequality: For almost every $t \in [0, T]$, find $\mathbf{u} \in \mathcal{K}$ with $\mathbf{u}(\cdot, t) \in C([0, T], \mathbf{H}^1)$ and $\dot{\mathbf{u}} \in \mathbf{W}^{1,2}(0, T; \mathbf{H}^1, \mathbf{L}_2)$ such that for all $\mathbf{v} \in \mathcal{K}$

$$\langle \ddot{\mathbf{u}}, \mathbf{v} - \mathbf{u} \rangle_{(\mathbf{H}^1)^* \times \mathbf{H}^1} + \langle \mathbf{F}(\mathbf{u}), \mathbf{v} - \mathbf{u} \rangle_{(\mathbf{H}^1)^* \times \mathbf{H}^1} + \langle \mathbf{G}(\dot{\mathbf{u}}), \mathbf{v} - \mathbf{u} \rangle_{(\mathbf{H}^1)^* \times \mathbf{H}^1} \geq 0 \quad (7)$$

and

$$\mathbf{u}(0) = \mathbf{u}_0, \quad \dot{\mathbf{u}}(0) = \dot{\mathbf{u}}_0. \quad (8)$$

Incorporating the constraints $\mathbf{v}(t) \in \mathcal{K}$ for almost every $t \in [0, T]$ by the characteristic functional $I_{\mathcal{K}}(\mathbf{v})$, the variational inequality (7) can equivalently be formulated as the variational inclusion

$$0 \in \ddot{\mathbf{u}} + \mathbf{F}(\mathbf{u}) + \mathbf{G}(\dot{\mathbf{u}}) + \partial I_{\mathcal{K}}(\mathbf{u}) \quad (9)$$

utilizing the subdifferential $\partial I_{\mathcal{K}}$ of $I_{\mathcal{K}}$ [7]. As shown for instance in [2], the unilateral contact problem between a viscoelastic body and a rigid foundation has at least one weak solution.

3 The Contact-Stabilized Newmark Method

In the community of computational mechanics, the most popular family of time discretization schemes are the classical Newmark methods [24]. Due to their excellent characteristics in the absence of contact constraints, these methods are often transferred into time-stepping schemes for solving problems from contact mechanics. However, the classical handling of the additional non-penetration condition causes an uncontrollable behavior of the total energy during time integration. Moreover, the methods evoke spurious

instabilities at dynamical contact boundaries, which show up as unwanted oscillations in displacements, velocities, and contact stresses [3, 21, 13].

In the last years, several variants of the classical Newmark method for dynamical contact problems have been designed to avoid these deficits. In order to overcome the poor energy conservation, Kane, Repetto, Ortiz, and Marsden [11] developed an energy-dissipative version of the scheme by proposing a fully implicit treatment of the contact forces. In 2007, Deuffhard, Krause, and Ertel [3] designed a contact-stabilized variant of this algorithm, which completely removes the spurious oscillations at contact boundaries and is still energy dissipative in the presence of contact. Both methods have originally been formulated in pure linear elasticity, but they have been generalized to the viscoelastic case in [16].

Unfortunately, the original contact-stabilization by Deuffhard et al. leads to constant normal velocities at active contact boundaries although vanishing normal components of the velocities are expected by reason of the well-known persistency condition [23, 13]. Hence, in this section, a further improvement of the contact-stabilized Newmark method will be presented, which overcomes this unsatisfactory behavior. As it will turn out later, the vanishing normal velocities are crucial for a beneficial behavior regarding the consistency of the spatiotemporal discretization.

Discretization in space. In a first step, the space is discretized by piecewise linear finite elements. Let Ω_h be a polyhedral domain partitioned into triangles or tetrahedra with $h > 0$ the maximal diameter, and let the sequence of triangulations be shape regular. Denote the corresponding finite element space by \mathbf{S}_h and the set of vertices contained in $\Omega_h \cup \Gamma_{h,N} \cup \Gamma_{h,C}$ by \mathcal{N}_h . In this setting, the discrete approximation $\mathcal{K}_h \subset \mathbf{S}_h$ of the set of admissible displacements is the set

$$\mathcal{K}_h := \{ \mathbf{v}_h \in \mathbf{S}_h \mid [\mathbf{v}_h \cdot \boldsymbol{\nu}_h]_{\phi_h} \leq g_h \quad \forall p \in \mathcal{N}_h \cap \Gamma_{h,C} \}, \quad (10)$$

where $\boldsymbol{\nu}_{\phi_h}$, ϕ_h , and g_h are suitable approximations of $\boldsymbol{\nu}_\phi$, ϕ , and g . Details of the spatial discretization can be found in [12, 18, 19, 20, 27].

Discretization in time. For the temporal discretization, let the continuous time interval $[0, T]$ be subdivided by $N + 1$ discrete timepoints $0 = t_0 < t_1 < \dots < t_N = T$ with $t_n = n \cdot \tau$ for $n = 0, \dots, N$ and τ denoting a given timestep.

With these algorithmic preparations, both the contact-implicit Newmark method by Kane et al. and a novel version of the contact-stabilized Newmark method by Deuffhard et al. can be given.

Contact-implicit Newmark method (N-CI)_h.

$$\mathbf{u}_{h,\text{CI}}^0 = \mathbf{u}(t_0) \quad (11a)$$

$$\dot{\mathbf{u}}_{h,\text{CI}}^0 = \dot{\mathbf{u}}(t_0) \quad (11b)$$

$$0 \in \mathbf{u}_{h,\text{CI}}^{n+1} - \mathbf{u}_{h,\text{CI}}^n - \tau \dot{\mathbf{u}}_{h,\text{CI}}^n + \frac{\tau^2}{2} \left(\mathbf{F} \left(\frac{\mathbf{u}_{h,\text{CI}}^n + \mathbf{u}_{h,\text{CI}}^{n+1}}{2} \right) + \mathbf{G} \left(\frac{\mathbf{u}_{h,\text{CI}}^{n+1} - \mathbf{u}_{h,\text{CI}}^n}{\tau} \right) + \partial I_{\mathcal{K}}(\mathbf{u}_{h,\text{CI}}^{n+1}) \right) \quad (11c)$$

$$\dot{\mathbf{u}}_{h,\text{CI}}^{n+1} = \dot{\mathbf{u}}_{h,\text{CI}}^n + \frac{2}{\tau} (\mathbf{u}_{h,\text{CI}}^{n+1} - \mathbf{u}_{h,\text{CI}}^n - \tau \dot{\mathbf{u}}_{h,\text{CI}}^n) \quad (11d)$$

Improved contact-stabilized Newmark method (N-CS++)_h.

$$\mathbf{u}_h^0 = \mathbf{u}(t_0) \quad (12a)$$

$$\dot{\mathbf{u}}_{h,\text{pred}}^0 = \dot{\mathbf{u}}(t_0) \quad (12b)$$

$$0 \in \dot{\mathbf{u}}_h^0 - \dot{\mathbf{u}}_{h,\text{pred}}^0 + \partial I_{\mathcal{K}}(\mathbf{u}_h^0 + \tau \dot{\mathbf{u}}_h^0) \quad (12c)$$

$$0 \in \mathbf{u}_h^{n+1} - \mathbf{u}_h^n - \tau \dot{\mathbf{u}}_h^n + \frac{\tau^2}{2} \left(\mathbf{F} \left(\frac{\mathbf{u}_h^n + \mathbf{u}_h^{n+1}}{2} \right) + \mathbf{G} \left(\frac{\mathbf{u}_h^{n+1} - \mathbf{u}_h^n}{\tau} \right) + \partial I_{\mathcal{K}}(\mathbf{u}_h^{n+1}) \right) \quad (12d)$$

$$\dot{\mathbf{u}}_{h,\text{pred}}^{n+1} = \dot{\mathbf{u}}_h^n + \frac{2}{\tau} (\mathbf{u}_h^{n+1} - \mathbf{u}_h^n - \tau \dot{\mathbf{u}}_h^n) \quad (12e)$$

$$0 \in \dot{\mathbf{u}}_h^{n+1} - \dot{\mathbf{u}}_{h,\text{pred}}^{n+1} + \partial I_{\mathcal{K}}(\mathbf{u}_h^{n+1} + \tau \dot{\mathbf{u}}_h^{n+1}). \quad (12f)$$

Variational problem. In (11c) and (12d), both Newmark methods require the solution of a nonlinear variational inclusion, which is equivalent to a convex minimization problem under non-penetration constraints [13, 25]. A suitable algorithm for solving these stationary contact problems in each timestep is the adaptive monotone multigrid method by Kornhuber and Krause [20, 18, 19, 17] or its recent improvement by Gräser and Kornhuber, the so-called truncated nonsmooth Newton multigrid method (TNNMG) [8].

Once the variational problem is solved, the contact forces $\mathbf{F}_{\text{con}}(\mathbf{u}_{h,\text{CI}}^{n+1})$ are defined as the residuals of the variational inequalities, i.e.

$$\begin{aligned} & \frac{\tau^2}{2} \langle \mathbf{F}_{\text{con}}(\mathbf{u}_{h,\text{CI}}^{n+1}), \mathbf{v}_h \rangle \\ & := \left\langle \mathbf{u}_{h,\text{CI}}^{n+1} - \mathbf{u}_{h,\text{CI}}^n - \tau \dot{\mathbf{u}}_{h,\text{CI}}^n + \frac{\tau^2}{2} \left(\mathbf{F} \left(\frac{\mathbf{u}_{h,\text{CI}}^n + \mathbf{u}_{h,\text{CI}}^{n+1}}{2} \right) + \mathbf{G} \left(\frac{\mathbf{u}_{h,\text{CI}}^{n+1} - \mathbf{u}_{h,\text{CI}}^n}{\tau} \right) \right), \mathbf{v}_h \right\rangle, \quad \mathbf{v}_h \in \mathbf{H}^1, \end{aligned} \quad (13)$$

and the variational problems (11c) and (12d) can equivalently be formulated as

$$\langle \mathbf{F}_{\text{con}}(\mathbf{u}_{h,\text{CI}}^{n+1}), \mathbf{u}_{h,\text{CI}}^{n+1} - \mathbf{v}_h \rangle_{(\mathbf{H}^1)^* \times \mathbf{H}^1} \leq 0, \quad \forall \mathbf{v}_h \in \mathcal{K}_h. \quad (14)$$

L₂-projection. The contact–stabilization procedure in the novel version (N-CS++)_h of the Newmark method by Deuffhard et al. adds a special non-linear corrector step (12f) to the linear velocity update (11d) and (12e). In addition, the velocities $\dot{\mathbf{u}}_h^0$ at initial time are given by the corrector step (12b) instead of the prescribed velocities $\dot{\mathbf{u}}(t_0)$ from the variational problem (7). These variational inclusions can equivalently be written as the constrained, convex minimization problems

$$\min_{\mathbf{v}_h \in \mathcal{K}_h} \left\| \mathbf{v}_h - \mathbf{u}_h^{n+1} - \tau \dot{\mathbf{u}}_{h,\text{pred}}^{n+1} \right\|_{\mathbf{L}_2(\Omega_h)}. \quad (15)$$

Hence, the corrector steps can be considered as the \mathbf{L}_2 -projections of the finite element functions $\mathbf{u}_h^{n+1} + \tau \dot{\mathbf{u}}_{h,\text{pred}}^{n+1} \in \mathbf{S}_h$ onto the discrete set \mathcal{K}_h of admissible displacements in each timestep. In the case of a full mass matrix, the \mathbf{L}_2 -projection can be solved by a monotone multigrid method [19]. If a lumped mass matrix is used instead, the \mathbf{L}_2 -projection can even be realized by a pointwise projection of the normal trace on the possible contact boundaries.

If the \mathbf{L}_2 -projection is carried out, the predictor step can be rewritten as

$$\left(\mathbf{G}_{\text{con}}(\mathbf{u}_h^{n+1} + \tau \dot{\mathbf{u}}_h^{n+1}), \mathbf{u}_h^{n+1} + \tau \dot{\mathbf{u}}_h^{n+1} - \mathbf{v}_h \right)_{\mathbf{L}_2} \leq 0, \quad \forall \mathbf{v}_h \in \mathcal{K}_h \quad (16)$$

with \mathbf{G}_{con} defined via

$$\left(\mathbf{G}_{\text{con}}(\mathbf{u}_h^{n+1} + \tau \dot{\mathbf{u}}_h^{n+1}), \mathbf{v}_h \right)_{\mathbf{L}_2} = \left(\dot{\mathbf{u}}_h^{n+1} - \dot{\mathbf{u}}_{h,\text{pred}}^{n+1}, \mathbf{v}_h \right)_{\mathbf{L}_2}, \quad \mathbf{v}_h \in \mathcal{K}_h. \quad (17)$$

The positive effects of this \mathbf{L}_2 -projection in view of conservation and stabilization properties will be discussed in detail in the following two sections.

3.1 Conservation properties

In the absence of contact constraints, the symmetric classical Newmark method preserves the linear momentum and the total energy of the discrete evolution [22, 9, 13]. In the constrained case, however, the situation must be reexamined due to the nonlinearity of the contact forces.

For the original contact-stabilized Newmark method by Deuffhard et al., conservation of linear momentum as well as dissipativity of total energy has been proven in [3]. These results can easily be translated to the novel variant of this scheme presented above. However, the proofs will be skipped here, since the calculations are very similar to those performed in [13] for the original contact-stabilized Newmark scheme.

Linear momentum conservation. The improved version of the contact-stabilized Newmark method preserves the linear momentum of the system both in the absence and in the presence of contact.

Theorem 3.1. *The improved contact-stabilized Newmark method (N-CS++)_h conserves the linear momentum if $f_{\text{ext}} = 0$ and $\Gamma_D = \emptyset$.*

Energy dissipativity. The energy of the discrete evolution of the improved contact–stabilized Newmark method is still preserved in the absence of contact (including the viscous energy). In the presence of contact, the implicit handling of the non-penetration constraints leads to a dissipative behavior, which is preserved by the contact–stabilization even for the latest variant presented above.

Theorem 3.2. *Consider the improved contact-stabilized Newmark method $(\text{N-CS}++)_h$ with $f_{\text{ext}} = 0$. If \mathbf{u}_h^{n+1} and $\mathbf{u}_h^{n+1} + \tau \dot{\mathbf{u}}_h^{n+1}$ are not in contact, the algorithm is energy conserving (including the viscous energy). Otherwise, the algorithm is energy dissipative.*

3.2 The contact–stabilization

As mentioned above, numerical instabilities arise in many discretization schemes for dynamical contact problems. The main cause for this undesirable effect is that discretization in space assigns a mass to the discrete contact boundaries, while the boundaries in the continuous problem have measure zero. In consequence, the entries of the discrete mass matrix are transferred into contributions to the contact forces at contact interfaces, which destroy the force equilibrium on account of Newton’s third law of motion. Therefore, the key idea of the contact–stabilization procedure by Deuffhard et al. is to remove the unphysical part of the discrete contact forces. As a result, the spurious oscillations in displacements and contact forces disappear, see the numerical examples in [3, 13].

The original contact–stabilization by Deuffhard et al. can be shown to produce constant normal components of the velocities during phases of active contact [3, 13]. However, the solution of the continuous problem fulfills the well-known persistency condition [23, 13] meaning that the normal velocities are equal to zero in the case of active contact constraints. The presented modification $(\text{N-CS}++)_h$ overcomes the lack of non-vanishing velocity values by performing the \mathbf{L}_2 -projection of the contact–stabilization at the end and not at the beginning of a timestep.

Vanishing normal velocity components and avoidance of artificial oscillations. Assume that contact is found on a part of the possible contact boundaries in some timestep, i.e.

$$[\mathbf{u}_h^{n+1} \cdot \boldsymbol{\nu}_h]_{\phi_h} = g_h \quad \text{on } \Gamma_{C,h}^* \subset \Gamma_{C,h}. \quad (18)$$

In a first step, the variational inequality (16) is evaluated for an admissible finite element function \mathbf{v}_h defined by $\mathbf{v}_h = \mathbf{u}_h^{n+1} \in \mathcal{K}_h$ at the nodes of $\Gamma_{C,h}^*$ and $\mathbf{v}_h = \mathbf{u}_h^{n+1} + \tau \dot{\mathbf{u}}_h^{n+1} \in K_h$ at the nodes of $\Omega_h / \Gamma_{C,h}^*$. This gives the inequality

$$(\mathbf{G}_{\text{con}}(\mathbf{u}_h^{n+1} + \tau \dot{\mathbf{u}}_h^{n+1}), \tau \dot{\mathbf{u}}_h^{n+1})_{\mathbf{L}_2(S_{C,h}^*)} \leq 0,$$

where $S_{C,h}^*$ denotes the stripe of finite elements along the active contact boundary $\Gamma_{C,h}^*$. A second admissible test function is given by $\mathbf{v}_h = \mathbf{u}_h^{n+1} + 2\tau\dot{\mathbf{u}}_h^{n+1}$ at the nodes of $\Gamma_{C,h}^*$ and $\mathbf{v}_h = \mathbf{u}_h^{n+1} + \tau\dot{\mathbf{u}}_h^{n+1}$ at the nodes of $\Omega_h / \Gamma_{C,h}^*$ since

$$\tau[\dot{\mathbf{u}}_h^{n+1} \cdot \boldsymbol{\nu}_h]_{\phi_h} = [\mathbf{u}_h^{n+1} + \tau\dot{\mathbf{u}}_h^{n+1} \cdot \boldsymbol{\nu}_h]_{\phi_h} - [\mathbf{u}_h^{n+1} \cdot \boldsymbol{\nu}_h]_{\phi_h} \leq 0 \quad \text{on } \Gamma_{C,h}^* \subset \Gamma_{C,h}.$$

This choice leads to

$$\begin{aligned} & (\mathbf{G}_{\text{con}}(\mathbf{u}_h^{n+1} + \tau\dot{\mathbf{u}}_h^{n+1}), \mathbf{u}_h^{n+1} + \tau\dot{\mathbf{u}}_h^{n+1} - \mathbf{v}_h)_{\mathbf{L}_2} \\ &= (\mathbf{G}_{\text{con}}(\mathbf{u}_h^{n+1} + \tau\dot{\mathbf{u}}_h^{n+1}), -\tau\dot{\mathbf{u}}_h^{n+1})_{\mathbf{L}_2(S_{C,h}^*)} \\ &\leq 0, \end{aligned}$$

and finally, the combination of the two inequalities yields

$$(\mathbf{G}_{\text{con}}(\mathbf{u}_h^{n+1} + \tau\dot{\mathbf{u}}_h^{n+1}), \dot{\mathbf{u}}_h^{n+1})_{\mathbf{L}_2(\Gamma_{C,h}^*)} = 0. \quad (19)$$

This expression is interpretable in the sense that the improved contact-stabilized Newmark method $(\text{N-CS}++)_h$ enables vanishing normal velocities on the active parts of the possible contact boundaries. This physically reasonable behavior even covers the first timestep with active contact constraints, which is in contrast to an earlier variant suggested by the author in [13].

The vanishing normal velocities of $(\text{N-CS}++)_h$ at active contact nodes finally guarantee that the displacements remain uninterrupted in phases of permanent active contact. This geometric argument explains the desired removal of artificial oscillations at discrete contact interfaces by means of the contact-stabilization.

The contact-stabilization in function space. In view of a consistency and convergence analysis for a spatiotemporal integration scheme, there are two principal choices differing in the sequence of discretization (compare e.g., [4]): the popular method of lines (MOL), in which discretization is performed first in space and then in time, and the method of time layers (MOT), also known as Rothe method, which discretizes first the time and then the space. Regarding Newmark methods for dynamical contact problems, the focus in the literature has been on the method of time layers up to now. Here, the results of this approach shall be briefly collected for a later comparison with the novel theory presented in this paper.

In the framework of the method of time layers, the spatial grid is refined initially, a process analyzed for the contact-stabilized and for the contact-implicit Newmark method in [16, 13]. The spatiotemporal algorithm $(\text{N-CS}++)_h$ differs from $(\text{N-CI})_h$ in the additional variational inclusion in the velocity update. As noted above, this corrector step can equivalently be

formulated as the \mathbf{L}_2 -projection of a finite element function onto the discrete admissible set \mathcal{K}_h , which is related to the pointwise behavior along the possible contact boundaries, cf. (10). Hence, the corrector mainly acts as a modification of the velocities near the contact interfaces, while the nodes in the interior of the domain are only slightly changed. As a result, the effect of the projection completely vanishes, if h tends to zero, due to the measure zero of the contact boundaries in the continuous problem [16, 13]. In consequence, the modified update formula for the improved velocities $\dot{\mathbf{u}}_h^{n+1}$ converges to $\dot{\mathbf{u}}_{h,\text{pred}}^{n+1}$ in \mathbf{L}_2 if the spatial grid vanishes.

In addition to this result, the analysis in [16, 13] has turned out a convergence result concerning the spatial limit of the space-discretized Newmark schemes: if h tends to zero, the solution of the contact-stabilized Newmark method possesses the same limit as the contact-implicit variant. The identical continuous counterpart of both Newmark algorithms in function space reads:

Contact-implicit/improved contact-stabilized Newmark method in function space (N-CI/CS++).

$$\begin{aligned}
\mathbf{u}^0 &= \mathbf{u}(t_0) \\
\dot{\mathbf{u}}^0 &= \dot{\mathbf{u}}(t_0) \\
0 &\in \mathbf{u}^{n+1} - \mathbf{u}^n - \tau \dot{\mathbf{u}}^n + \frac{1}{2} \tau^2 \left(\mathbf{F}^{1/2}(\mathbf{u}^n, \mathbf{u}^{n+1}) \right. \\
&\quad \left. + \mathbf{G}\left(\frac{\mathbf{u}^{n+1} - \mathbf{u}^n}{\tau}\right) + \partial I_{\mathcal{K}}(\mathbf{u}^{n+1}) \right) \\
\dot{\mathbf{u}}^{n+1} &= \dot{\mathbf{u}}^n + \frac{2}{\tau} (\mathbf{u}^{n+1} - \mathbf{u}^n - \tau \dot{\mathbf{u}}^n).
\end{aligned} \tag{20}$$

In view of the convergence result, the admissible set $\mathcal{K} \subset \mathbf{H}_D^1$ is assumed to be approximated by the sets $\mathcal{K}_h \subset \mathbf{S}_h$ in the following way.

Assumption 3.3.

- (i) $\forall \mathbf{v} \in \mathcal{K}, \exists \mathbf{v}_h \in \mathcal{K}_h$ such that $\|\mathbf{v}_h - \mathbf{v}\|_{\mathbf{H}^1} \rightarrow 0$ as $h \rightarrow 0$, and
- (ii) for $\mathbf{w}_h \in \mathcal{K}_h$ “ $\mathbf{w}_h \rightarrow \mathbf{w}$ weakly as $h \rightarrow 0$ ” implies $\mathbf{w} \in \mathcal{K}$

These preparations allow to formulate the following theorem on the convergence behavior of the spatiotemporal Newmark schemes for fixed temporal step size τ .

Theorem 3.4. ([16, 13]) *Assume \mathbf{S}_H to be fixed, $\mathbf{u}_H^n, \dot{\mathbf{u}}_H^n \in \mathbf{S}_H$, and \mathbf{S}_h a family of quasiuniform refinements of \mathbf{S}_H with $h \rightarrow 0$. Let Assumption 3.3 hold. Then, $(\text{N-CI})_h$ and $(\text{N-CS++})_h$ converge to the same limit (N-CI/CS++) for $h \rightarrow 0$, i.e.,*

$$\lim_{h \rightarrow 0} \left(\|\mathbf{u}_{h,(CI)}^{n+1} - \mathbf{u}^{n+1}\|_{\mathbf{H}^1} + \|\dot{\mathbf{u}}_{h,(CI)}^{n+1} - \dot{\mathbf{u}}^{n+1}\|_{\mathbf{L}_2} \right) = 0. \tag{21}$$

4 Consistency Error in Physical Energy Norm

By means of the preceding preparations, this section of the paper contains the main results concerning the consistency error of the improved contact-stabilized Newmark method in time and space.

In the unconstrained case, the symmetric Newmark method is well-known to be pointwise second-order consistent [22]. In the presence of contact constraints, however, the situation is completely different due to the high non-linearity and non-regularity of the arising contact forces. In this case, the consistency and convergence behavior of the contact-stabilized Newmark method has been discussed within the framework of the method of time layers up to now [16, 13]. These results cover both the contact-stabilized and the contact-implicit Newmark scheme in function space, since the \mathbf{L}_2 -projection of the contact-stabilization vanishes in the spatial limit as mentioned above. Hence, the error estimates in this approach can not benefit from the advantages of the contact-stabilization, which motivates the following consistency theory including both spatial and temporal discretization parameters.

After a brief recall of the consistency results in function space in Section 4.1, an estimate for the spatiotemporal consistency error of the improved contact-stabilized Newmark method will be derived in the subsequent Section 4.2. In the last Section 4.3, the novel consistency results will be discussed in view of convergence properties of the spatiotemporal Newmark scheme.

4.1 Consistency error in function space

For the sake of comparison and for using similar results in the proofs to follow, the consistency theory for the common counterpart of the contact-implicit and of the contact-stabilized Newmark method in function space will be revealed in this section. As a preparatory step, some basics concepts will be given.

In order to estimate the local and global discretization errors of the Newmark schemes, the physical energy norm suggested in [15] will be exploited.

Physical energy norm. For a function $\bar{\mathbf{v}} = (\mathbf{v}, \dot{\mathbf{v}}) : [t, t + \tau] \rightarrow \mathbf{H}^1 \times \mathbf{L}_2$ with $\dot{\mathbf{v}} \in \mathbf{L}_2(t, t + \tau; \mathbf{H}^1)$, the physical energy norm is defined as

$$\|\bar{\mathbf{v}}\|_{\mathcal{E}(t, \tau)}^2 := \|\bar{\mathbf{v}}(t + \tau)\|_E^2 + \int_t^{t+\tau} \|\dot{\mathbf{v}}(s)\|_b^2 ds \quad (22)$$

in terms of the reduced norm

$$\|\bar{\mathbf{v}}(t + \tau)\|_E^2 := \frac{1}{2} \|\dot{\mathbf{v}}(t + \tau)\|_{\mathbf{L}_2}^2 + \frac{1}{2} \|\mathbf{v}(t + \tau)\|_a^2. \quad (23)$$

This norm may be interpreted as a sum of the kinetic energy, measured in \mathbf{L}_2 , and the potential energy, measured in the energy norm in \mathbf{H}^1 , including the viscous part. For the later consistency theory, the following variant using a finite difference for the velocities in \mathbf{H}^1 will become important:

$$\|\bar{\mathbf{u}}_h^{n+1} - \bar{\mathbf{u}}(t+\tau)\|_{\tilde{\mathcal{E}}(t,\tau)}^2 := \|\bar{\mathbf{u}}_h^{n+1} - \bar{\mathbf{u}}(t+\tau)\|_E^2 + \int_t^{t+\tau} \left\| \frac{\mathbf{u}_h^{n+1} - \mathbf{u}(t)}{\tau} - \dot{\mathbf{u}}(t+s) \right\|_b^2 ds. \quad (24)$$

In the presence of active contact constraints, the velocities, accelerations, and contact forces of the variational problem (7) become highly irregular. However, with respect to adequate Sobolev spaces, these quantities might be continuous with respect to time except at countable many timepoints. In general, such a behavior is shown by functions of bounded total variation.

Bounded variation. Let $(\mathbf{V}; \|\cdot\|_{\mathbf{V}})$ be a Banach space. The total variation of a function $\mathbf{v} : [t_0, t] \rightarrow \mathbf{V}$ is defined as

$$\text{TV}(\mathbf{v}, [t_0, t], \mathbf{V}) := \sup \left\{ \sum_{j=1}^n \|\mathbf{v}(t_j) - \mathbf{v}(t_{j-1})\|_{\mathbf{V}} : t_0 < t_1 < \dots < t_n = t \right\},$$

and the set of functions from $[t_0, t]$ into \mathbf{V} that have bounded variation is denoted by $\mathbf{BV}([t_0, t], \mathbf{V})$. Moreover, the intriguing property

$$\text{TV}(\mathbf{v}, [t_0, t_1], \mathbf{V}) + \text{TV}(\mathbf{v}, [t_1, t], \mathbf{V}) = \text{TV}(\mathbf{v}, [t_0, t], \mathbf{V}), \quad \text{for } t_0 < t_1 < t \quad (25)$$

holds for every function of bounded variation (compare, e.g., [26]).

In this paper, the considerations are restricted to dynamical contact problems satisfying the following regularity assumptions.

Assumption 4.1.

$$\dot{\mathbf{u}} \in \mathbf{BV}([0, T], \mathbf{H}^1), \quad \ddot{\mathbf{u}} \in \mathbf{BV}([0, T], (\mathbf{H}^1)^*)$$

Due to this assumption, the quantity

$$\begin{aligned} R(\mathbf{u}, [t, t+\tau]) &:= \text{TV}(\mathbf{u}, [t, t+\tau], \mathbf{H}^1) + \text{TV}(\dot{\mathbf{u}}, [t, t+\tau], \mathbf{H}^1) \\ &\quad + \text{TV}(\ddot{\mathbf{u}}, [t, t+\tau], (\mathbf{H}^1)^*) \end{aligned} \quad (26)$$

will arise in all results on the error behavior of Newmark schemes in the presence of contact constraints. The following theorem presents a previous estimate for the consistency error of the common contact-implicit and contact-stabilized Newmark method in function space.

Theorem 4.2. ([16]) *Let Assumption 4.1 hold. Then, for initial values $\mathbf{u}^n = \mathbf{u}(t)$ and $\dot{\mathbf{u}}^n = \dot{\mathbf{u}}(t)$, the local error $\bar{\mathbf{u}}^{n+1} - \bar{\mathbf{u}}(t + \tau) = (\mathbf{u}^{n+1} - \mathbf{u}(t + \tau), \dot{\mathbf{u}}^{n+1} - \dot{\mathbf{u}}(t + \tau))$ of (N-CI/CS++) satisfies*

$$\|\bar{\mathbf{u}}^{n+1} - \bar{\mathbf{u}}(t + \tau)\|_{\tilde{\mathcal{E}}(t,\tau)} = R(\mathbf{u}, [t, t + \tau]) \cdot O(\tau^{1/2}). \quad (27)$$

In general, the term $R(\mathbf{u}, [t, t + \tau])$ on the right-hand side of the consistency result does not contribute to any order in τ . In fact, the consistency order of the scheme may only be of order 1/2 at single timepoints. Therefore, applying standard proof techniques (as ‘‘Lady Windermere’s Fan’’, [9]) does not allow to show convergence of the method due to the principle loss of one order in τ . However, by means of the telescoping property (25), the terms $R(\mathbf{u}, [t, t + \tau])$ on the right-hand side sum up to $R(\mathbf{u}, [0, T])$ over the whole time interval. Hence, the common Newmark method in function space can be shown to be globally convergent of order 1/2. For more details see [13].

4.2 Consistency error in time and space

This section is finally devoted to the main topic of the paper, the consistency error of the improved contact-stabilized Newmark method after discretization in time and in space.

In general, the consistency error of a spatiotemporal discretization only tends to zero if both the timestep τ and the spatial parameter h tend to zero. In the following considerations, interest is mainly on the τ -dependence of the error. If h is fixed, the Newmark scheme is expected to converge to the original dynamical contact problem formulated on the finite dimensional admissible set $\mathcal{K}_h \subset \mathbf{S}_h$ instead of $\mathcal{K} \subset \mathbf{H}^1$.

Variational problem on \mathcal{K}_h . For almost every $t \in [0, T]$, find $\mathbf{u}_h \in \mathcal{K}_h$ with $\mathbf{u}_h(\cdot, t) \in C([0, T], \mathbf{H}^1)$ and $\dot{\mathbf{u}}_h \in \mathbf{W}^{1,2}(0, T; \mathbf{H}^1, \mathbf{L}_2)$ such that for all $\mathbf{v}_h \in \mathcal{K}_h$

$$0 \in \ddot{\mathbf{u}}_h + \mathbf{F}(\mathbf{u}_h) + \mathbf{G}(\dot{\mathbf{u}}_h) + \partial I_{\mathcal{K}}(\mathbf{u}_h) \quad (28)$$

and

$$\mathbf{u}_h(0) = \mathbf{u}_0, \quad \dot{\mathbf{u}}_h(0) = \dot{\mathbf{u}}_0. \quad (29)$$

The solution of this variational problem as well as the algorithmic solution of the improved contact-stabilized Newmark method depend on the spatial discretization parameter h . Both the h -dependent problem (28) and the spatiotemporal Newmark method (N-CS++) $_h$ are formulated by using Sobolev spaces, although all norms are equivalent in the finite dimensional case. Please note that this is in view of the later investigations of the convergence behavior for $h \rightarrow 0$.

By means of the auxiliary dynamical contact problem (28), the consistency error of the spatiotemporal Newmark algorithm measured in physical

energy norm can be split up as follows:

$$\begin{aligned} & \|\bar{\mathbf{u}}_h^{n+1} - \bar{\mathbf{u}}(t + \tau)\|_{\tilde{\mathcal{E}}(t,\tau)} \\ & \leq \|\bar{\mathbf{u}}_h(t + \tau) - \bar{\mathbf{u}}(t + \tau)\|_{\tilde{\mathcal{E}}(t,\tau)} + \|\bar{\mathbf{u}}_h^{n+1} - \bar{\mathbf{u}}_h(t + \tau)\|_{\tilde{\mathcal{E}}(t,\tau)}. \end{aligned} \quad (30)$$

The first part of this estimate describes the difference between the h -dependent solution of (28) and the continuous solution of the original contact problem (9). This term is expected to converge to zero if the spatial parameter h tends to zero. Since the quantity does not depend on the timestep τ in a special way, it will not be discussed further. The second part of the estimate can be interpreted as the consistency error of $(\text{N-CS}++)_h$ compared to the h -dependent solution of (28). Hence, this is the quantity of interest in the following investigations.

Due to the high correspondence of the variational problems (9) and (28), the regularity assumptions for the space-continuous dynamical contact problem are also reasonable for the space-discretized problem. This observation refers to the requirement of total bounded variation.

Assumption 4.3.

$$\dot{\mathbf{u}}_h \in \mathbf{BV}([0, T], \mathbf{H}^1), \quad \ddot{\mathbf{u}}_h \in \mathbf{BV}([0, T], (\mathbf{H}^1)^*)$$

Under this assumption, the second component of the error estimate (30) can be further estimated as follows.

Lemma 4.4. *Let Assumption 4.3 hold. Then, for initial values $\mathbf{u}_h^n = \mathbf{u}_h(t)$ and $\dot{\mathbf{u}}_{h,\text{pred}}^n = \dot{\mathbf{u}}_h(t)$, the local error $\bar{\mathbf{u}}_h^{n+1} - \bar{\mathbf{u}}_h(t + \tau) = (\mathbf{u}_h^{n+1} - \mathbf{u}_h(t + \tau), \dot{\mathbf{u}}_h^{n+1} - \dot{\mathbf{u}}_h(t + \tau))$ of $(\text{N-CS}++)_h$ satisfies*

$$\begin{aligned} & \|\bar{\mathbf{u}}_h^{n+1} - \bar{\mathbf{u}}_h(t + \tau)\|_{\tilde{\mathcal{E}}(t,\tau)} \\ & \leq C \left(R(\mathbf{u}_h, [t, t + \tau]) \cdot \tau^{1/2} + \|\dot{\mathbf{u}}_h^n - \dot{\mathbf{u}}_{h,\text{pred}}^n\|_{\mathbf{L}_2} + \|\dot{\mathbf{u}}_h^{n+1} - \dot{\mathbf{u}}_{h,\text{pred}}^{n+1}\|_{\mathbf{L}_2} \right). \end{aligned} \quad (31)$$

Proof. As a start, the local error in velocities can be split up via

$$\dot{\mathbf{u}}_h^{n+1} - \dot{\mathbf{u}}_h(t + \tau) = (\dot{\mathbf{u}}_{h,\text{pred}}^{n+1} - \dot{\mathbf{u}}_h(t + \tau)) + (\dot{\mathbf{u}}_h^{n+1} - \dot{\mathbf{u}}_{h,\text{pred}}^{n+1}).$$

Using the triangle inequality, this expression allows to estimate the energy norm of the discretization error via

$$\begin{aligned} & \|\bar{\mathbf{u}}_h^{n+1} - \bar{\mathbf{u}}_h(t + \tau)\|_{\tilde{\mathcal{E}}(t,\tau)} \\ & \leq \left(\frac{1}{2} \|\dot{\mathbf{u}}_{h,\text{pred}}^{n+1} - \dot{\mathbf{u}}_h(t + \tau)\|_{\mathbf{L}_2}^2 + \frac{1}{2} \|\mathbf{u}_h^{n+1} - \mathbf{u}_h(t + \tau)\|_a^2 \right. \\ & \quad \left. + \int_t^{t+\tau} \left\| \frac{\mathbf{u}_h^{n+1} - \mathbf{u}_h^n}{\tau} - \dot{\mathbf{u}}_h(t + s) \right\|_b^2 ds \right)^{1/2} + \frac{1}{\sqrt{2}} \|\dot{\mathbf{u}}_h^{n+1} - \dot{\mathbf{u}}_{h,\text{pred}}^{n+1}\|_{\mathbf{L}_2}. \end{aligned}$$

In a next step, the discrete solution $(\mathbf{u}_{h,\text{CI}}^{n+1}, \dot{\mathbf{u}}_{h,\text{CI}}^{n+1})$ of the contact-implicit Newmark method is inserted into the physical energy norm on the right-hand side of this inequality:

$$\begin{aligned}
& \left(\frac{1}{2} \|\dot{\mathbf{u}}_{h,\text{pred}}^{n+1} - \dot{\mathbf{u}}_h(t+\tau)\|_{\mathbf{L}_2}^2 + \frac{1}{2} \|\mathbf{u}_h^{n+1} - \mathbf{u}_h(t+\tau)\|_a^2 \right. \\
& \quad \left. + \int_t^{t+\tau} \left\| \frac{\mathbf{u}_h^{n+1} - \mathbf{u}_h^n}{\tau} - \dot{\mathbf{u}}_h(t+s) \right\|_b^2 ds \right)^{1/2} \\
& \leq \left(\frac{1}{2} \|\dot{\mathbf{u}}_{h,\text{pred}}^{n+1} - \dot{\mathbf{u}}_{h,\text{CI}}^{n+1}\|_{\mathbf{L}_2}^2 + \frac{1}{2} \|\mathbf{u}_h^{n+1} - \mathbf{u}_{h,\text{CI}}^{n+1}\|_a^2 + \tau \|\mathbf{u}_h^{n+1} - \mathbf{u}_{h,\text{CI}}^{n+1}\|_b^2 \right)^{1/2} \\
& \quad + \|\bar{\mathbf{u}}_{h,\text{CL}}^{n+1} - \bar{\mathbf{u}}_h(t+\tau)\|_{\tilde{\mathcal{E}}(t,\tau)}.
\end{aligned}$$

By means of

$$\dot{\mathbf{u}}_{h,\text{pred}}^{n+1} - \dot{\mathbf{u}}_{h,\text{CI}}^{n+1} = -(\dot{\mathbf{u}}_h^n - \dot{\mathbf{u}}_{h,\text{pred}}^n) + \frac{2}{\tau} (\mathbf{u}_h^{n+1} - \mathbf{u}_{h,\text{CI}}^{n+1}),$$

the defining equations of the Newmark algorithms lead to

$$\begin{aligned}
& \frac{1}{2} \|\dot{\mathbf{u}}_{h,\text{pred}}^{n+1} - \dot{\mathbf{u}}_{h,\text{CI}}^{n+1}\|_{\mathbf{L}_2}^2 + \frac{1}{2} \|\mathbf{u}_h^{n+1} - \mathbf{u}_{h,\text{CI}}^{n+1}\|_a^2 + \tau \|\mathbf{u}_h^{n+1} - \mathbf{u}_{h,\text{CI}}^{n+1}\|_b^2 \\
& = -\frac{1}{2} (\dot{\mathbf{u}}_{h,\text{pred}}^{n+1} - \dot{\mathbf{u}}_{h,\text{CI}}^{n+1}, \dot{\mathbf{u}}_h^n - \dot{\mathbf{u}}_{h,\text{pred}}^n)_{\mathbf{L}_2} \\
& \quad + \frac{1}{\tau} (\dot{\mathbf{u}}_{h,\text{pred}}^{n+1} - \dot{\mathbf{u}}_{h,\text{CI}}^{n+1}, \mathbf{u}_h^{n+1} - \mathbf{u}_{h,\text{CI}}^{n+1})_{\mathbf{L}_2} \\
& \quad + \frac{1}{2} \|\mathbf{u}_h^{n+1} - \mathbf{u}_{h,\text{CI}}^{n+1}\|_a^2 + \tau \|\mathbf{u}_h^{n+1} - \mathbf{u}_{h,\text{CI}}^{n+1}\|_b^2 \\
& = -\frac{1}{2} (\dot{\mathbf{u}}_{h,\text{pred}}^{n+1} - \dot{\mathbf{u}}_{h,\text{CI}}^{n+1}, \dot{\mathbf{u}}_h^n - \dot{\mathbf{u}}_{h,\text{pred}}^n)_{\mathbf{L}_2} \\
& \quad + \frac{1}{\tau} (\dot{\mathbf{u}}_h^n - \dot{\mathbf{u}}_{h,\text{pred}}^n, \mathbf{u}_h^{n+1} - \mathbf{u}_{h,\text{CI}}^{n+1})_{\mathbf{L}_2} \\
& \quad + \langle \mathbf{F}_{\text{con}}(\mathbf{u}_h^{n+1}) - \mathbf{F}_{\text{con}}(\mathbf{u}_{h,\text{CI}}^{n+1}), \mathbf{u}_h^{n+1} - \mathbf{u}_{h,\text{CI}}^{n+1} \rangle_{\mathbf{H}^1 \times (\mathbf{H}^1)^*} \\
& = \frac{1}{2} \|\dot{\mathbf{u}}_h^n - \dot{\mathbf{u}}_{h,\text{pred}}^n\|_{\mathbf{L}_2}^2 \\
& \quad + \langle \mathbf{F}_{\text{con}}(\mathbf{u}_h^{n+1}) - \mathbf{F}_{\text{con}}(\mathbf{u}_{h,\text{CI}}^{n+1}), \mathbf{u}_h^{n+1} - \mathbf{u}_{h,\text{CI}}^{n+1} \rangle_{\mathbf{H}^1 \times (\mathbf{H}^1)^*} \\
& \leq \frac{1}{2} \|\dot{\mathbf{u}}_h^n - \dot{\mathbf{u}}_{h,\text{pred}}^n\|_{\mathbf{L}_2}^2.
\end{aligned}$$

Combining this with the result of Theorem 4.2 applied to the variational

problem (28) gives

$$\begin{aligned}
& \left(\frac{1}{2} \left\| \dot{\mathbf{u}}_{h,\text{pred}}^{n+1} - \dot{\mathbf{u}}_h(t + \tau) \right\|_{\mathbf{L}_2}^2 + \frac{1}{2} \left\| \mathbf{u}_h^{n+1} - \mathbf{u}_h(t + \tau) \right\|_a^2 \right. \\
& \quad \left. + \int_t^{t+\tau} \left\| \frac{\mathbf{u}_h^{n+1} - \mathbf{u}_h^n}{\tau} - \dot{\mathbf{u}}_h(t + s) \right\|_b^2 ds \right)^{1/2} \\
& \leq CR(\mathbf{u}_h, [t, t + \tau]) \tau^{1/2} + \frac{1}{\sqrt{2}} \left\| \dot{\mathbf{u}}_h^n - \dot{\mathbf{u}}_{h,\text{pred}}^n \right\|_{\mathbf{L}_2},
\end{aligned} \tag{32}$$

which yields the result of the lemma.

In summary, the local discretization error of the improved contact-stabilized Newmark method can be split up into one component describing the consistency error of the contact-implicit Newmark scheme and a second one containing the error contributions of the \mathbf{L}_2 -projections. \square

Now, the main challenge is to find a sharp estimate for the contributions of the \mathbf{L}_2 -projections to the consistency error estimate (31). For this purpose, the interpretation of the discrete \mathbf{L}_2 -projections as minimization problems gives valuable insights.

Lemma 4.5. *Let Assumption 4.1 hold. Then, for initial values $\mathbf{u}_h^n = \mathbf{u}_h(t)$ and $\dot{\mathbf{u}}_{h,\text{pred}}^n = \dot{\mathbf{u}}_h(t)$, the \mathbf{L}_2 -projections satisfy*

$$\left\| \dot{\mathbf{u}}_h^n - \dot{\mathbf{u}}_{h,\text{pred}}^n \right\|_{\mathbf{L}_2} \leq \min_{\mathbf{v}_h \in \mathcal{K}_h} \left\| \frac{\mathbf{v}_h - \mathbf{u}_h(t)}{\tau} - \dot{\mathbf{u}}_h(t) \right\|_{\mathbf{L}_2} \tag{33}$$

and

$$\begin{aligned}
\left\| \dot{\mathbf{u}}_h^{n+1} - \dot{\mathbf{u}}_{h,\text{pred}}^{n+1} \right\|_{\mathbf{L}_2} & \leq \min_{\mathbf{v}_h \in \mathcal{K}_h} \left\| \frac{\mathbf{v}_h - \mathbf{u}_h(t + \tau)}{\tau} - \dot{\mathbf{u}}_h(t + \tau) \right\|_{\mathbf{L}_2} \\
& \quad + \min_{\mathbf{v}_h \in \mathcal{K}_h} \left\| \frac{\mathbf{v}_h - \mathbf{u}_h(t)}{\tau} - \dot{\mathbf{u}}_h(t) \right\|_{\mathbf{L}_2} \\
& \quad + CR(\mathbf{u}_h, [t, t + \tau]) \cdot \tau^{1/2}
\end{aligned} \tag{34}$$

with

(i)

$$\min_{\mathbf{v}_h \in \mathcal{K}_h} \left\| \frac{\mathbf{v}_h - \mathbf{u}_h(t + \tau)}{\tau} - \dot{\mathbf{u}}_h(t + \tau) \right\|_{\mathbf{L}_2} \leq CR^{1/2}(\mathbf{u}_h, [t + \tau, t + 2\tau]) \cdot \tau^{1/2} \tag{35}$$

(ii)

$$\begin{aligned}
& \min_{\mathbf{v}_h \in \mathcal{K}_h} \left\| \frac{\mathbf{v}_h - \mathbf{u}_h(t + \tau)}{\tau} - \dot{\mathbf{u}}_h(t + \tau) \right\|_{\mathbf{L}_2} \\
& \leq C(h \cdot \text{meas}\Gamma_{h,C}^*[t + \tau, t + 2\tau])^{\frac{1}{2} - \frac{1}{p}} R(\mathbf{u}_h, [t + \tau, t + 2\tau])
\end{aligned} \tag{36}$$

where $0 < \frac{1}{2} - \frac{1}{p} < \frac{1}{d}$ and

$$\begin{aligned} \Gamma_{h,C}^*[t + \tau, t + 2\tau] &:= \{[\mathbf{u}_h(t + \tau) \cdot \boldsymbol{\nu}]_{\phi_h} < g_h \\ &\text{and } [(\mathbf{u}_h(t + \tau) + \tau \dot{\mathbf{u}}_h(t + \tau)) \cdot \boldsymbol{\nu}_h]_{\phi_h} > g_h\}. \end{aligned} \quad (37)$$

Proof. Due to the minimization property (15) of the discrete \mathbf{L}_2 -projections in (12c) and (12f) of $(\text{N-CS}++)_h$,

$$\|\dot{\mathbf{u}}_h^k - \dot{\mathbf{u}}_{h,\text{pred}}^k\|_{\mathbf{L}_2} \leq \min_{\mathbf{v}_h \in \mathcal{K}_h} \left\| \frac{\mathbf{v}_h - \mathbf{u}_h^k}{\tau} - \dot{\mathbf{u}}_{h,\text{pred}}^k \right\|_{\mathbf{L}_2}, \quad \text{for } k = n, n + 1.$$

For $k = n$, the initial values $\mathbf{u}_h^n = \mathbf{u}_h(t)$ and $\dot{\mathbf{u}}_{h,\text{pred}}^n = \dot{\mathbf{u}}_h(t)$ give the first estimate of the lemma. For $k = n + 1$, equation (12e) for the velocities in $(\text{N-CS}++)_h$ yields

$$\begin{aligned} &\|\dot{\mathbf{u}}_h^{n+1} - \dot{\mathbf{u}}_{h,\text{pred}}^{n+1}\|_{\mathbf{L}_2} \\ &\leq \min_{\mathbf{v}_h \in \mathcal{K}_h} \left\| \frac{\mathbf{v}_h - \mathbf{u}_h^{n+1}}{\tau} - \dot{\mathbf{u}}_{h,\text{pred}}^{n+1} \right\|_{\mathbf{L}_2} \\ &\leq \min_{\mathbf{v}_h \in \mathcal{K}_h} \left\| \frac{\mathbf{v}_h - \mathbf{u}_h(t + \tau)}{\tau} - \dot{\mathbf{u}}_h(t + \tau) \right\|_{\mathbf{L}_2} + \|\dot{\mathbf{u}}_h(t + \tau) - \dot{\mathbf{u}}_{h,\text{pred}}^{n+1}\|_{\mathbf{L}_2} \\ &\quad + \frac{1}{\tau} \|\mathbf{u}_h(t + \tau) - \mathbf{u}_h^{n+1}\|_{\mathbf{L}_2} \\ &= \min_{\mathbf{v}_h \in \mathcal{K}_h} \left\| \frac{\mathbf{v}_h - \mathbf{u}_h(t + \tau)}{\tau} - \dot{\mathbf{u}}_h(t + \tau) \right\|_{\mathbf{L}_2} + \|\dot{\mathbf{u}}_h(t + \tau) - \dot{\mathbf{u}}_{h,\text{pred}}^{n+1}\|_{\mathbf{L}_2} \\ &\quad + \frac{1}{\tau} \left\| \mathbf{u}_h(t + \tau) - \mathbf{u}_h(t) - \frac{\tau}{2} (\dot{\mathbf{u}}_h^n + \dot{\mathbf{u}}_{h,\text{pred}}^{n+1}) \right\|_{\mathbf{L}_2} \\ &\leq \min_{\mathbf{v}_h \in \mathcal{K}_h} \left\| \frac{\mathbf{v}_h - \mathbf{u}_h(t + \tau)}{\tau} - \dot{\mathbf{u}}_h(t + \tau) \right\|_{\mathbf{L}_2} \\ &\quad + \frac{1}{2} \|\dot{\mathbf{u}}_h^n - \dot{\mathbf{u}}_{h,\text{pred}}^n\|_{\mathbf{L}_2} + \frac{3}{2} \|\dot{\mathbf{u}}_h(t + \tau) - \dot{\mathbf{u}}_{h,\text{pred}}^{n+1}\|_{\mathbf{L}_2} \\ &\quad + \frac{1}{\tau} \left\| \mathbf{u}_h(t + \tau) - \mathbf{u}_h(t) - \frac{\tau}{2} (\dot{\mathbf{u}}_h(t) + \dot{\mathbf{u}}_h(t + \tau)) \right\|_{\mathbf{L}_2}. \end{aligned}$$

Using the previous estimate (33), the consistency result (32), and

$$\begin{aligned} &\left\| \mathbf{u}_h(t + \tau) - \mathbf{u}_h(t) - \frac{\tau}{2} (\dot{\mathbf{u}}_h(t) + \dot{\mathbf{u}}_h(t + \tau)) \right\|_{\mathbf{L}_2}^2 \\ &= \left\| \int_t^{t+\tau} (\dot{\mathbf{u}}_h(t + s) - \dot{\mathbf{u}}_h(t)) - \frac{1}{2} (\dot{\mathbf{u}}_h(t + \tau) - \dot{\mathbf{u}}_h(t)) ds \right\|_{\mathbf{L}_2}^2 \\ &= \left\langle \int_t^{t+\tau} \left(\int_t^{t+s} \ddot{\mathbf{u}}_h(t + \eta) d\eta - \frac{1}{2} \int_t^{t+\tau} \ddot{\mathbf{u}}_h(t + \zeta) d\zeta \right) ds, \right. \\ &\quad \left. \int_t^{t+\tau} (\dot{\mathbf{u}}_h(t + s) - \dot{\mathbf{u}}_h(t)) - \frac{1}{2} (\dot{\mathbf{u}}_h(t + \tau) - \dot{\mathbf{u}}_h(t)) ds \right\rangle_{(\mathbf{H}^1)^* \times \mathbf{H}^1} \end{aligned}$$

$$\begin{aligned}
&= \left\langle \int_t^{t+\tau} \left(\int_t^{t+s} (\ddot{\mathbf{u}}_h(t+\eta) - \ddot{\mathbf{u}}_h(t)) d\eta - \frac{1}{2} \int_t^{t+\tau} (\ddot{\mathbf{u}}_h(t+\zeta) - \ddot{\mathbf{u}}_h(t)) d\zeta \right) ds, \right. \\
&\quad \left. \int_t^{t+\tau} \left(\dot{\mathbf{u}}_h(t+s) - \dot{\mathbf{u}}_h(t) \right) - \frac{1}{2} \left(\dot{\mathbf{u}}_h(t+\tau) - \dot{\mathbf{u}}_h(t) \right) ds \right\rangle_{(\mathbf{H}^1)^* \times \mathbf{H}^1} \\
&\leq \int_t^{t+\tau} \left(\int_t^{t+s} \|\ddot{\mathbf{u}}_h(t+\eta) - \ddot{\mathbf{u}}_h(t)\|_{(\mathbf{H}^1)^*} d\eta \right. \\
&\quad \left. + \frac{1}{2} \int_t^{t+\tau} \|\ddot{\mathbf{u}}_h(t+\zeta) - \ddot{\mathbf{u}}_h(t)\|_{(\mathbf{H}^1)^*} d\zeta \right) ds \\
&\quad \int_t^{t+\tau} \left(\|\dot{\mathbf{u}}_h(t+s) - \dot{\mathbf{u}}_h(t)\|_{\mathbf{H}^1} + \frac{1}{2} \|\dot{\mathbf{u}}_h(t+\tau) - \dot{\mathbf{u}}_h(t)\|_{\mathbf{H}^1} \right) ds \\
&\leq CR(\mathbf{u}_h, [t, t+\tau])^2 \cdot \tau^3,
\end{aligned}$$

the validity of the second estimate (34) is proven.

(i) Since $\mathbf{v}_h = \mathbf{u}_h(t+2\tau) \in \mathcal{K}_h$,

$$\begin{aligned}
&\min_{\mathbf{v}_h \in \mathcal{K}_h} \left\| \frac{\mathbf{v}_h - \mathbf{u}_h(t+\tau)}{\tau} - \dot{\mathbf{u}}_h(t+\tau) \right\|_{\mathbf{L}_2} \\
&\leq \left\| \frac{\mathbf{u}_h(t+2\tau) - \mathbf{u}_h(t+\tau)}{\tau} - \dot{\mathbf{u}}_h(t+\tau) \right\|_{\mathbf{L}_2} \\
&\leq \frac{1}{\tau} \int_0^\tau \|\dot{\mathbf{u}}_h(t+\tau+s) - \dot{\mathbf{u}}_h(t+\tau)\|_{\mathbf{L}_2} ds \\
&\leq \sup_{s \in [0, \tau]} \|\dot{\mathbf{u}}_h(t+\tau+s) - \dot{\mathbf{u}}_h(t+\tau)\|_{\mathbf{L}_2} \tau
\end{aligned}$$

holds where

$$\begin{aligned}
&\frac{1}{2} \|\dot{\mathbf{u}}_h(t+\tau+s) - \dot{\mathbf{u}}_h(t+\tau)\|_{\mathbf{L}_2}^2 \\
&= \frac{1}{2} \|\dot{\mathbf{u}}_h(t+\tau+\zeta) - \dot{\mathbf{u}}_h(t+\tau)\|_{\mathbf{L}_2}^2 \Big|_{\zeta=0}^{\zeta=s} \\
&= \int_0^s \langle \ddot{\mathbf{u}}_h(t+\tau+\zeta), \dot{\mathbf{u}}_h(t+\tau+\zeta) - \dot{\mathbf{u}}_h(t+\tau) \rangle_{(\mathbf{H}^1)^* \times \mathbf{H}^1} d\zeta \\
&\leq \int_0^s \|\ddot{\mathbf{u}}_h(t+\tau+\zeta)\|_{(\mathbf{H}^1)^*} \|\dot{\mathbf{u}}_h(t+\tau+\zeta) - \dot{\mathbf{u}}_h(t+\tau)\|_{\mathbf{H}^1} d\zeta \\
&\leq s \sup_{\zeta \in [0, s]} \|\ddot{\mathbf{u}}_h(t+\tau+\zeta)\|_{(\mathbf{H}^1)^*} \sup_{\zeta \in [0, s]} \|\dot{\mathbf{u}}_h(t+\tau+\zeta) - \dot{\mathbf{u}}_h(t+\tau)\|_{\mathbf{H}^1} \\
&\leq CR(\mathbf{u}_h, [t+\tau, t+2\tau]) \cdot \tau.
\end{aligned}$$

This yields the first estimate (35) for the minimum.

(ii) Let $S_h^*[t+\tau, t+2\tau]$ denote the small stripe of finite elements along the possible contact boundaries where $[\mathbf{u}_h(t+\tau) \cdot \boldsymbol{\nu}_h]_{\phi_h} < g_h$, but $[(\mathbf{u}_h(t+\tau) + \tau \dot{\mathbf{u}}_h(t+\tau)) \cdot \boldsymbol{\nu}_h]_{\phi_h} > g_h$. Choose $\mathbf{v}_h = \mathbf{u}_h(t+2\tau) \in \mathcal{K}_h$ at this part of the contact boundaries and $\mathbf{v}_h = \mathbf{u}_h(t+\tau) + \tau \dot{\mathbf{u}}_h(t+\tau)$ at

the remaining parts of the contact boundaries as well as in the interior of the domain. Then, \mathbf{v}_h is admissible since the first derivatives $\dot{\mathbf{u}}_h$ of the continuous solution are equal to zero in active contact due to the persistency condition [23, 13]. This leads to

$$\begin{aligned} & \min_{\mathbf{v}_h \in \mathcal{K}_h} \left\| \frac{\mathbf{v}_h - \mathbf{u}_h(t + \tau)}{\tau} - \dot{\mathbf{u}}_h(t + \tau) \right\|_{\mathbf{L}_2} \\ & \leq \left\| \frac{\mathbf{u}_h(t + 2\tau) - \mathbf{u}_h(t + \tau)}{\tau} - \dot{\mathbf{u}}_h(t + \tau) \right\|_{\mathbf{L}_2(S_h^*[t+\tau, t+2\tau])} \\ & \leq \frac{1}{\tau} \int_{\tau}^{2\tau} \left\| \dot{\mathbf{u}}_h(t + s) - \dot{\mathbf{u}}_h(t + \tau) \right\|_{\mathbf{L}_2(S_h^*[t+\tau, t+2\tau])} ds, \end{aligned}$$

and applying Hölder's inequality gives

$$\begin{aligned} & \left\| \dot{\mathbf{u}}_h(t + s) - \dot{\mathbf{u}}_h(t + \tau) \right\|_{\mathbf{L}_2(S_h^*[t+\tau, t+2\tau])} \\ & \leq (\text{meas}(S_h^*[t + \tau, t + 2\tau]))^{\frac{1}{2} - \frac{1}{p}} \left\| \dot{\mathbf{u}}_h(t + s) - \dot{\mathbf{u}}_h(t + \tau) \right\|_{\mathbf{L}_p(S_h^*[t+\tau, t+2\tau])} \end{aligned}$$

with the Lebesgue space \mathbf{L}_p for $p > 2$ (see, e.g., [1, Theorem 2.14]). Denoting by $\Gamma_{h,C}^*[t + \tau, t + 2\tau]$ the part of the possible contact boundaries where $[\mathbf{u}_h(t + \tau) \cdot \boldsymbol{\nu}_h]_{\phi_h} < g_h$, but $[(\mathbf{u}_h(t + \tau) + \tau \dot{\mathbf{u}}_h(t + \tau)) \cdot \boldsymbol{\nu}_h]_{\phi_h} > g_h$, the rough estimate

$$\text{meas}(S_h^*[t + \tau, t + 2\tau]) = O(h) \cdot \text{meas}(\Gamma_{h,C}^*[t + \tau, t + 2\tau])$$

leads to

$$\begin{aligned} & \left\| \dot{\mathbf{u}}_h(t + s) - \dot{\mathbf{u}}_h(t + \tau) \right\|_{\mathbf{L}_2(S_h^*[t+\tau, t+2\tau])} \\ & \leq C(h \cdot \text{meas}(\Gamma_{h,C}^*[t + \tau, t + 2\tau]))^{\frac{1}{2} - \frac{1}{p}} \left\| \dot{\mathbf{u}}_h(t + s) - \dot{\mathbf{u}}_h(t + \tau) \right\|_{\mathbf{L}_p(\Omega_h)}. \end{aligned}$$

Since $\dot{\mathbf{u}}_h \in \mathbf{H}^1$ and $\mathbf{H}^1 \hookrightarrow \mathbf{L}_p$ for $p < \frac{2d}{d-2}$ with the space dimension $d = 2, 3$ (Sobolev embedding, cf., e.g., [1]), there exists a constant C independent of h such that

$$\begin{aligned} & \left\| \dot{\mathbf{u}}_h(t + s) - \dot{\mathbf{u}}_h(t + \tau) \right\|_{\mathbf{L}_2(S_h^*[t+\tau, t+2\tau])} \\ & \leq C(h \cdot \text{meas}(\Gamma_{h,C}^*[t + \tau, t + 2\tau]))^{\frac{1}{2} - \frac{1}{p}} \left\| \dot{\mathbf{u}}_h(t + s) - \dot{\mathbf{u}}_h(t + \tau) \right\|_{\mathbf{H}^1(\Omega_h)} \\ & \leq C(h \cdot \text{meas}(\Gamma_{h,C}^*[t + \tau, t + 2\tau]))^{\frac{1}{2} - \frac{1}{p}} R(\mathbf{u}_h, [t + \tau, t + 2\tau]) \end{aligned}$$

for $2 < p < \frac{2d}{d-2}$. This gives the last estimate (36) of the lemma. \square

The lemma above contains two different possibilities for estimating the contributions of the \mathbf{L}_2 -projections to the consistency error estimate of

Lemma 4.4. The first one mainly bounds the \mathbf{L}_2 -norm against the quantity $R(\mathbf{u}_h)^{1/2}\tau^{1/2}$, which is equal to $\tau^{1/2}$ in the worst case that the total variations do not contribute to any order. Unfortunately, this estimate does not allow to prove the convergence of the scheme for $\tau \rightarrow 0$ since the square roots of total variations do not sum up. The second estimate additionally shows a dependency of the \mathbf{L}_2 -projections on the measure of the critical boundary set $\Gamma_{h,C}^*$. This set represents those points at the possible contact boundaries where $\mathbf{u}_h(t + \tau)$ at the beginning of the timestep is not active, but $\mathbf{u}_h(t + \tau) + \tau\dot{\mathbf{u}}_h(t + \tau) \approx \mathbf{u}_h(t + 2\tau)$ is active. Therefore, the term $(\text{meas}\Gamma_{h,C}^*)^{\frac{1}{2}-\frac{1}{p}}R(\mathbf{u}_h)$ does not need to tend to zero for $\tau \rightarrow 0$ in every single time interval, but it may be of higher order than the first one in many intervals.

In order to cope with these observations adequately, a combination of the two error estimates of Lemma 4.5 shall be inserted into the result of the preceding Lemma 4.4. This directly leads to the following central theorem on the local discretization error of the improved contact-stabilized Newmark method in space and time.

Theorem 4.6. *Let Assumption 4.1 hold. Then, for initial values $\mathbf{u}_h^n = \mathbf{u}_h(t)$ and $\dot{\mathbf{u}}_{h,\text{pred}}^n = \dot{\mathbf{u}}_h(t)$, the local error $\bar{\mathbf{u}}_h^{n+1} - \bar{\mathbf{u}}_h(t + \tau) = (\mathbf{u}_h^{n+1} - \mathbf{u}_h(t + \tau), \dot{\mathbf{u}}_h^{n+1} - \dot{\mathbf{u}}_h(t + \tau))$ of (N-CS++)_h satisfies*

$$\begin{aligned} & \left\| \bar{\mathbf{u}}_h^{n+1} - \bar{\mathbf{u}}_h(t + \tau) \right\|_{\tilde{\mathcal{E}}(t,\tau)} \\ &= R(\mathbf{u}_h, [t, t + \tau]) \cdot O(\tau^{1/2}) \\ & \quad + (S^\alpha(\mathbf{u}_h, [t, t + \tau]) + S^\alpha(\mathbf{u}_h, [t + \tau, t + 2\tau])) \cdot O(\tau^{(1-\alpha)/2}) \end{aligned} \quad (38)$$

where

$$S^\alpha(\mathbf{u}_h, [t, t + \tau]) := (h \cdot \text{meas}\Gamma_{h,C}^*[t, t + \tau])^{\alpha(\frac{1}{2}-\frac{1}{p})} R^{\frac{1+\alpha}{2}}(\mathbf{u}_h, [t, t + \tau]) \quad (39)$$

with $\alpha \in [0, 1]$, $0 < \frac{1}{2} - \frac{1}{p} < \frac{1}{d}$ and

$$\begin{aligned} \Gamma_{h,C}^*[t + \tau, t + 2\tau] &:= \{[\mathbf{u}_h(t + \tau) \cdot \boldsymbol{\nu}]_{\phi_h} < g_h \\ & \quad \text{and } [(\mathbf{u}_h(t + \tau) + \tau\dot{\mathbf{u}}_h(t + \tau)) \cdot \boldsymbol{\nu}]_{\phi_h} > g_h\}. \end{aligned} \quad (40)$$

Proof. Weighting (35) and (36) of Lemma 4.4 with the factor $\alpha \in [0, 1]$ and $(1 - \alpha) \in [0, 1]$, respectively, and inserting the results into the error estimate of Lemma 4.5. \square

The theorem above provides an estimate for the consistency order of the improved contact-stabilized Newmark method in time and space that depends on the total variations of the h -dependent continuous solution and the measure of the critical boundary $\Gamma_{h,C}^*$. Under the assumption of bounded total variation, both quantities do not need to provide an order of τ at every single timepoint, but they may show a certain order at most timepoints. The behavior of the local discretization error as well as its consequences on the global error of the scheme will be discussed in the following section.

4.3 Discussion of consistency and convergence order

The final section of this paper is devoted to a detailed discussion of the consistency and convergence behavior of the improved contact-stabilized Newmark method after discretization both in time and in space.

Focussing on the local discretization error first, the quantities $R(\mathbf{u}_h)$ and $\text{meas}\Gamma_{h,C}^*$ arising in Theorem 4.6 do not need to contribute to any consistency order of the Newmark scheme in each timestep. In fact, the regularity assumption of bounded total variation only provides

$$R(\mathbf{u}, [t, t + \tau]) < C, \quad R(\mathbf{u}, [t + \tau, t + 2\tau]) < C, \quad (41)$$

and

$$\text{meas}\Gamma_C^*([t + \tau, t + 2\tau]) \leq \text{meas}\Gamma_C, \quad (42)$$

in general. With these bounds, the best error estimate is obtained by the choice $\alpha = 0$ leading to

$$\|\bar{\mathbf{u}}_h^{n+1} - \bar{\mathbf{u}}_h(t + \tau)\|_{\tilde{\mathcal{E}}(t,\tau)} = O(\tau^{1/2}). \quad (43)$$

Thus, the improved contact-stabilized Newmark method in time and space shows the same worst case consistency order $1/2$ as the corresponding Newmark method in function space, cf. Section 4.1. But, fortunately, the continuous solution of bounded variation may be of much higher regularity at most of the timepoints, which also results in a higher consistency order.

In view of the global discretization error of the spatiotemporal Newmark algorithm, the natural idea is to utilize the same proof technique as for the convergence of the corresponding time discretization in function space, compare [13]. This approach is based on a less popular version of the classical Lady Windermere's fan by Hairer, Nørsett, and Wanner [10] which needs the stability of the algorithmic solution under perturbations of the initial data. Up to now, the existence of such a perturbation result has only been analyzed for the common Newmark method in function space [13]. A perturbation analysis for the improved contact-stabilized Newmark method in time and space shall not be the content of this paper. Instead, the existence of such a perturbation result will just be assumed in the following.

The modified Lady Windermere's fan means to sum up the consistency errors along the solution of the dynamical contact problem over the whole time interval of interest. With regard to the telescoping property of the total variations, the first attempt is to bound the measures of the critical sets $\Gamma_{h,C}^*$ against the whole contact boundaries and to choose the parameter $\alpha = 1$. Then, Theorem 4.6 yields

$$\begin{aligned} & \|\bar{\mathbf{u}}_h^{n+1} - \bar{\mathbf{u}}_h(t + \tau)\|_{\tilde{\mathcal{E}}(t,\tau)} \\ &= R(\mathbf{u}_h, [t, t + \tau]) \cdot O(\tau^{1/2}) \\ & \quad + (R(\mathbf{u}_h, [t, t + \tau]) + R(\mathbf{u}_h, [t + \tau, t + 2\tau])) \cdot O(h^{\frac{1}{2} - \frac{1}{p}}) \end{aligned}$$

for the local error, and the telescoping property

$$\sum_{n=0}^{N-1} R(\mathbf{u}_h, [t_n, t_{n+1}]) = R(\mathbf{u}_h, [t_0, T])$$

leads to the global error

$$\|\bar{\mathbf{u}}_h^N - \bar{\mathbf{u}}_h(T)\|_E = R(\mathbf{u}_h, [t_0, T]) \cdot O(\tau^{1/2} + h^{\frac{1}{2} - \frac{1}{p}}). \quad (44)$$

Unfortunately, this estimate yields convergence of the improved contact-stabilized Newmark method only if both the spatial parameter h and the timestep τ tend to zero. A proof of convergence for $\tau \rightarrow 0$ but for fixed h requires a more advanced concept.

For this purpose, the measures $\text{meas}\Gamma_{h,C}^*$ of the critical boundary sets play a substantial role. Instead of being only bounded, the sum of critical measures over the whole time interval has to fulfill the following assumption.

Assumption 4.7.

$$\sum_{n=0}^{N-1} \text{meas}\Gamma_{h,C}^*[t_n, t_{n+1}] \leq C_{\text{meas}}$$

The assumption reflects the expectable behavior of the possible contact boundaries to become critical only at a finite number of time intervals. Hence, the sum of critical measures over time should be uniformly bounded, which corresponds to the earlier Assumption 4.3 of bounded variation for the continuous solution.

By means of Hölder's inequality, Assumption 4.3 leads to

$$\begin{aligned} & \sum_{n=0}^{N-1} S^\alpha(\mathbf{u}_h, [t_n, t_{n+1}]) \\ &= \sum_{n=0}^{N-1} (h \cdot \text{meas}\Gamma_{h,C}^*[t_n, t_{n+1}])^{\alpha(\frac{1}{2} - \frac{1}{p})} R^{\frac{1+\alpha}{2}}(\mathbf{u}_h, [t_n, t_{n+1}]) \\ &\leq \left(\sum_{n=0}^{N-1} \text{meas}\Gamma_{h,C}^*[t_n, t_{n+1}]^{\frac{2\alpha}{1-\alpha}(\frac{1}{2} - \frac{1}{p})} \right)^{\frac{1-\alpha}{2}} \left(\sum_{n=0}^{N-1} R(\mathbf{u}_h, [t_n, t_{n+1}]) \right)^{\frac{1+\alpha}{2}} h^{\alpha(\frac{1}{2} - \frac{1}{p})} \\ &= \left(\sum_{n=0}^{N-1} \text{meas}\Gamma_{h,C}^*[t_n, t_{n+1}]^{\frac{2\alpha}{1-\alpha}(\frac{1}{2} - \frac{1}{p})} \right)^{\frac{1-\alpha}{2}} R(\mathbf{u}_h, [t_0, T])^{\frac{1+\alpha}{2}} h^{\alpha(\frac{1}{2} - \frac{1}{p})}. \end{aligned}$$

Setting $\alpha = \frac{1}{2(1-1/p)}$ and using the novel Assumption 4.7 for the critical

measures gives

$$\begin{aligned}
& \sum_{n=0}^{N-1} S^\alpha(\mathbf{u}_h, [t_n, t_{n+1}]) \\
& \leq \left(\sum_{n=0}^{N-1} \text{meas}\Gamma_{h,C}^*[t_n, t_{n+1}] \right)^{\frac{1-\alpha}{2}} R(\mathbf{u}_h, [t_0, T])^{\frac{1+\alpha}{2}} \cdot h^{\frac{1-\alpha}{2}} \\
& \leq C_{\text{meas}} R(\mathbf{u}_h, [t_0, T])^{\frac{1+\alpha}{2}} \cdot h^{\frac{1-\alpha}{2}}.
\end{aligned}$$

Finally, by means of a discrete perturbation result for the algorithmic solution of $(\text{N-CS}++)_h$, Lady Windermere's fan yields a sharpened convergence result for the improved contact-stabilized Newmark method:

$$\|\bar{\mathbf{u}}_h^N - \bar{\mathbf{u}}_h(T)\|_E = O(\tau^{1/2}) + O\left((h\tau)^{\frac{1}{2}} \left(1 - \frac{1}{2(1-1/p)}\right)\right). \quad (45)$$

In the special case of space dimension $d = 3 < \left(\frac{1}{2} - \frac{1}{p}\right)^{-1}$, this estimate provides the global convergence of the Newmark algorithm in time and space for $\tau \rightarrow 0$ with an order less than $1/5$.

5 Conclusion

In this paper, a consistency theory for an improved version of the contact-stabilized Newmark method after discretization in time and in space has been worked out. The error estimate in physical energy norm is given for solutions of bounded total variation and depends on the measure of the boundary parts, where active contact is gained in the considered timestep. Under an additional assumption for these critical sets, the consistency result even allows to prove the convergence of the improved contact-stabilized Newmark method if the timestep tends to zero.

Acknowledgment. The author thanks Anton Schiela, Technical University of Berlin, and especially Peter Deuffhard, Zuse Institute Berlin, for helpful discussions on the topic of this paper.

References

- [1] R. A. Adams and J. J. F. Fournier. *Sobolev Spaces*, volume 140 of *Pure and Applied Mathematics*. Elsevier, Oxford Amsterdam, 2nd edition, 2003.
- [2] J. Ahn and D. E. Stewart. Dynamic frictionless contact in linear viscoelasticity. *IMA J. Numer. Anal.*, 29(1):43–71, 2009.

- [3] P. Deuffhard, R. Krause, and S. Ertel. A contact-stabilized Newmark method for dynamical contact problems. *Internat. J. Numer. Methods Engrg.*, 73(9):1274–1290, 2007.
- [4] P. Deuffhard and M. Weiser. *Numerische Mathematik 3: Adaptive Lösung partieller Differentialgleichungen*. de Gruyter Lehrbuch. Walter de Gruyter, Berlin, 2011.
- [5] G. Duvaut and J. L. Lions. *Inequalities in Mechanics and Physics*. Springer-Verlag, Berlin Heidelberg New York, 1976.
- [6] C. Eck. *Existenz und Regularität der Lösungen für Kontaktprobleme mit Reibung*. PhD thesis, Universität Stuttgart, 1996.
- [7] I. Ekeland and R. Temam. *Convex Analysis and Variational Problems*. Number 28 in Classics in Applied Mathematics. Society for Industrial and Applied Mathematics, Amsterdam New York, 1987.
- [8] C. Gräser and R. Kornhuber. Multigrid methods for obstacle problems. *J. Comp. Math.*, 27(1):1–44, 2009.
- [9] E. Hairer, C. Lubich, and G. Wanner. *Geometric Numerical Integration. Structure-Preserving Algorithms for Ordinary Differential Equations*. Number 31 in Computational Mathematics. Springer-Verlag, Berlin Heidelberg New York, 2nd edition, 2006.
- [10] E. Hairer, S. P. Nørsett, and G. Wanner. *Solving Ordinary Differential Equations I. Nonstiff Problems*. Springer-Verlag, Berlin Heidelberg New York, 1987.
- [11] C. Kane, E. A. Repetto, M. Ortiz, and J. E. Marsden. Finite element analysis of nonsmooth contact. *Comput. Methods Appl. Mech. Engrg.*, 180(1-2):1–26, 1999.
- [12] N. Kikuchi and J. T. Oden. *Contact Problems in Elasticity*. SIAM Studies in Applied Mathematics. Society for Industrial and Applied Mathematics, Philadelphia, 1988.
- [13] C. Klapproth. *Adaptive numerical integration for dynamical contact problems*. PhD thesis, Freie Universität Berlin, 2011. Also published at Cuvillier Verl. Göttingen.
- [14] C. Klapproth, P. Deuffhard, and A. Schiela. A perturbation result for dynamical contact problems. Preprint, Zuse-Institut Berlin, 2008. ZR 08-27. <http://opus.kobv.de/zib/volltexte/2008/1113/>.
- [15] C. Klapproth, P. Deuffhard, and A. Schiela. A perturbation result for dynamical contact problems. *Numer. Math. Theor. Meth. Appl.*, 2(3):237–257, 2009.

- [16] C. Klapproth, A. Schiela, and P. Deuffhard. Consistency results on Newmark methods for dynamical contact problems. *Numer. Math.*, 116(1):65–94, 2010.
- [17] R. Kornhuber. *Adaptive Monotone Multigrid Methods for Nonlinear Variational Problems*. B. G. Teubner, Stuttgart, 1997.
- [18] R. Kornhuber and R. Krause. Adaptive multigrid methods for Signorini’s problem in linear elasticity. *Comput. Vis. Sci.*, 4(1):9–20, 2001.
- [19] R. Kornhuber, R. Krause, O. Sander, P. Deuffhard, and S. Ertel. A monotone multigrid solver for two body contact problems in biomechanics. *Comput. Vis. Sci.*, 11(1):3–15, 2008.
- [20] R. Krause. *Monotone Multigrid Methods for Signorini’s Problem with Friction*. PhD thesis, Freie Universität Berlin, 2000. <http://www.diss.fu-berlin.de/2001/240/indexe.html>.
- [21] R. Krause and M. Walloth. Presentation and comparison of selected algorithms for dynamic contact based on the Newmark scheme. Technical report, Institute of Computational Science, University of Lugano, 2009. ICS Preprint 2009-08.
- [22] T. A. Laursen. *Computational Contact and Impact Mechanics. Fundamentals of Modeling Interfacial Phenomena in Nonlinear Finite Element Analysis*. Springer-Verlag, Berlin Heidelberg New York, 2003.
- [23] T. A. Laursen and V. Chawla. Design of energy conserving algorithms for frictionless dynamic contact problems. *Internat. J. Numer. Methods Engrg.*, 40(5):863–886, 1997.
- [24] N. M. Newmark. A method of computation for structural dynamics. *J. Engrg. Mech. Div., ASCE*, pages 67–94, 1959.
- [25] O. Sander, C. Klapproth, J. Youett, R. Kornhuber, and P. Deuffhard. Towards an efficient numerical simulation of complex 3D knee joint motion. *SIAM J. Sci. Comput.*, submitted. Preprint ZR 12-06, Zuse Institute Berlin, 2012, <http://opus4.kobv.de/opus4-zib/frontdoor/index/index/docId/1451>.
- [26] E. Schechter. *Handbook of Analysis and Its Foundations*. Academic Press, San Diego, 1997.
- [27] B. Wohlmuth and R. Krause. Monotone methods on nonmatching grids for nonlinear contact problems. *SIAM J. Sci. Comput.*, 25(1):324–347, 2003.
- [28] E. Zeidler. *Nonlinear Functional Analysis and Applications II*. Springer Verlag, Berlin Heidelberg New York, 1986.