Computer Methods in Applied Mechanics and Engineering 418 (2024) 116496

Contents lists available at ScienceDirect

Comput. Methods Appl. Mech. Engrg.

journal homepage: www.elsevier.com/locate/cma

Lobatto-type variational integrators for mechanical systems with frictional contact

Giuseppe Capobianco [a,*], Jonas Harsch [b], Sigrid Leyendecker [a]

[a] Institute of Applied Dynamics, Friedrich-Alexander-Universität Erlangen-Nürnberg, Immerwahrstrasse 1, 91058 Erlangen, Germany
[b] Institute for Nonlinear Mechanics, University of Stuttgart, Pfaffenwaldring 9, 70569 Stuttgart, Germany

ARTICLE INFO

Keywords:
Nonsmooth contact dynamics
Variational integrator
Unilateral constraints
Coulomb friction

ABSTRACT

This paper introduces a family of Lobatto IIIA-IIIB methods for simulating mechanical systems with frictional contact. These methods extend the existing schemes by addressing both bilateral and unilateral constraints, as well as set-valued Coulomb friction. The Lobatto IIIA-IIIB methods presented in this paper make a substantial contribution to the ongoing endeavor of developing event-capturing versions of high-order schemes.

By stating a generalized version of the principle of virtual action for nonsmooth mechanical systems, the integrators are derived from an appropriate discretization of it. The discrete contact laws are found by discretizing the impenetrabilty condition as well as the Newton-type frictional impact laws in an event-capturing way. The presented discrete contact laws exhibit no contact penetration and satisfy the involved unilateral constraints both on position and on velocity level. This behavior is showcased using benchmark examples.

1. Introduction

Variational integrators are well-suited for the simulation of (bilaterally) constrained mechanical systems due to several key advantages they offer, see [1]. Their good long-term energetic consistency, which bears on their symplecticity, helps to maintain the accuracy and stability of the simulation over long time periods. By their geometric nature, variational integrators are able to accurately preserve bilateral constraints on both position and velocity level. Moreover, variational integrators with high orders of accuracy have been developed. A very prominent example is the family of Lobatto IIIA-IIIB methods, which, depending on the number of stages s, are convergent with order $2s-2$, see [2,3].

In this paper, we present a family of Lobatto IIIA-IIIB methods for the simulation of mechanical systems with frictional contact. In addition to bilateral constraints, the presented methods can cope with unilateral constraints and set-valued Coulomb friction, which are the two main ingredients for the description of frictional contact. Hence, the presented Lobatto IIIA-IIIB methods are an extension of the existing ones.

It is well known that the dynamics of multibody systems with frictional contact is nonsmooth, i.e., these systems can for example exhibit velocity jumps due to impacts whenever a unilateral constraint becomes active, see [4–6]. For the numerical simulation of nonsmooth mechanical systems, two classes of schemes can be distinguished, namely the event-driven and the event-capturing schemes, confer [7] for an overview. The distinguishing property of event-driven schemes is that they accurately resolve every discontinuity point, i.e., every time instant at which for example a slip-stick transition or an impact occurs is precisely determined, see for example [8,9]. In contrast, the event-capturing schemes approximate all discontinuities occurring within a time

* Corresponding author.
E-mail address: giuseppe.capobianco@fau.de (G. Capobianco).

https://doi.org/10.1016/j.cma.2023.116496
Received 24 July 2023; Received in revised form 26 September 2023; Accepted 27 September 2023
Available online 14 October 2023

step qualitatively, e.g., [6,10,11]. Accurately resolving every single discontinuity point is computationally very expensive, especially if the system's dynamics exhibits a large number of discontinuities. This is typically the case for engineering systems, since they have many possible contact points. There are even cases where the dynamics shows an accumulation point, i.e., an infinite number of discontinuities occurring in a finite interval of time. It is immediately clear that an event-driven scheme will not be able to overcome such an accumulation point. It is the major advantage of event-capturing schemes that they can overcome accumulation points and that they can efficiently simulate motions with many discontinuities. The presented Lobatto IIIA-IIIB methods are all event-capturing and hence have this beneficial property.

There has been a big endeavor to develop event-capturing versions of schemes with order higher than one. For example nonsmooth versions of the generalized-α scheme [12,13] or of the RATTLE scheme [14], which are both event-capturing extensions of second-order schemes, have been devised. A family of nonsmooth variants of high-order schemes, i.e., with orders higher than two, were presented in [15]. All mentioned schemes, are nonsmooth versions of schemes with orders that are higher than one. For these, it is well known that the event-capturing way of discretizing the contact laws leads to schemes that are only convergent of order one for motions with discontinuities. However, for motions with persistent contact, the higher order of the original method is retrieved, see [15].

The presented Lobatto IIIA-IIIB schemes are a substantial contribution to the endeavor of developing event-capturing versions of high order schemes, as they cover a complete family of s-stage schemes of arbitrary high order, i.e., of order $2s-2$. Moreover, the presented schemes can cope with both bilateral constraints on position and velocity level, and hence with holonomic and nonholonomic constraints. In contrast, [15] only includes position level constraints. For the derivation of the schemes, we have generalized the principle of virtual action presented in [10] such that the dynamics of the system is described in terms of the generalized coordinates, the velocities and the momenta as independent variables. This principle of virtual action serves as the starting point for the discretization. In a first step, as it is usual for variational integrators, a discrete version of the principle of virtual action is derived. The time-stepping scheme then results as the necessary and sufficient conditions of the discrete principle of virtual action. Finally, the contact laws are discretized in an event-capturing manner. For that, we capitalized on the particular structure of the Butcher tableau of the Lobatto IIIB scheme. The developed discrete contact law numerically satisfies the unilateral constraints on position and velocity level and fulfills a Newton-type impact law in an integral way over the whole time step. Hence, the presented family of Lobatto schemes do not exhibit contact penetration, i.e., the presented schemes do not show unilateral (and bilateral) constraint drift. Moreover, the presented discrete contact law is consistent with the discretization of the bilateral constraints. This implies that for motions where the contact is always closed, e.g., rolling wheels in vehicle dynamics simulations, the convergence rate $2s-2$ is retrieved. As all event-capturing schemes, also the presented schemes exhibit the drop in convergence order for motions with discontinuities. This behavior is showcased in the last numerical example of this paper.

The paper is organized as follows. In Section 2, the mathematical framework for the description of nonsmooth motion arising due to the velocity jumps is introduced. The subsequent Section 3 briefly revises the mechanics of systems with frictional contact in continuous time and introduces the principle of virtual action. In Section 4, the family of Lobatto IIIA/IIB schemes is derived. Section 5 summarizes the scheme and concisely contains all the information needed for the implementation. To validate the schemes numerically, several benchmark examples are studied in Section 6. Concluding remarks can be found in Section 7.

2. Nonsmooth functions and the differential measure

This section is meant as a brief summary of the mathematical properties of special functions of locally bounded variations (SLBV functions) and sets the notation used within this paper. For a more complete treatise of this topic, we refer the reader to Chapter 4 in [4] as well as to [16].

Let $\boldsymbol{x} : \mathbb{R} \to \mathbb{R}^n$ be a function of locally bounded variation. It is well known, see [16], that at each discontinuity point t_i of $\boldsymbol{x}$ the right limit $\boldsymbol{x}^+(t_i)$ and the left limit $\boldsymbol{x}^-(t_i)$ exist. Moreover, the derivative $\dot{\boldsymbol{x}}$ of $\boldsymbol{x}$ exists almost everywhere (a.e.). A function of locally bounded variation can be decomposed as $\boldsymbol{x} = \boldsymbol{x}_A + \boldsymbol{x}_C + \boldsymbol{x}_S$, where $\boldsymbol{x}_A$ is absolutely continuous, $\boldsymbol{x}_C$ is a piecewise constant function with a countable set of discontinuities, and $\boldsymbol{x}_S$ is a singular function (a continuous function with $\dot{\boldsymbol{x}}_S = 0$ almost everywhere, e.g., the Cantor function). This decomposition is unique up to an additive constant. As introduced by [17], the special functions of locally bounded variations are defined as functions of locally bounded variations with no singular part, i.e., as the subspace defined by $\boldsymbol{x}_S = 0$. Finally, the (Riemann–Stieltjes) integral can be associated to $\boldsymbol{x}$ and has the property

$$\boldsymbol{x}^+(t_r) - \boldsymbol{x}^-(t_l) = \int_{[t_l,t_r]} \mathrm{d}\boldsymbol{x}, \tag{1}$$

where $\mathrm{d}\boldsymbol{x}$ is called the differential measure of $\boldsymbol{x}$. For any continuously differentiable function $\boldsymbol{y} : \mathbb{R} \to \mathbb{R}^n$, the integration by parts formula

$$\int_{[t_l,t_r]} \boldsymbol{x}^{\mathrm{T}} \dot{\boldsymbol{y}}\, \mathrm{d}t = -\int_{[t_l,t_r]} \boldsymbol{y}^{\mathrm{T}} \mathrm{d}\boldsymbol{x} + \boldsymbol{x}^+(t_r)\boldsymbol{y}(t_r) - \boldsymbol{x}^-(t_l)\boldsymbol{y}(t_l) \tag{2}$$

is a consequence of Corollary 11.4 in [16].

Let $\mathcal{D}_{\boldsymbol{x}} = \{t_1, t_2, \dots\}$ be the set of discontinuity points of a SLBV function $\boldsymbol{x}$, then the differential measure of $\boldsymbol{x}$ is

$$\mathrm{d}\boldsymbol{x} = \dot{\boldsymbol{x}}\,\mathrm{d}t + \sum_{t_i \in \mathcal{D}_{\boldsymbol{x}}} [\boldsymbol{x}^+ - \boldsymbol{x}^-]\mathrm{d}\delta_{t_i}, \tag{3}$$

where $\mathrm{d}t$ is the Lebesgue measure on $\mathbb{R}$ and $\mathrm{d}\delta_{t_i}$ is the Dirac point measure at t_i. It is sometimes useful to introduce the measure $\mathrm{d}\mu$ and the density $\boldsymbol{x}'$ of $\boldsymbol{x}$ as

$$\mathrm{d}\boldsymbol{x} = \boldsymbol{x}'\,\mathrm{d}\mu \quad \text{with} \quad \mathrm{d}\mu = \mathrm{d}t + \sum_{t_i \in D_{\boldsymbol{x}}} \mathrm{d}\delta_{t_i}. \tag{4}$$

It follows from the comparison of (3) with (4) that $\boldsymbol{x}' = \dot{\boldsymbol{x}}$ almost everywhere and $\boldsymbol{x}' = \boldsymbol{x}^+ - \boldsymbol{x}^-$ at the discontinuity points of $\boldsymbol{x}$.

3. Mechanical systems with frictional contact

For all times t, let $\boldsymbol{q}(t) \in \mathbb{R}^{n_q}$ denote the generalized coordinates of a finite-dimensional mechanical system. Since the velocity of the system may jump due to impacts, the generalized velocity $\boldsymbol{v}$ is assumed to be an SLBV function. The positions and the velocities are linked by the kinematic equation

$$\mathrm{d}\boldsymbol{q} = \boldsymbol{v}\,\mathrm{d}t \quad \Longleftrightarrow \quad \dot{\boldsymbol{q}} = \boldsymbol{v} \quad \text{a.e.} \tag{5}$$

Let $L(t, \boldsymbol{q}, \boldsymbol{v}) = T(t, \boldsymbol{q}, \boldsymbol{v}) - V(t, \boldsymbol{q})$ be the Lagrangian of the system, which is defined as the difference between the system's kinetic energy T and its potential energy V. With that, the generalized momentum of the system is defined as

$$\boldsymbol{\pi}(t) = \left(\frac{\partial L}{\partial \boldsymbol{v}}\right)^{\mathrm{T}}(t, \boldsymbol{q}(t), \boldsymbol{v}(t)), \tag{6}$$

which inherits the jump discontinuities of the velocity and is therefore an SLBV function. The equations of motion of a mechanical system with frictional contact take the form of an equality of measures and links the change in momentum to the forces acting on the system:

$$\mathrm{d}\boldsymbol{\pi} - \left(\frac{\partial L}{\partial \boldsymbol{q}}\right)^{\mathrm{T}} \mathrm{d}t = \boldsymbol{f}\,\mathrm{d}t + \boldsymbol{W}\,\mathrm{d}\boldsymbol{P}. \tag{7}$$

Herein, $\boldsymbol{W}(t, \boldsymbol{q})\,\mathrm{d}\boldsymbol{P}$ are the constraint and contact forces with generalized force directions $\boldsymbol{W}$, and $\boldsymbol{f}(t, \boldsymbol{q}, \boldsymbol{v})\,\mathrm{d}t$ describes all forces which are not contemplated by L and are not constraint or contact forces. The percussion measure $\mathrm{d}\boldsymbol{P}$ combines the effects of nonimpulsive forces λ and impulsive forces $\boldsymbol{\Lambda}$ in the sense that

$$\mathrm{d}\boldsymbol{P} = \lambda\,\mathrm{d}t + \sum_{t_i \in D_{\boldsymbol{P}}} \boldsymbol{\Lambda}\,\mathrm{d}\delta_{t_i} = \boldsymbol{R}\,\mathrm{d}\mu, \tag{8}$$

in accordance with (3) and (4). Indeed, using this and (3) for the momentum measure $\mathrm{d}\boldsymbol{\pi}$, the equality of measures (7) can be reformulated as

$$\left[\dot{\boldsymbol{\pi}} - \left(\frac{\partial L}{\partial \boldsymbol{q}}\right)^{\mathrm{T}} - \boldsymbol{f} - \boldsymbol{W}\lambda\right]\mathrm{d}t + \sum_{t_i \in D}\left[\boldsymbol{\pi}^+ - \boldsymbol{\pi}^- - \boldsymbol{W}\boldsymbol{\Lambda}\right]\mathrm{d}\delta_{t_i} = \boldsymbol{0}, \tag{9}$$

with the set of discontinuity points $D := D_{\boldsymbol{P}} = D_{\boldsymbol{\pi}}$. Hence, using (6), the equality of measure is satisfied if and only if Lagrange's equations of the second kind

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \boldsymbol{v}}\right)^{\mathrm{T}} - \left(\frac{\partial L}{\partial \boldsymbol{q}}\right)^{\mathrm{T}} = \boldsymbol{f} + \boldsymbol{W}\lambda \tag{10}$$

are satisfied almost everywhere in time and the impact equations

$$\boldsymbol{\pi}^+ - \boldsymbol{\pi}^- = \boldsymbol{W}\boldsymbol{\Lambda} \tag{11}$$

describe the jump discontinuities of the momentum. Even though this splitting of the equations of motion into an impulsive part and a nonimpulsive part is very illustrative, it is not used in the remainder of this paper. Instead, we will exploit that $\boldsymbol{R} = \lambda$ almost everywhere and $\boldsymbol{R} = \boldsymbol{\Lambda}$ at the discontinuity points $t_i \in D$. This allows for example to formulate the contact laws jointly for phases of persistent contact, where the velocity is continuous, and for the impact time instants, where the velocity jumps.

To address the constraint and contact forces separately, the notation

$$\boldsymbol{W}\,\mathrm{d}\boldsymbol{P} = \boldsymbol{W}_g\,\mathrm{d}\boldsymbol{P}_g + \boldsymbol{W}_\gamma\,\mathrm{d}\boldsymbol{P}_\gamma + \boldsymbol{W}_N\,\mathrm{d}\boldsymbol{P}_N + \boldsymbol{W}_F\,\mathrm{d}\boldsymbol{P}_F \tag{12}$$

is introduced, i.e., we split[1] the percussion measures as $\mathrm{d}\boldsymbol{P} = (\mathrm{d}\boldsymbol{P}_g, \mathrm{d}\boldsymbol{P}_\gamma, \mathrm{d}\boldsymbol{P}_N, \mathrm{d}\boldsymbol{P}_F)$ and partition the generalized force directions $\boldsymbol{W} = (\boldsymbol{W}_g\ \boldsymbol{W}_\gamma\ \boldsymbol{W}_N\ \boldsymbol{W}_F)$ accordingly. The force measures $\boldsymbol{W}_g\mathrm{d}\boldsymbol{P}_g$ and $\boldsymbol{W}_\gamma\mathrm{d}\boldsymbol{P}_\gamma$ in (12) are the ideal constraint forces of the position and velocity level constraints

$$\boldsymbol{g}(t, \boldsymbol{q}) = \boldsymbol{0}, \qquad \text{and} \qquad \boldsymbol{\gamma}(t, \boldsymbol{q}, \boldsymbol{v}) = \boldsymbol{W}_\gamma^{\mathrm{T}}(t, \boldsymbol{q})\,\boldsymbol{v} + \chi_\gamma(t, \boldsymbol{q}) = \boldsymbol{0}, \tag{13}$$

respectively. The generalized force directions of the bilateral constraints are given by

$$\boldsymbol{W}_g^{\mathrm{T}} = \frac{\partial \boldsymbol{g}}{\partial \boldsymbol{q}} \quad \text{and} \quad \boldsymbol{W}_\gamma^{\mathrm{T}} = \frac{\partial \boldsymbol{\gamma}}{\partial \boldsymbol{v}}. \tag{14}$$

[1] We use the following notation: For $\boldsymbol{x} \in \mathbb{R}^n$ and $\boldsymbol{y} \in \mathbb{R}^m$, $(\boldsymbol{x}, \boldsymbol{y}) := (\boldsymbol{x}^{\mathrm{T}}\ \boldsymbol{y}^{\mathrm{T}})^{\mathrm{T}} \in \mathbb{R}^{n+m}$.

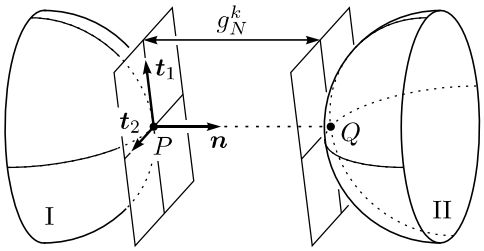

Fig. 1. Kinematics of the contact k consisting of the contact points P and Q on the bodies I and II, respectively.

The remaining parts of (12), i.e., $\boldsymbol{W}_N \mathrm{d}\boldsymbol{P}_N + \boldsymbol{W}_F \mathrm{d}\boldsymbol{P}_F$, are the forces describing the frictional contact. Consider a pair of contact points P and Q on either contacting body, which we will simply refer to as contact k, see Fig. 1. Let the tangent planes of the pair of contact points be spanned by the vectors $\boldsymbol{t}_1$ and $\boldsymbol{t}_2$, and let $\boldsymbol{n}$ denote the outward pointing normal vector at P. In order to describe the distance between the tangent planes, the gap function $g_N^k = \boldsymbol{n}^{\mathrm{T}} \boldsymbol{r}_{PQ}$ is introduced, where $\boldsymbol{r}_{PQ}$ denotes the vector pointing from P to Q. When $g_N^k > 0$, the contacting bodies are separated and we say that the contact is open. Similarly, we speak of a closed contact if $g_N^k = 0$ and the bodies are touching, and of a penetrating contact if the bodies are penetrating each other and $g_N^k < 0$. The impenetrability of contact k is enforced by the ideal unilateral constraint formulated as Signorini's law

$$g_N^k(t,\boldsymbol{q}) \in \mathcal{N}_{\mathbb{R}_0^-}(-R_N^k) \quad \Longleftrightarrow \quad g_N^k(t,\boldsymbol{q}) \geq 0,\ R_N^k \geq 0,\ g_N^k(t,\boldsymbol{q})\, R_N^k = 0, \tag{15}$$

where $\mathcal{N}_{\mathbb{R}_0^-}(x)$ denotes the normal cone to the set $\mathbb{R}_0^- = \{z \in \mathbb{R} \mid z \leq 0\}$ of non-positive numbers evaluated at x. The equivalence is easily proven using the definition of the normal cone, which for any nonempty convex set $C \subset \mathbb{R}^n$ reads as

$$\mathcal{N}_C(\boldsymbol{x}) = \{\boldsymbol{y} \in \mathbb{R}^n \mid \boldsymbol{y}^{\mathrm{T}}(\boldsymbol{x}^* - \boldsymbol{x}) \leq 0 \quad \forall \boldsymbol{x}^* \in C\}$$

for any $\boldsymbol{x} \in C$. Otherwise, for $\boldsymbol{x} \notin C$, the normal cone is empty, i.e., $\mathcal{N}_C(\boldsymbol{x}) = \emptyset$. A look at Signorini's law, see also Fig. 2, reveals that due to the complementarity condition the contact force can only be positive if $g_N^k(t,\boldsymbol{q}) = 0$, i.e., if the contact is closed and the contacting bodies touch. Moreover, it is clear that the contact force is introduced such that for positive values it pushes the contacting bodies apart.

In order to formulate an impact law, we first combine the gap functions of all n_N contacts into a vector $\boldsymbol{g}_N(t,\boldsymbol{q}) \in \mathbb{R}^{n_N}$ and define the gap velocity

$$\dot{\boldsymbol{g}}_N(t,\boldsymbol{q},\boldsymbol{v}) = \frac{\partial \boldsymbol{g}_N}{\partial \boldsymbol{q}}(t,\boldsymbol{q})\,\dot{\boldsymbol{q}} + \frac{\partial \boldsymbol{g}_N}{\partial t}(t,\boldsymbol{q}) \overset{(5)}{=} \boldsymbol{W}_N^{\mathrm{T}}(t,\boldsymbol{q})\boldsymbol{v} + \boldsymbol{\chi}_N(t,\boldsymbol{q}), \tag{16}$$

where we have introduced the generalized force directions

$$\boldsymbol{W}_N^{\mathrm{T}} = \frac{\partial \dot{\boldsymbol{g}}_N}{\partial \boldsymbol{v}} = \frac{\partial \boldsymbol{g}_N}{\partial \boldsymbol{q}} \tag{17}$$

arising in (12). Moreover, we define the set of active contacts as

$$A(t,\boldsymbol{q}) = \left\{k = 1,\dots,n_N \;\middle|\; g_N^k(t,\boldsymbol{q}) \leq 0\right\}, \tag{18}$$

together with its complement $\bar{A} = \{1,\dots,n_N\} \setminus A$, the set of inactive (open) contacts. A Newton-type impact law is chosen. Namely,

$$\begin{aligned} k \in A: &\quad \xi_N^k(t,\boldsymbol{q},\boldsymbol{v}^-,\boldsymbol{v}^+) \in \mathcal{N}_{\mathbb{R}_0^-}(-R_N^k) \\ k \in \bar{A}: &\quad R_N^k = 0 \end{aligned} \tag{19}$$

where the kinematic quantity

$$\xi_N^k(t,\boldsymbol{q},\boldsymbol{v}^-,\boldsymbol{v}^+) = \dot{g}_N^k(t,\boldsymbol{q},\boldsymbol{v}^+) + e_N^k\, \dot{g}_N^k(t,\boldsymbol{q},\boldsymbol{v}^-) \tag{20}$$

is introduced and $e_N^k \in [0,1]$ is the restitution coefficient for the normal impact. See Fig. 2 for a graphical representation of the impact law. This impact law ensures that if the contact k is open, i.e., $k \in \bar{A}$, the contact force is zero. Moreover, it follows from the definition of the normal cone that for an active contact, i.e., $k \in A$, the impact law is equivalent to the inequality complementarity condition

$$\xi_N^k(t,\boldsymbol{q},\boldsymbol{v}^-,\boldsymbol{v}^+) \geq 0, \quad R_N^k \geq 0, \quad \xi_N^k(t,\boldsymbol{q},\boldsymbol{v}^-,\boldsymbol{v}^+)\, R_N^k = 0. \tag{21}$$

This implies, that whenever there is a contact force $R_N^k > 0$, then

$$\dot{g}_N^k(t,\boldsymbol{q},\boldsymbol{v}^+) = -e_N^k\, \dot{g}_N^k(t,\boldsymbol{q},\boldsymbol{v}^-) \tag{22}$$

due to the complementarity condition. Moreover, it contains the exceptional case where active contacts $k \in A$ with $R_N^k = 0$ allow for

$$\dot{g}_N^k(t,\boldsymbol{q},\boldsymbol{v}^+) \geq -e_N^k\, \dot{g}_N^k(t,\boldsymbol{q},\boldsymbol{v}^-), \tag{23}$$

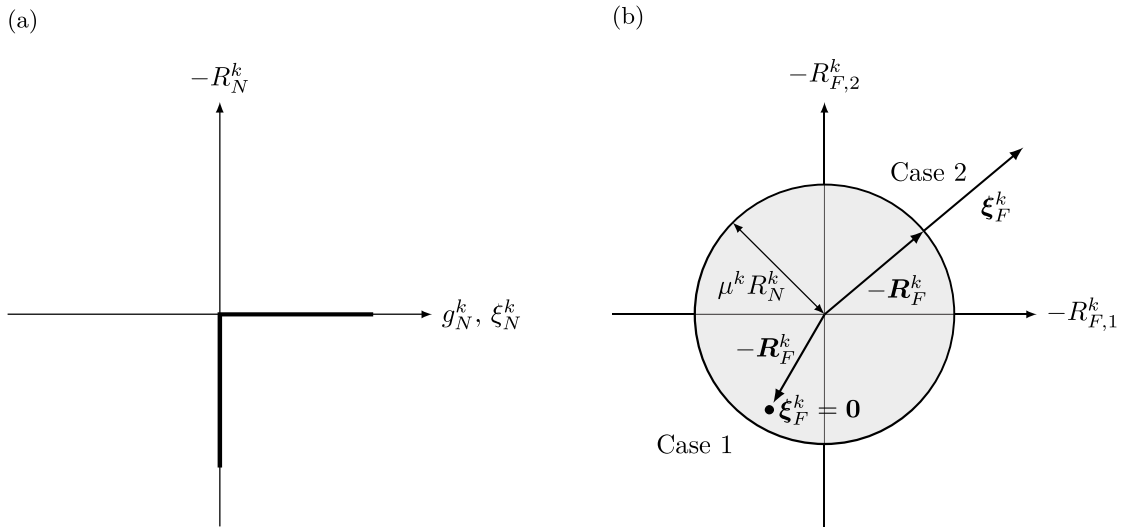

Fig. 2. Graphical representation of the contact laws. (a) Signorini's law $g_N^k(t,\boldsymbol{q}) \in \mathcal{N}_{\mathbb{R}_0^-}(-R_N^k)$ as well as the Newton-type impact law $\xi_N^k \in \mathcal{N}_{\mathbb{R}_0^-}(-R_N^k)$. (b) Frictional impact law with isotropic Coulomb friction $\boldsymbol{\xi}_F^k \in \mathcal{N}_{C_F^k(R_N^k)}(-\boldsymbol{R}_F^k)$, where $\boldsymbol{R}_F^k = (R_{F,1}^k, R_{F,2}^k)$. Case 1: For $\|\boldsymbol{R}_F^k\| \le \mu^k R_N^k$ this model enforces the constraint $\boldsymbol{\xi}_F^k = \mathbf{0}$. Case 2: $\|\boldsymbol{\xi}_F^k\| > 0$ results in the force law $\boldsymbol{R}_F^k = -\mu R_N^k \boldsymbol{\xi}_F^k / \|\boldsymbol{\xi}_F^k\|$.

which can happen for multi-contact configurations, see [18,19]. Hence, the contact law (19) models a Newton-type impact law. But not only that. It can be shown that the Newton-type impact law (19) implies Signorini's law (15). In fact, since the transition from open contact $g_N^k > 0$ to active contact $g_N^k \le 0$ implies $\dot{g}_N^k(t,\boldsymbol{q},\boldsymbol{v}^-) \le 0$ at the time instant where the contact k gets active, the first inequality in (21) implies

$$k \in A(t,\boldsymbol{q}): \quad \dot{g}_N^k(t,\boldsymbol{q},\boldsymbol{v}^+) \ge 0 \tag{24}$$

which can be recognized as a unilateral constraint on velocity level. It is a consequence of the viability lemma of Moreau, Proposition 2.4 in [6], that (24) implies the unilateral constraint on position level. Hence, in continuous time, it suffices to impose the contact law (19) to model both impenetrability and Newton-type impact. In discrete time however, this will not be sufficient as numerical drift would lead to contact penetration.

Finally, set-valued Coulomb friction is added to the contact model. For that, the tangent velocity

$$\boldsymbol{\gamma}_F^k = \begin{pmatrix} \boldsymbol{t}_1^{\mathrm{T}} \dot{\boldsymbol{r}}_{PQ} \\ \boldsymbol{t}_2^{\mathrm{T}} \dot{\boldsymbol{r}}_{PQ} \end{pmatrix}$$

describing the relative velocity of the pair of contact points in the tangent plane is introduced, cf. Fig. 1. The generalized force direction of the friction forces is defined by

$$\boldsymbol{W}_F = (\boldsymbol{W}_F^1 \ \dots \ \boldsymbol{W}_F^{n_N}) \quad \text{with} \quad (\boldsymbol{W}_F^k)^{\mathrm{T}} = \frac{\partial \boldsymbol{\gamma}_F^k}{\partial \boldsymbol{v}} \tag{25}$$

and we use the notation $\mathrm{d}\boldsymbol{P}_F = \boldsymbol{R}_F \mathrm{d}\mu$ with $\boldsymbol{R}_F = (\boldsymbol{R}_F^1, \dots, \boldsymbol{R}_F^{n_F})$. Using the tangent velocity $\boldsymbol{\gamma}_F^k(t,\boldsymbol{q},\boldsymbol{v})$, the kinematic quantity

$$\boldsymbol{\xi}_F^k(t,\boldsymbol{q},\boldsymbol{v}^-,\boldsymbol{v}^+) = \boldsymbol{\gamma}_F^k(t,\boldsymbol{q},\boldsymbol{v}^+) + e_F^k\, \boldsymbol{\gamma}_F^k(t,\boldsymbol{q},\boldsymbol{v}^-) \tag{26}$$

is introduced with tangential restitution coefficient $e_F^k \in [0,1]$. With that, the Newton-type frictional impact law modeling Coulomb friction is introduced as

$$\forall k: \quad \boldsymbol{\xi}_F^k(t,\boldsymbol{q},\boldsymbol{v}^-,\boldsymbol{v}^+) \in \mathcal{N}_{C_F^k(R_N^k)}(-\boldsymbol{R}_F^k), \tag{27}$$

where for isotropic Coulomb friction, the set of admissible (negative) friction forces is

$$C_F^k(R_N^k) = B_2(\mu^k R_N^k), \quad \text{where} \quad B_2(r) = \left\{ \boldsymbol{x} \in \mathbb{R}^2 \;\middle|\; \|\boldsymbol{x}\| \le r \right\} \tag{28}$$

with friction coefficient μ^k. Similarly as before, $\mathcal{N}_{C_F^k(R_N^k)}$ denotes the normal cone to the set $C_F^k(R_N^k)$. A graphical representation of the impact law can be found in Fig. 2. It is worth noticing that the impact law (27) contains Coulomb's friction law for the time instants where no impact occurs and is hence sufficient to include friction into the contact model. To see that, note that if no impact occurs, the motion is continuous and $\boldsymbol{v} = \boldsymbol{v}^- = \boldsymbol{v}^+$. Hence, in that case the kinematic quantity (26) reduces to $(1 + e_F^k)\, \boldsymbol{\gamma}_F^k(t,\boldsymbol{q},\boldsymbol{v})$

implying that the friction law (27) is equivalent[2] to

$$\forall k: \quad \boldsymbol{\gamma}_F^k(t,\boldsymbol{q},\boldsymbol{v}) \in \mathcal{N}_{C_F^k(\lambda_N^k)}(-\boldsymbol{\lambda}_F^k) \tag{29}$$

since $\boldsymbol{R}_N^k = \boldsymbol{\lambda}_N^k$ and $\boldsymbol{R}_F^k = \boldsymbol{\lambda}_F^k$ for nonimpulsive time instants, $\mathcal{N}_{C_F^k(\lambda_N^k)}$ is a cone and $(1+e_F^k) > 0$. For more details about the intricacies of the introduced contact model we refer to [4,20].

Instead of characterizing the motion of the system for $t \in \mathcal{I} \subset \mathbb{R}$ using (5), (6) and (7), the motion can be described by the principle of virtual action in the form

$$\delta A := \int_{\mathcal{I}} \delta[L(t,\boldsymbol{q},\boldsymbol{v}) + \boldsymbol{\pi}^{\mathrm{T}}(\dot{\boldsymbol{q}} - \boldsymbol{v})]\,\mathrm{d}t + \int_{\mathcal{I}} \delta\boldsymbol{q}^{\mathrm{T}}(\boldsymbol{f}\,\mathrm{d}t + \boldsymbol{W}\,\mathrm{d}\boldsymbol{P}) = 0 \qquad \forall \delta\boldsymbol{q}, \delta\boldsymbol{v}, \delta\boldsymbol{\pi} \tag{30}$$

where the test function $\delta\boldsymbol{q}$ is assumed to vanish on the boundary of $\mathcal{I}$. This form of the principle of virtual action generalizes the one presented in [10] in the sense that in (30) the kinematic equation $\boldsymbol{v} = \dot{\boldsymbol{q}}$ is not explicitly used, but imposed using the generalized momentum $\boldsymbol{\pi}$ as a Lagrange multiplier. This way of imposing the kinematic equation with the help of the generalized momentum as a Lagrange multiplier is well known in classical mechanics literature, see for example [21,22], and is here applied analogously in the nonsmooth setting. To see that (30) indeed describes the motion of a mechanical system with frictional contact, we first perform the variation of the term in square brackets, i.e.,

$$\delta A = \int_{\mathcal{I}} \left[\frac{\partial L}{\partial \boldsymbol{q}}\delta\boldsymbol{q} + \frac{\partial L}{\partial \boldsymbol{v}}\delta\boldsymbol{v} + \delta\boldsymbol{\pi}^{\mathrm{T}}(\dot{\boldsymbol{q}} - \boldsymbol{v}) + \boldsymbol{\pi}^{\mathrm{T}}(\delta\dot{\boldsymbol{q}} - \delta\boldsymbol{v})\right]\mathrm{d}t + \int_{\mathcal{I}} \delta\boldsymbol{q}^{\mathrm{T}}(\boldsymbol{f}\,\mathrm{d}t + \boldsymbol{W}\,\mathrm{d}\boldsymbol{P}). \tag{31}$$

Using the integration by parts rule (2), the principle virtual action can be recast to

$$\delta A = \int_{\mathcal{I}} \left[\frac{\partial L}{\partial \boldsymbol{q}}\delta\boldsymbol{q} + \frac{\partial L}{\partial \boldsymbol{v}}\delta\boldsymbol{v} + \delta\boldsymbol{\pi}^{\mathrm{T}}(\dot{\boldsymbol{q}} - \boldsymbol{v}) - \boldsymbol{\pi}^{\mathrm{T}}\delta\boldsymbol{v}\right]\mathrm{d}t + \int_{\mathcal{I}} \delta\boldsymbol{q}^{\mathrm{T}}(-\mathrm{d}\boldsymbol{\pi} + \boldsymbol{f}\,\mathrm{d}t + \boldsymbol{W}\,\mathrm{d}\boldsymbol{P}) \tag{32a}$$

$$= \int_{\mathcal{I}} \delta\boldsymbol{\pi}^{\mathrm{T}}[\dot{\boldsymbol{q}} - \boldsymbol{v}]\,\mathrm{d}t + \int_{\mathcal{I}} \delta\boldsymbol{v}^{\mathrm{T}}\left[\left(\frac{\partial L}{\partial \boldsymbol{v}}\right)^{\mathrm{T}} - \boldsymbol{\pi}\right]\mathrm{d}t + \int_{\mathcal{I}} \delta\boldsymbol{q}^{\mathrm{T}}\left[\left(\frac{\partial L}{\partial \boldsymbol{q}}\right)^{\mathrm{T}}\mathrm{d}t - \mathrm{d}\boldsymbol{\pi} + \boldsymbol{f}\,\mathrm{d}t + \boldsymbol{W}\,\mathrm{d}\boldsymbol{P}\right] \tag{32b}$$

$$= 0 \qquad \forall \delta\boldsymbol{q}, \delta\boldsymbol{v}, \delta\boldsymbol{\pi},$$

where the boundary terms vanish because $\delta\boldsymbol{q} = 0$ on the boundary of $\mathcal{I}$. The motion satisfies the principle of virtual action if and only if the square brackets of (32b) vanish, which is equivalent to the motion satisfying (5), (6) and (7).

4. Lobatto IIIA-IIIB discretization

To derive a numerical scheme for the simulation of mechanical systems with frictional contact, we discretize the principle of virtual action (30). For that, we introduce equidistant time nodes t_n with constant time step $\Delta t = t_{n+1} - t_n$ $(n = 1, \dots, N-1)$ and choose $\mathcal{I} = [t_1, t_N)$.

We discretize the kinematic Eq. (5) using an s-stage Lobatto IIIA method

$$t_n^i = t_n + c_i\Delta t, \quad \boldsymbol{Q}_n^i = \boldsymbol{q}_n + \Delta t\sum_{j=1}^{s} a_{ij}\boldsymbol{V}_n^j, \quad \boldsymbol{q}_{n+1} = \boldsymbol{q}_n + \Delta t\sum_{j=1}^{s} b_j\boldsymbol{V}_n^j \tag{33}$$

where $i = 1, \dots, s$. Herein, $\boldsymbol{q}_n$ and $\boldsymbol{q}_{n+1}$ approximate the generalized coordinates $\boldsymbol{q}$ at t_n and t_{n+1}, respectively. Moreover, $\boldsymbol{Q}_n^i$ and $\boldsymbol{V}_n^i$ are the respective approximants of $\boldsymbol{q}$ and $\boldsymbol{v}$ at t_n^i. The conditions for defining the coefficients a_{ij}, b_i and c_i can be found in [23]. Especially, it holds that $a_{si} = b_i$ and $a_{1i} = 0$ for all stages i. Moreover, $c_1 = 0$ and $c_s = 1$. Hence, the Butcher tableau of the Lobatto IIIA methods have the particular structure

$c_1 = 0$	0	0	$\dots$	0	0
c_2	a_{21}	a_{22}	$\dots$	$a_{2,s-1}$	a_{2s}
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
c_{s-1}	$a_{s-1,1}$	$a_{s-1,2}$	$\dots$	$a_{s-1,s-1}$	$a_{s-1,s}$
$c_s = 1$	b_1	b_2	$\dots$	b_{s-1}	b_s
	b_1	b_2	$\dots$	b_{s-1}	b_s

From this Butcher tableau and (33), it is evident that the Lobatto IIIA methods have the special feature that $\boldsymbol{Q}_n^1 = \boldsymbol{q}_n$ and $\boldsymbol{Q}_n^s = \boldsymbol{q}_{n+1}$, which is exploited later for the formulation of the constraints. Specific Butcher tableaus for $s = 2, 3, 4, 5$ can be found in [23]. For consistency, it is assumed that the used Lobatto scheme satisfies

$$\sum_{i=1}^{s} b_i = 1 \quad \text{and} \quad b_i \neq 0 \quad \text{for} \quad i = 1, \dots, s. \tag{34}$$

[2] By definition, it holds that a set $C \subset \mathbb{R}^n$ is a cone if for any $\boldsymbol{x} \in C$ it follows that $\lambda\boldsymbol{x} \in C$ for all $\lambda > 0$.

In order to discretize the virtual action (30), we rewrite the integral over $\mathcal{I}$ as a sum of integrals over the time steps $\mathcal{I}_n = [t_n, t_{n+1})$ as

$$\delta A = \sum_{n=1}^{N-1} (\delta A_n^1 + \delta A_n^2) \quad \text{with} \quad \begin{cases} \delta A_n^1 = \delta \int_{\mathcal{I}_n} [L(t, \boldsymbol{q}, \boldsymbol{v}) + \boldsymbol{\pi}^{\mathrm{T}}(\dot{\boldsymbol{q}} - \boldsymbol{v})] \, \mathrm{d}t \\ \delta A_n^2 = \int_{\mathcal{I}_n} \delta \boldsymbol{q}^{\mathrm{T}} (\boldsymbol{f} \, \mathrm{d}t + \boldsymbol{W} \boldsymbol{R} \, \mathrm{d}\mu), \end{cases} \tag{35}$$

where we have used that $\mathrm{d}\boldsymbol{P} = \boldsymbol{R} \, \mathrm{d}\mu$, see (8). We approximate the first integral appearing in δA_n^1 using the Lobatto quadrature with the same coefficients as in (33). Instead of using a quadrature for the second part, i.e., the integral over $\boldsymbol{\pi}^{\mathrm{T}}(\dot{\boldsymbol{q}} - \boldsymbol{v})$ enforcing the kinematic equation, we enforce the discrete kinematic Eq. (33) using discrete Lagrange multipliers, as done in [3,22]. We follow [3] and approximate δA_n^1 by

$$\delta A_{d,n}^1 = \delta \left\{ \Delta t \sum_{i=1}^{s} b_i \left[L(t_n^i, \boldsymbol{Q}_n^i, \boldsymbol{V}_n^i) + (\bar{\boldsymbol{\Pi}}_n^i)^{\mathrm{T}} \left(\frac{\boldsymbol{Q}_n^i - \boldsymbol{q}_n}{\Delta t} - \sum_{j=1}^{s} a_{ij} \boldsymbol{V}_n^j \right) + \boldsymbol{\pi}_{n+1}^{\mathrm{T}} \left(\frac{\boldsymbol{q}_{n+1} - \boldsymbol{q}_n}{\Delta t} - \sum_{j=1}^{s} b_j \boldsymbol{V}_n^j \right) \right] \right\}. \tag{36}$$

Herein, $\boldsymbol{\pi}_{n+1}$ approximates $\boldsymbol{\pi}$ at t_{n+1}. However, it will be seen later that $\bar{\boldsymbol{\Pi}}_n^i$ is not the approximation of $\boldsymbol{\pi}$ at t_n^i but can only be interpreted as a Lagrange multiplier for the discretized kinematic equation. The approximation of δA_n^2 is chosen as

$$\delta A_{d,n}^2 = \sum_{i=1}^{s} b_i \, (\delta \boldsymbol{Q}_n^i)^{\mathrm{T}} \left[\Delta t \, \boldsymbol{f}(t_n^i, \boldsymbol{Q}_n^i, \boldsymbol{V}_n^i) + \boldsymbol{W}(t_n^i, \boldsymbol{Q}_n^i) \boldsymbol{R}_n^i \right]. \tag{37}$$

It can be seen, that for the integral over $\mathrm{d}t$ appearing in δA_n^2, we have chosen the Lobatto quadrature with the same coefficients as in (33). For the integral over $\mathrm{d}\mu$ however, we have chosen the same Lobatto quadrature up to the factor Δt. This can be motivated by the following.

In essence, we have approximated the change in percussion

$$\Delta \boldsymbol{P}(t; \Delta t) = \boldsymbol{P}^+(t + \Delta t) - \boldsymbol{P}^-(t) \overset{(1)}{=} \int_{[t, t+\Delta t]} \mathrm{d}\boldsymbol{P} = \int_{[t, t+\Delta t]} \boldsymbol{R} \, \mathrm{d}\mu \tag{38}$$

over a time step by the quadrature formula

$$\Delta \boldsymbol{P}(t; \Delta t) \approx \sum_{i=1}^{s} b_i \, \boldsymbol{R}_n^i \tag{39}$$

instead of using the usual Lobatto quadrature rule

$$\Delta \boldsymbol{P}(t; \Delta t) \approx \Delta t \sum_{i=1}^{s} b_i \, \hat{\boldsymbol{R}}_n^i. \tag{40}$$

Hence, we have basically used a version of the quadrature with rescaled constrained and contact forces in the discrete setting, as obviously $\boldsymbol{R}_n^i = \Delta t \, \hat{\boldsymbol{R}}_n^i$. This scaling is important in order to avoid unbounded growth of discrete constraint and contact forces for $\Delta t \to 0$. In fact, if $t \in \mathcal{D}$ is a discontinuity point, then $\Delta \boldsymbol{P}(t; 0) = \boldsymbol{P}^+(t) - \boldsymbol{P}^-(t) \neq 0$ is finite. It is clear that in the limit $\Delta t \to 0$, the right-hand side of (40) can only be finite if $\hat{\boldsymbol{R}}_n^i \to \infty$ for some i, which is numerically unfavorable. Hence, the chosen discretizaton (39) is preferred, as it does not have this issue.

With the approximations (36) and (37), the discretized principle of virtual action reads as

$$\delta A_d = \sum_{n=1}^{N-1} (\delta A_{d,n}^1 + \delta A_{d,n}^2) = 0 \quad \forall \delta \boldsymbol{q}_n, \delta \boldsymbol{Q}_n^i, \delta \boldsymbol{V}_n^i, \delta \bar{\boldsymbol{\Pi}}_n^i, \delta \boldsymbol{\pi}_{n+1}, \; n = 1, \dots, N, \; i = 1, \dots, s, \tag{41}$$

where $\delta \boldsymbol{q}_1 = \delta \boldsymbol{q}_N = 0$.

Under the assumptions (34), the following necessary and sufficient conditions for the discrete principle of virtual action (41) can be derived by choosing all but the indicated test quantity to be zero:

$$\delta \boldsymbol{Q}_n^i : \quad \boldsymbol{0} = \left(\frac{\partial L}{\partial \boldsymbol{q}} \right)^{\mathrm{T}} (t_n^i, \boldsymbol{Q}_n^i, \boldsymbol{V}_n^i) + \frac{\bar{\boldsymbol{\Pi}}_n^i}{\Delta t} + \boldsymbol{f}(t_n^i, \boldsymbol{Q}_n^i, \boldsymbol{V}_n^i) + \boldsymbol{W}(t_n, \boldsymbol{Q}_n^i) \frac{\boldsymbol{R}_n^i}{\Delta t} \tag{42a}$$

$$\delta \boldsymbol{V}_n^i : \quad \boldsymbol{0} = b_i \left(\frac{\partial L}{\partial \boldsymbol{v}} \right)^{\mathrm{T}} (t_n^i, \boldsymbol{Q}_n^i, \boldsymbol{V}_n^i) - \sum_{j=1}^{s} b_j a_{ji} \bar{\boldsymbol{\Pi}}_n^j - b_i \boldsymbol{\pi}_{n+1} \tag{42b}$$

$$\delta \boldsymbol{q}_n : \quad \boldsymbol{0} = -\sum_{i=1}^{s} b_i \frac{\bar{\boldsymbol{\Pi}}_n^i}{\Delta t} - \frac{\boldsymbol{\pi}_{n+1}}{\Delta t} + \frac{\boldsymbol{\pi}_n}{\Delta t} \tag{42c}$$

$$\delta \bar{\boldsymbol{\Pi}}_n^i : \quad \boldsymbol{Q}_n^i = \boldsymbol{q}_n + \Delta t \sum_{j=1}^{s} a_{ij} \boldsymbol{V}_n^j \tag{42d}$$

$$\delta \boldsymbol{\pi}_{n+1} : \quad \boldsymbol{q}_{n+1} = \boldsymbol{q}_n + \Delta t \sum_{j=1}^{s} b_j \boldsymbol{V}_n^j \tag{42e}$$

Condition (42c) can be rearranged to

$$\boldsymbol{\pi}_{n+1} = \boldsymbol{\pi}_n + \sum_{j=1}^{s} b_j \left(-\bar{\boldsymbol{\Pi}}_n^j \right). \tag{43}$$

To further manipulate the condition (42b), we define the discrete momentum

$$\boldsymbol{\Pi}_n^i = \left(\frac{\partial L}{\partial \boldsymbol{v}} \right)^{\mathrm{T}} (t_n^i, \boldsymbol{Q}_n^i, \boldsymbol{V}_n^i). \tag{44}$$

With that at hand, we can use (43) in the stationarity condition (42b) and obtain

$$\boldsymbol{\Pi}_n^i = \sum_{j=1}^{s} \frac{b_j a_{ji}}{b_i} \bar{\boldsymbol{\Pi}}_n^j + \boldsymbol{\pi}_{n+1} = \boldsymbol{\pi}_n + \sum_{j=1}^{s} \left(b_j - \frac{b_j a_{ji}}{b_i} \right) \left(-\bar{\boldsymbol{\Pi}}_n^j \right). \tag{45}$$

It is notationally favorable to introduce

$$\hat{a}_{ij} = b_j - \frac{b_j a_{ji}}{b_i}, \tag{46}$$

which can be recognized as the coefficients of the Lobatto IIIB method, see [24]. Moreover, it is convenient to introduce the notation $\boldsymbol{F}_n^i = -\bar{\boldsymbol{\Pi}}_n^i$, such that the stationarity conditions (42) can be equivalently stated as

$$\boldsymbol{q}_{n+1} = \boldsymbol{q}_n + \Delta t \sum_{j=1}^{s} b_j \boldsymbol{V}_n^j \tag{47a}$$

$$\boldsymbol{\pi}_{n+1} = \boldsymbol{\pi}_n + \sum_{j=1}^{s} b_j \boldsymbol{F}_n^j \tag{47b}$$

$$\boldsymbol{Q}_n^i = \boldsymbol{q}_n + \Delta t \sum_{j=1}^{s} a_{ij} \boldsymbol{V}_n^j \quad (i = 1, \dots, s) \tag{47c}$$

$$\boldsymbol{\Pi}_n^i = \boldsymbol{\pi}_n + \sum_{j=1}^{s} \hat{a}_{ij} \boldsymbol{F}_n^j \quad (i = 1, \dots, s) \tag{47d}$$

$$\boldsymbol{\Pi}_n^i = \left(\frac{\partial L}{\partial \boldsymbol{v}} \right)^{\mathrm{T}} (t_n^i, \boldsymbol{Q}_n^i, \boldsymbol{V}_n^i) \quad (i = 1, \dots, s) \tag{47e}$$

$$\boldsymbol{F}_n^i = \Delta t \left(\frac{\partial L}{\partial \boldsymbol{q}} \right)^{\mathrm{T}} (t_n^i, \boldsymbol{Q}_n^i, \boldsymbol{V}_n^i) + \Delta t \, \boldsymbol{f}(t_n^i, \boldsymbol{Q}_n^i, \boldsymbol{V}_n^i) + \boldsymbol{W}(t_n, \boldsymbol{Q}_n^i) \boldsymbol{R}_n^i \quad (i = 1, \dots, s). \tag{47f}$$

With the above scheme we have hence derived a Lobatto IIIA-IIIB discretization of the partitioned system (5)–(7), i.e.,

$$\mathrm{d}\boldsymbol{q} = \boldsymbol{v} \, \mathrm{d}t, \quad \mathrm{d}\boldsymbol{\pi} = \left(\frac{\partial L}{\partial \boldsymbol{q}} \right)^{\mathrm{T}} \mathrm{d}t + \boldsymbol{f} \, \mathrm{d}t + \boldsymbol{W} \boldsymbol{R} \, \mathrm{d}\mu, \quad \boldsymbol{\pi} = \left(\frac{\partial L}{\partial \boldsymbol{v}} \right)^{\mathrm{T}}, \tag{48}$$

As approximant of the velocity of the system at t_n, we implicitly define $\boldsymbol{v}_n$ by

$$\boldsymbol{\pi}_n = \left(\frac{\partial L}{\partial \boldsymbol{v}} \right)^{\mathrm{T}} (t_n, \boldsymbol{q}_n, \boldsymbol{v}_n). \tag{49}$$

Before addressing the discretization of the constraints and the frictional contact, it is worthwhile having a closer look at the structure of the coefficients of the Lobatto IIIB method. Since the Lobatto IIIA method has the property $a_{si} = b_i$ and $a_{1i} = 0$ for all i, it follows from (46) that the coefficients of the Lobatto IIIB method satisfy $\hat{a}_{is} = 0$ and $\hat{a}_{i1} = b_1$. The resulting Butcher tableau has the form

$c_1 = 0$	b_1	$\hat{a}_{12}$	$\dots$	$\hat{a}_{1,s-1}$	0
c_2	b_1	$\hat{a}_{22}$	$\dots$	$\hat{a}_{2,s-1}$	0
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
c_{s-1}	b_1	$\hat{a}_{s-1,2}$	$\dots$	$\hat{a}_{s-1,s-1}$	0
$c_s = 1$	b_1	$\hat{a}_{s2}$	$\dots$	$\hat{a}_{s,s-1}$	0
	b_1	b_2	$\dots$	b_{s-1}	b_s

Specific Butcher tableaus for $s = 2, 3, 4, 5$ can be found in [23]. The last column of the Butcher tabelau is always zero which by (47d) implies that the generalized momenta $\boldsymbol{\Pi}_n^i$ of all stages i do not depend on the discrete force $\boldsymbol{F}_n^s$ of the last stage. Hence, it can be seen from (47f) that the Lagrange multiplier $\boldsymbol{R}_n^s$ does not influence $\boldsymbol{\Pi}_n^i$ and consequently has no influence on $\boldsymbol{V}_n^i$, see (47e). Ultimately, by (47a), this means that the position update $\boldsymbol{q}_{n+1}$ is not influenced by $\boldsymbol{R}_n^s$. This peculiarity of the Lobatto IIIB method is exploited for the enforcement of the constraints as well as for the formulation of the discrete contact laws.

We impose the bilateral constraints discretely as in [3] and demand

$$\boldsymbol{g}(t_n^i, \boldsymbol{Q}_n^i) = \boldsymbol{0} \ (i = 2, \dots, s), \quad \dot{\boldsymbol{g}}(t_{n+1}, \boldsymbol{q}_{n+1}, \boldsymbol{v}_{n+1}) = \boldsymbol{0} \tag{50}$$

$$\boldsymbol{\gamma}(t_n^i, \boldsymbol{Q}_n^i, \boldsymbol{V}_n^i) = \boldsymbol{0} \ (i = 2, \dots, s), \quad \boldsymbol{\gamma}(t_{n+1}, \boldsymbol{q}_{n+1}, \boldsymbol{v}_{n+1}) = \boldsymbol{0}. \tag{51}$$

Due to the property of the Lobatto IIIA method that $(t_n^s, \boldsymbol{Q}_n^s) = (t_{n+1}, \boldsymbol{q}_{n+1})$, the position level constraints are satisfied at the end of the time step. Moreover, because of the particular structure of the Butcher tableau of the Lobatto IIIB scheme, the constraint $\boldsymbol{g}(t, \boldsymbol{q}) = \boldsymbol{0}$ can be imposed on both position and velocity level at the end of the time step, see (50). Namely, we exploit that the Lagrange multiplier $\boldsymbol{R}_{g,n}^s$ only affects the update of $\boldsymbol{v}_{n+1}$ and has no influence on $\boldsymbol{q}_{n+1}$. This decoupling allows for the simultaneous enforcement of both constraints. As a consequence of the splitting, we can attribute the Lagrange multiplier $\boldsymbol{R}_{g,n}^s$ to the enforcement of the velocity level constraint $\dot{\boldsymbol{g}}(t_{n+1}, \boldsymbol{q}_{n+1}, \boldsymbol{v}_{n+1}) = \boldsymbol{0}$. Hence, the remaining Lagrange multipliers $\boldsymbol{R}_{g,n}^{i-1}$, $i = 2, \dots, s$ are responsible for the fulfillment of the other constraints, i.e., $\boldsymbol{g}(t_n^i, \boldsymbol{Q}_n^i) = \boldsymbol{0}$, $i = 2, \dots, s$. It is crucial to see that there is an index shift between the constraint and the Lagrange multipliers, especially in view of the formulation of unilateral constraints.

Similar to the bilateral constraints, we impose the unilateral constraints, i.e., Signorini's law for the stages $i = 2, \dots, s$ as

$$\forall k: \quad g_N^k(t_n^i, \boldsymbol{Q}_n^i) \in \mathcal{N}_{\mathbb{R}_0^-}(-R_{N,n}^{k,i-1}). \tag{52}$$

Following the analogy to (50) further, it would be natural to additionally impose the unilateral constraint at velocity level at the end of the time step. However, from the discussion preceding (24) it can be concluded that instead the Newton-type impact law can be imposed. To do so, it is assumed that if the contact k is active at the end of the time step, i.e., $k \in A_{n+1} = A(t_{n+1}, \boldsymbol{q}_{n+1})$, then it has been active during the whole time step $\mathcal{I}_n = (t_n, t_{n+1}]$. Moreover, during that time step, the kinematic quantity ξ_N^k is assumed to be constant and to correspond to

$$\xi_{N,n+1}^k = \dot{g}_N^k(t_{n+1}, \boldsymbol{q}_{n+1}, \boldsymbol{v}_{n+1}) + e_N^k \, \dot{g}_N^k(t_n, \boldsymbol{q}_n, \boldsymbol{v}_n), \tag{53}$$

which is regarded as a discrete approximation of (20) over one time step. These assumptions allow to "integrate" (19) over one time step, resulting in

$$\begin{aligned} k \in A_{n+1}: &\quad \xi_{N,n+1}^k \in \mathcal{N}_{\mathbb{R}_0^-}(-\textstyle\int_{\mathcal{I}_n} R_N^k \, \mathrm{d}\mu) \\ k \in \bar{A}_{n+1}: &\quad \textstyle\int_{\mathcal{I}_n} R_N^k \, \mathrm{d}\mu = 0. \end{aligned} \tag{54}$$

A detailed discussion of how this is done as well as an interpretation of the integrated contact law can be found in Section 6 of [13]. Using the Lobatto quadrature rule to approximate the integral as

$$\int_{\mathcal{I}_n} R_N^k \, \mathrm{d}\mu \approx \sum_{j=1}^{s} b_j R_{N,n}^{k,j} = \Delta P_{N,n+1}^k, \tag{55}$$

which is the discrete approximation of (38). Substitution of (55) into (54) results in the discrete velocity level contact law

$$\begin{aligned} k \in A(t_{n+1}, \boldsymbol{q}_{n+1}): &\quad \xi_{N,n+1}^k \in \mathcal{N}_{\mathbb{R}_0^-}(-\Delta P_{N,n+1}^k) \\ k \in \bar{A}(t_{n+1}, \boldsymbol{q}_{n+1}): &\quad \Delta P_{N,n+1}^k = 0. \end{aligned} \tag{56}$$

Summing up, the temporally discretized contact law in normal direction is given by (52) and (56). This means that the unilateral constraint is imposed on position level for all stages $i = 2, \dots, s$ and the Newton-type impact law is fulfilled in an integral way over the whole time step, where $\boldsymbol{v}_n$ and $\boldsymbol{v}_{n+1}$ are regarded as the pre- and post-impact velocities, respectively. It is straightforward to see, that in the case that a contact is closed during some interval of time, then the unilateral constraint looks like a bilateral constraint $g_N^k(t, \boldsymbol{q}) = 0$ during that time interval. It is an important feature of the presented discretized contact law that in that case the discrete contact law reduces to the discrete bilateral constraint Eqs. (50) formulated with g_N^k. Hence, it can be stated that the discretized contact law in normal direction is consistent with the discretization of the bilateral constraints.

For the discretization of the Coulomb friction, a similar path is taken. We impose Coulomb friction (29) for all stages $i = 2, \dots, s$ by

$$\forall k: \quad \boldsymbol{\gamma}_F^k(t_n^i, \boldsymbol{Q}_n^i, \boldsymbol{V}_n^i) \in \mathcal{N}_{C_F^k(R_{N,n}^{k,i-1})}(-\boldsymbol{R}_{F,n}^{k,i-1}) \tag{57}$$

together the integrated version of the Newton-type frictional impact law (27)

$$\forall k: \quad \boldsymbol{\xi}_{F,n+1}^k = \boldsymbol{\gamma}_F(t_{n+1}, \boldsymbol{q}_{n+1}, \boldsymbol{v}_{n+1}) + e_F^k \, \boldsymbol{\gamma}_F^k(t_n, \boldsymbol{q}_n, \boldsymbol{v}_n) \in \mathcal{N}_{C_F^k(\Delta P_{N,n+1}^k)}(-\Delta \boldsymbol{P}_{F,n+1}^k). \tag{58}$$

Again, the discrete percussions are defined using the Lobatto quadrature as

$$\int_{\mathcal{I}_n} \boldsymbol{R}_F^k \, \mathrm{d}\mu \approx \sum_{j=1}^{s} b_j \boldsymbol{R}_{F,n}^{k,j} = \Delta \boldsymbol{P}_{F,n+1}^k. \tag{59}$$

5. Implementation of the scheme

For the computation of one time step $(t_n, \boldsymbol{q}_n, \boldsymbol{v}_n) \mapsto (t_{n+1}, \boldsymbol{q}_{n+1}, \boldsymbol{v}_{n+1})$ of the scheme presented in the previous section, the scheme is brought into residual form and, similar to [13], is then solved using a semismooth Newton method [7,25].

In order to formulate the residual form of the scheme as $\boldsymbol{\Phi}(\boldsymbol{x}) = \boldsymbol{0}$, the vector of unknowns

$$\boldsymbol{x} = (\boldsymbol{V}_n^i, \boldsymbol{v}_{n+1}, \boldsymbol{R}_{g,n}^i, \boldsymbol{R}_{\gamma,n}^i, \boldsymbol{R}_{N,n}^i, \boldsymbol{R}_{F,n}^i) \tag{60}$$

is introduced, where for notational brevity we write $\boldsymbol{V}_n^i$ instead of $\boldsymbol{V}_n^1, \dots, \boldsymbol{V}_n^s$. The residual describing a time step is subdivided into three contributions as $\boldsymbol{\Phi} = (\boldsymbol{\Phi}_S, \boldsymbol{\Phi}_N, \boldsymbol{\Phi}_F)$. The first contribution contains Eqs. (47), (49), (50) and (51) describing the time step except for the contacts and is given by

$$\boldsymbol{\Phi}_S(\boldsymbol{x}) = \begin{bmatrix} \left(\frac{\partial L}{\partial \boldsymbol{v}}\right)^{\mathrm{T}}(t_n^i, \boldsymbol{Q}_n^i, \boldsymbol{V}_n^i) - \left(\frac{\partial L}{\partial \boldsymbol{v}}\right)^{\mathrm{T}}(t_n, \boldsymbol{q}_n, \boldsymbol{v}_n) + \sum_{j=1}^s \hat{a}_{ij} \boldsymbol{F}_n^j \\ \left(\frac{\partial L}{\partial \boldsymbol{v}}\right)^{\mathrm{T}}(t_{n+1}, \boldsymbol{q}_{n+1}, \boldsymbol{v}_{n+1}) - \left(\frac{\partial L}{\partial \boldsymbol{v}}\right)^{\mathrm{T}}(t_n, \boldsymbol{q}_n, \boldsymbol{v}_n) + \sum_{j=1}^s b_j \boldsymbol{F}_n^j \\ \boldsymbol{g}(t_n^i, \boldsymbol{Q}_n^i) \ (i = 2, \dots, s) \\ \dot{\boldsymbol{g}}(t_{n+1}, \boldsymbol{q}_{n+1}, \boldsymbol{v}_{n+1}) \\ \boldsymbol{\gamma}(t_n^i, \boldsymbol{Q}_n^i, \boldsymbol{V}_n^i) \ (i = 2, \dots, s) \\ \boldsymbol{\gamma}(t_{n+1}, \boldsymbol{q}_{n+1}, \boldsymbol{v}_{n+1}) \end{bmatrix}. \tag{61}$$

Unless indicated otherwise, the index i ranges from 1 to s. Moreover, the discrete positions are seen as $\boldsymbol{x}$-dependent quantities, where the dependence is given by (33) and the properties of the corresponding Butcher tableau, i.e.,

$$\boldsymbol{Q}_n^i = \boldsymbol{q}_n + \Delta t \sum_{j=1}^s a_{ij} \boldsymbol{V}_n^j \quad \text{and} \quad \boldsymbol{q}_{n+1} = \boldsymbol{Q}_n^s. \tag{62}$$

With that, also the discrete forces $\boldsymbol{F}_n^i$ depend on $\boldsymbol{x}$ only and read as

$$\boldsymbol{F}_n^i = \Delta t \left(\frac{\partial L}{\partial \boldsymbol{q}}\right)^{\mathrm{T}}(t_n^i, \boldsymbol{Q}_n^i, \boldsymbol{V}_n^i) + \Delta t \, \boldsymbol{f}(t_n^i, \boldsymbol{Q}_n^i, \boldsymbol{V}_n^i) + \boldsymbol{W}(t_n, \boldsymbol{Q}_n^i) \boldsymbol{R}_n^i. \tag{63}$$

The remainder of the residual $\boldsymbol{\Phi}$ is specified for each stage $i = 1, \dots, s$ and each contact $k = 1, \dots, n_N$ separately and is then gathered as $\boldsymbol{\Phi}_N = (\boldsymbol{\Phi}_N^1, \dots, \boldsymbol{\Phi}_N^s)$ with $\boldsymbol{\Phi}_N^i = (\Phi_N^{1,i}, \dots, \Phi_N^{n_N,i})$. Similarly, $\boldsymbol{\Phi}_F = (\boldsymbol{\Phi}_F^1, \dots, \boldsymbol{\Phi}_F^s)$ with $\boldsymbol{\Phi}_F^i = (\boldsymbol{\Phi}_F^{1,i}, \dots, \boldsymbol{\Phi}_F^{n_N,i})$. That being said, the appearing normal cone inclusions have to be equivalently reformulated as equations for the formulation of the contact laws (52), (56), (57) and (58) in residual form. To do so, we use the fact that for two vectors $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^f$ and a closed convex nonempty set $C \subset \mathbb{R}^f$, the following are equivalent relations

$$\boldsymbol{y} \in \mathcal{N}_C(-\boldsymbol{x}) \quad \Longleftrightarrow \quad -\boldsymbol{x} = \mathrm{prox}_C(r\boldsymbol{y} - \boldsymbol{x}) \quad \forall r > 0, \tag{64}$$

where

$$\mathrm{prox}_C : \mathbb{R}^f \to \mathbb{R}^f, \ \boldsymbol{p} \mapsto \boldsymbol{q} = \underset{\boldsymbol{p}^* \in C}{\arg\min} \left(\tfrac{1}{2} \|\boldsymbol{p} - \boldsymbol{p}^*\|^2 \right)$$

is the proximal point function to C, see [13,26]. The two proximal point functions needed for the implementation of the scheme are

$$\mathrm{prox}_{\mathbb{R}_0^-}(x) = \min\{x, 0\} \quad \text{and} \quad \mathrm{prox}_{B_2(r)}(\boldsymbol{x}) = \begin{cases} \boldsymbol{x} & \text{if} \quad \boldsymbol{x} \in B_2(r) \\ r \frac{\boldsymbol{x}}{\|\boldsymbol{x}\|} & \text{if} \quad \boldsymbol{x} \notin B_2(r). \end{cases} \tag{65}$$

Using (64), the discrete normal contact laws (52) and (56) are formulated in residual form as

$$\begin{aligned} \forall k: \quad & \boldsymbol{\Phi}_N^{k,i-1}(\boldsymbol{x}) = R_n^{k,i-1} + \mathrm{prox}_{\mathbb{R}_0^-}\left(r g_N^k(t_n^i, \boldsymbol{Q}_n^i) - R_n^{k,i-1}\right), \quad i = 2, \dots, s \\ k \in \mathcal{A}(t_{n+1}, \boldsymbol{q}_{n+1}): \quad & \boldsymbol{\Phi}_N^{k,s}(\boldsymbol{x}) = \Delta P_{N,n+1}^k + \mathrm{prox}_{\mathbb{R}_0^-}\left(r \xi_{N,n+1}^k - \Delta P_{N,n+1}^k\right) \\ k \in \bar{\mathcal{A}}(t_{n+1}, \boldsymbol{q}_{n+1}): \quad & \boldsymbol{\Phi}_N^{k,s}(\boldsymbol{x}) = \Delta P_{N,n+1}^k, \end{aligned} \tag{66}$$

where the set of active constraints is computed as

$$\mathcal{A}(t_{n+1}, \boldsymbol{q}_{n+1}) = \{k \mid r g_N^k(t_{n+1}, \boldsymbol{q}_{n+1}) - R_n^{k,s-1} \leq 0\}, \tag{67}$$

$r > 0$ is a numerical parameter and the dependence of the discrete percussion and $\xi_{N,n+1}^k$ on $\boldsymbol{x}$ is given by (53) and (55), i.e.,

$$\Delta P_{N,n+1}^k = \sum_{j=1}^s b_j R_{N,n}^{k,j} \quad \text{and} \quad \xi_{N,n+1}^k = \dot{g}_N^k(t_{n+1}, \boldsymbol{q}_{n+1}, \boldsymbol{v}_{n+1}) + e_N^k \, \dot{g}_N^k(t_n, \boldsymbol{q}_n, \boldsymbol{v}_n). \tag{68}$$

In the transition from (56) to (66) we have replaced the active set A by the active set $\mathcal{A}$ since it can be shown that $A(t_{n+1}, \boldsymbol{q}_{n+1}) = \mathcal{A}(t_{n+1}, \boldsymbol{q}_{n+1})$. However, using the active set $\mathcal{A}$ is numerically more robust, see the discussion of this topic in [14].

Similar to the normal contact law, also the discrete friction laws (57) and (58) are formulated in residual form with the help of (64). Specifically,

$$\begin{aligned} \forall k: \quad & \boldsymbol{\Phi}_F^{k,i-1}(\boldsymbol{x}) = \boldsymbol{R}_{F,n}^{k,i-1} + \mathrm{prox}_{C_F^k(R_{N,n}^{k,i-1})}\left(r \boldsymbol{\gamma}_F^k(t_n^i, \boldsymbol{Q}_n^i, \boldsymbol{V}_n^i) - \boldsymbol{R}_{F,n}^{k,i-1}\right), \quad i = 2, \dots, s \\ \forall k: \quad & \boldsymbol{\Phi}_F^{k,s}(\boldsymbol{x}) = \Delta \boldsymbol{P}_{F,n+1}^k + \mathrm{prox}_{C_F^k(\Delta P_{N,n+1}^k)}\left(r \boldsymbol{\xi}_{F,n+1}^k - \Delta \boldsymbol{P}_{F,n+1}^k\right) \end{aligned} \tag{69}$$

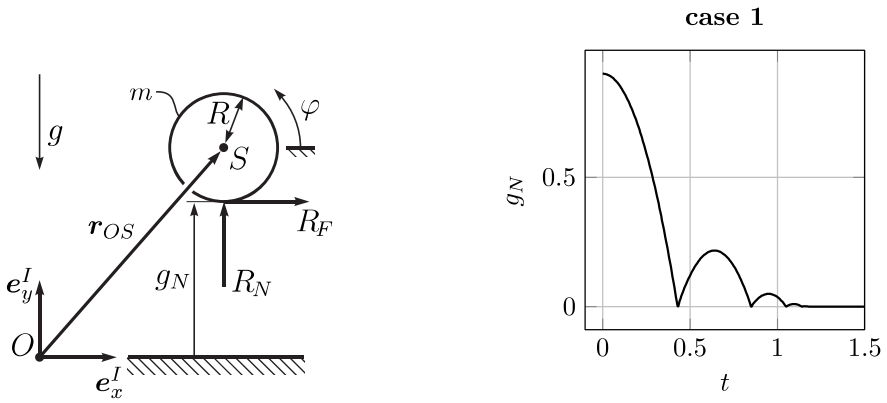

Fig. 3. Sketch of the bouncing ball system (left) and simulated time evolution of the contact distance g_N for case 1 (right).

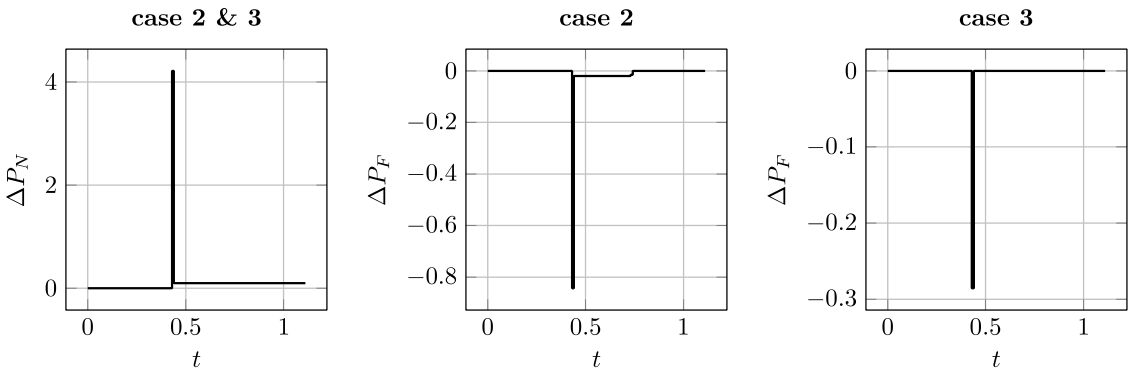

Fig. 4. Simulated normal and friction contact percussions for case 2 ($\omega = 50$) and case 3 ($\omega = 10$).

where the discrete percussion and $\xi_{F,n+1}^k$ depend on $\boldsymbol{x}$ by (58) and (59), which read as

$$\Delta \boldsymbol{P}_{F,n+1}^k = \sum_{j=1}^{s} b_j \boldsymbol{R}_{F,n}^{k,j} \quad \text{and} \quad \xi_{F,n+1}^k = \boldsymbol{\gamma}_F^k(t_{n+1}, \boldsymbol{q}_{n+1}, \boldsymbol{v}_{n+1}) + e_F^k \boldsymbol{\gamma}_F^k(t_n, \boldsymbol{q}_n, \boldsymbol{v}_n). \tag{70}$$

6. Numerical validation

In this section, we validate the presented scheme using selected benchmark examples. The bouncing ball example is used to affirm that the discrete contact law is able to properly resolve the contact dynamics in an event capturing way. With the simulation of the slider-crank mechanism, the ability to simulate systems with bilateral constraints and multiple contacts is showcased. Finally, the point mass on a slope is used to numerically analyze the convergence behavior of the scheme.

6.1. Rotating bouncing ball

To show that the presented scheme is able to capture the contact dynamics well, we use the rotating bouncing ball example from [13]. The system consists of a homogeneous sphere of radius $R = 0.1$ and mass $m = 1$ subjected to gravity with gravitational acceleration $g = 9.81$. The ball can come into contact with a horizontal plane, see Fig. 3. The system is planar and the position of the ball's center of mass is ${}_I\boldsymbol{r}_{OS} = (x,\ y,\ 0)$. The rotational degree of freedom of the sphere is described by the angle φ. Hence, the generalized coordinates of the system are $\boldsymbol{q} = (x,\ y,\ \varphi)$. For the description of planar Coulomb friction, $\mu = 0.2$ and $e_F = 0$ are chosen.

To examine the contact dynamics in full detail, three different cases are simulated. All share the initial conditions $\boldsymbol{q}(0) = (0,\ 1,\ 0)$ and $\boldsymbol{v}(0) = (0,\ 0,\ \omega)$, such that the ball has an initial rotational velocity ω. However, the cases differ by the values of ω and the restitution coefficient e_N. All simulations were performed using a constant step size $\Delta t = 10^{-2}$ and the proximal point parameter $r = 0.5$. The tolerance for the semismooth Newton method used to solve each time step was set to 10^{-8}. For this parameter set, the numerical solutions of the methods with $s = 2, 3, 4$ stages are nearly indistinguishable, which is why we used $s = 3$ for all simulations in Fig. 4.

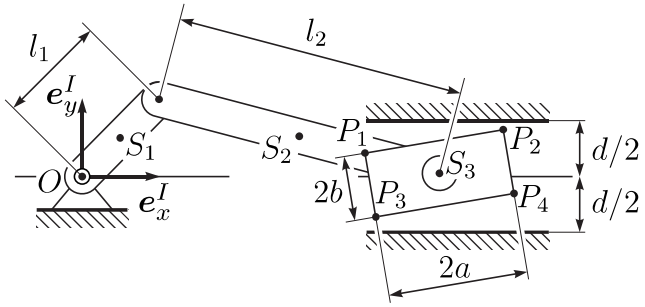

Fig. 5. Sketch of the slider-crank mechanism consisting of three rigid bodies, which are interconnected by revolute joints. The third body can come into contact with the walls through the points $P_1, \dots, P_4$.

Table 1
Geometry and inertia properties of the slider-crank mechanism.

l_1	l_2	$a = 2b$	d	$m_1 = m_2$	m_3	θ_{S_1}	θ_{S_2}	θ_{S_3}
0.153	0.306	0.05	0.052	0.038	0.076	$7.4 \cdot 10^{-5}$	$5.9 \cdot 10^{-4}$	$2.7 \cdot 10^{-6}$

Case 1:. ($\omega = 0$, $e_N = 0.5$) The motion of the system is the typical bouncing ball motion exhibiting the Zeno phenomenon. The simulation result shown in Fig. 3 confirms that the presented scheme can overcome accumulation points, as expected by an event capturing scheme. Moreover, no contact penetration is observed, since this is enforced by the scheme.

Case 2:. ($\omega = 50$, $e_N = 0$) Due to the inelastic impact, the contact remains closed after the impact and the friction behavior can be validated. For such a high rotational velocity, sliding occurs directly after the frictional impact. Hence, after an initial period of sliding contact, the ball slows down and a slip-stick transition occurs. At this point, the ball begins to move in a pure rolling motion. Since the rolling motion is characterized by constant velocities, the friction forces are zero. Consequently, at the slip-stick transition, the friction force instantly drops to zero. Fig. 4 illustrates that the presented scheme perfectly replicates this behavior.

Case 3:. ($\omega = 10$, $e_N = 0$) This case is similar to the second one. However, due to the lower initial velocity, the ball sticks upon impact and exhibits a pure rolling motion with constant velocities and zero friction forces afterwards. This behavior is captured by the scheme, see Fig. 4.

The numerical results of cases 2 and 3 perfectly reproduce the analytical solutions which can be found in [13]. Moreover, since the RATTLE scheme in essence is the Lobatto IIIA-IIIB pair with two stages, these simulation results are unsurprisingly identical to those in [14].

6.2. Slider-crank mechanism

This example demonstrates the ability of the proposed scheme to handle impulsive changes of bilateral constraint forces as well as a variety of high frequency contact patterns. The slider-crank mechanism consists of three planar rigid bodies with masses m_1, m_2, m_3 and rotational inertia θ_{S_1}, θ_{S_2}, θ_{S_3} with respect to their centers of mass S_1, S_2, S_3, respectively. Each body is subjected to gravity with $g = 9.81$. As in the previous example, each body is described by the coordinates x_i, y_i of its center of mass and an angle φ_i describing its rotation. Hence, for the slider-crank mechanism the generalized coordinates $\boldsymbol{q} = (x_1,\ y_1,\ \varphi_1,\ x_2,\ y_2,\ \varphi_2,\ x_3,\ y_3,\ \varphi_3)$ are chosen. As shown in Fig. 5, the three bodies are connected via revolute joints, which are modeled as ideal bilateral constraints. Moreover, the four corner points P_i of the third body can come into contact with either the upper or the lower wall of the cylinder. Consequently, the gap functions g_{Ni} describing the contact are the distance of the respective corner point with the closest wall. For all contacts, the parameters $e_N = 0.4$, $e_F = 0$ and $\mu = 0.01$ are chosen. The equations of motion as well as the kinematic quantities used for the definition of the constraints and contacts can be found in [14].

The geometry and inertia properties of the mechanism are listed in Table 1. The initial conditions are $\boldsymbol{q}(0) = (0.0765,\ 0,\ 0,\ 0.306,\ 0,\ 0,\ 0.459,\ 0,\ 0.017)$ and $\boldsymbol{v}(0) = (0,\ 11.475,\ 150,\ 0,\ 11.475,\ -75,\ 0,\ 0,\ 0)$. The simulations were performed using the three stage method with a constant step size $\Delta t = 10^{-4}$ and the proximal point parameter $r = 1$. The tolerance for the semismooth Newton method was set to 10^{-8}.

The simulation results are depicted in Fig. 6. Since the simulation started with a small perturbation of the slider's orientation ($\varphi_3 \approx 1°$), it is apparent from Fig. 6(c), that the slider's orientation is stabilized after a short time ($t \approx 0.01$). Moreover, Figs. 6 (a) and (b) show the evolution of two contact distances and their time derivatives. As enforced by the scheme, no penetration is present. The spatial trajectory of the center of mass of the third body is shown by (d) of the same figure.

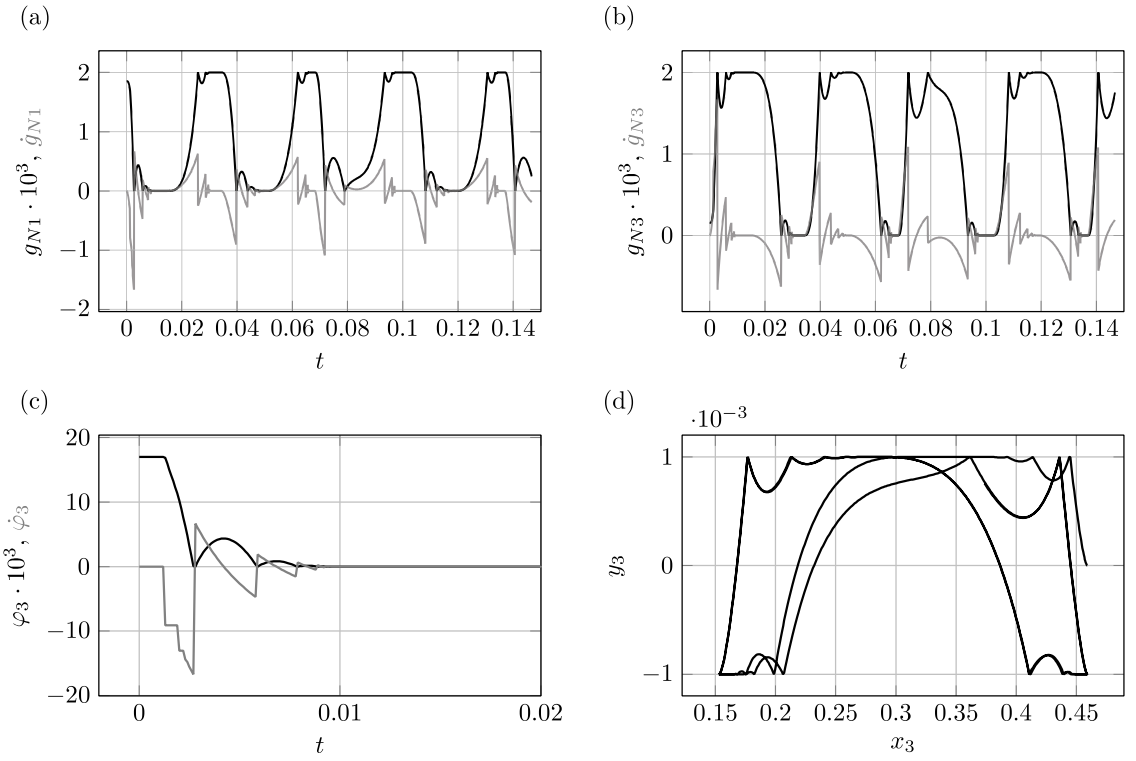

Fig. 6. Exemplary results of the slider-crank mechanism. (a) temporal evolution of the gap function and its time derivative for the contact point P_1. (b) temporal evolution of the gap function and its time derivative for the contact point P_3. (c) temporal evolution of the rotation angle and angular velocity of body 3. (d) spatial trajectory of the point S_3.

6.3. Point mass on slope

Consider a point mass (mass $m = 1$, gravitational acceleration $g = 10$) falling and sliding on a slope described by the exponential function. The position of the mass is described by the Cartesian coordinates $\boldsymbol{q} = (x,\ y)$ with corresponding velocities $\boldsymbol{v} = (v_x, v_y)$. The curve $\boldsymbol{f} : \mathbb{R} \to \mathbb{R}^2$, defining the slope, and its first derivative are given by

$$\boldsymbol{f}(x) = \big(x,\ \exp(-x)\big), \quad \boldsymbol{f}'(x) = \big(1,\ -\exp(-x)\big), \tag{71}$$

see Fig. 7. The tangential contact velocity γ_F used to describe the friction between the point mass and the slope is computed using the tangent vector $\boldsymbol{t}$, i.e.,

$$\gamma_F(\boldsymbol{q}, \boldsymbol{v}) = \boldsymbol{t}(x)^{\mathrm{T}} \boldsymbol{v} \qquad \text{with} \qquad \boldsymbol{t}(x) = \frac{\boldsymbol{f}'(x)}{\|\boldsymbol{f}'(x)\|}. \tag{72}$$

Using the outward pointing normal $\boldsymbol{n}$ of the slope, the normal contact distance is

$$g_N(\boldsymbol{q}) = \boldsymbol{n}(x)^{\mathrm{T}} \big(\boldsymbol{q} - \boldsymbol{f}(x)\big) \quad \text{with} \quad \boldsymbol{n}(x) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \boldsymbol{t}(x). \tag{73}$$

We assume inelastic contacts, i.e., $e_N = e_F = 0$, and choose the friction parameter as $\mu = 0.3$. Depending on the chosen initial conditions, two cases are investigated, see Fig. 7.

Case 1:. (sliding) The point mass starts from rest and lies on the slope, i.e., it starts with a closed contact. We therefore choose $\boldsymbol{q}(0) = \boldsymbol{f}(0) = (0,\ 1)$ and $\boldsymbol{v}(0) = (0,\ 0)$. For this initial condition and the chosen friction parameter, the point mass slides downhill and comes to rest at $t \approx 2.1$ due to friction.

Case 2:. (sliding after impact) The point mass starts at rest with $\boldsymbol{q}(0) = (0,\ 1.5)$, i.e., it starts from a resting position situated above the slope. Hence, the point mass will fall down and eventually come into contact with the slope, where an inelastic impact occurs. Subsequently, the point mass starts sliding on the slope until coming to rest.

The trajectories for both cases are depicted in Fig. 8. It can be seen that in case 1 the velocity of the point mass is continuous, whereas in case 2 there is a velocity jump at $t \approx 0.3$ due to the impact. Between $t = 2$ and $t = 2.5$, the motions of both cases show a kink in the velocity. At that time instant, the velocity is not differentiable. This behavior is the result of a slip-stick transition, which causes the acceleration to jump.

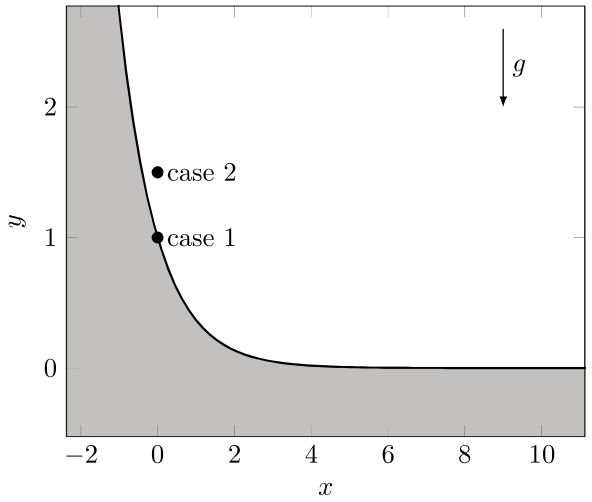

Fig. 7. Sketch of the initial conditions for the point mass (falling and) sliding on a slope. The contact surface is given by the function $y = \exp(-x)$.

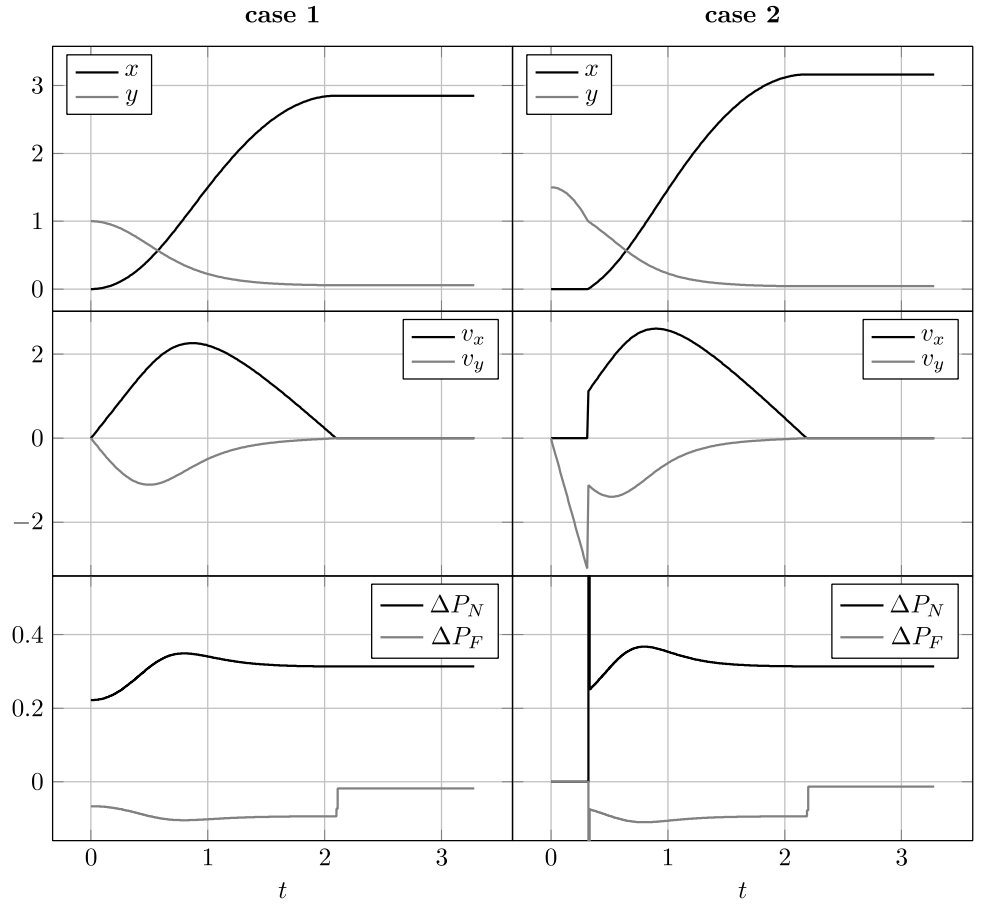

Fig. 8. Simulation results of the point mass on slope example for cases 1 and 2 computed using the presented method with $s = 3$ and $\Delta t = 10^{-2}$.

We investigate the convergence of the presented schemes numerically. For that, the cases 1 and 2 are simulated with different step sizes $\Delta t \in \{3.2 \cdot 10^{-3},\ 6.4 \cdot 10^{-3},\ 1.28 \cdot 10^{-2},\ 2.56 \cdot 10^{-2},\ 5.12 \cdot 10^{-2},\ 1.024 \cdot 10^{-1},\ 2.048 \cdot 10^{-1},\ 4.096 \cdot 10^{-1}\}$ and the convergence behavior of the solutions with respect to an accurate numerical solution computed with $\Delta t^{\text{ref}} = 5 \cdot 10^{-5}$ is studied. Let $\boldsymbol{x}_n$ and $\boldsymbol{x}_n^{\text{ref}}$ ($n = 1, 2, \dots$) be the time evolution of a quantity (e.g. $\boldsymbol{x} = \boldsymbol{q}$ or $\boldsymbol{x} = \boldsymbol{v}$) computed with the step sizes Δt and Δt^{ref}, respectively.

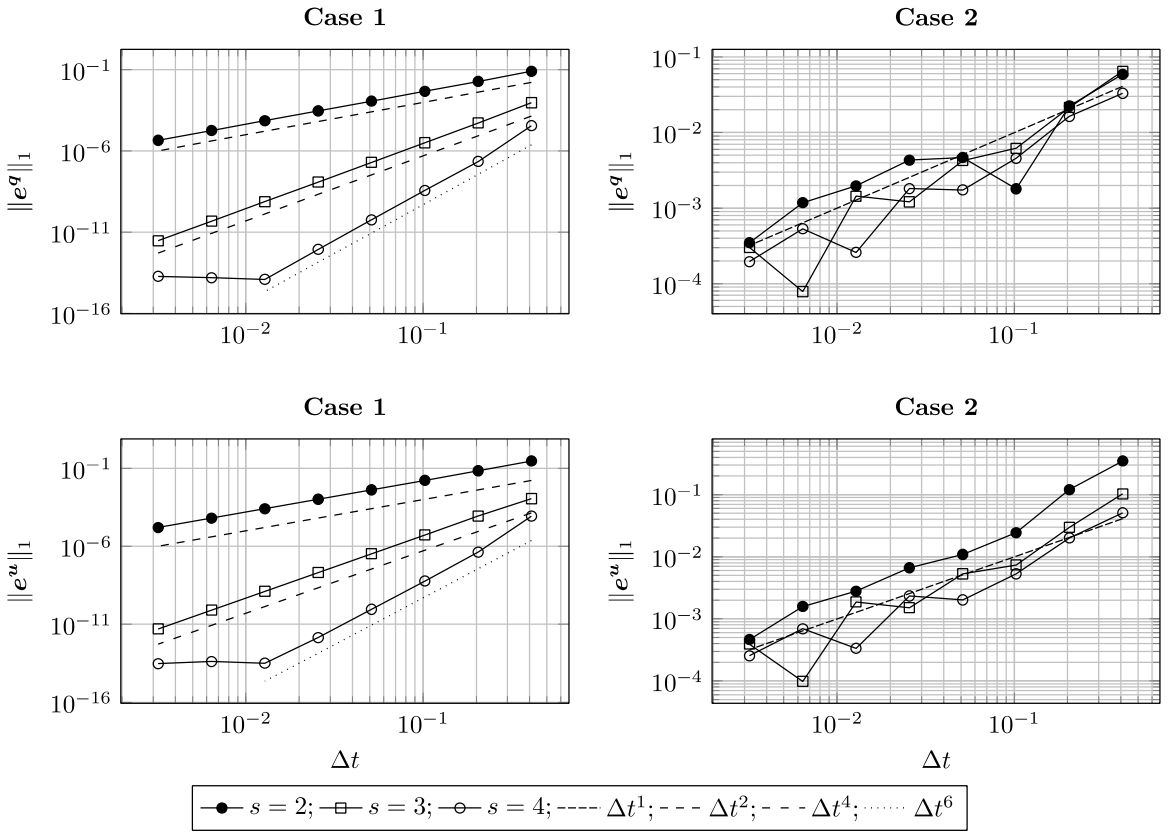

Fig. 9. Convergence behavior of the s-stage Lobatto IIIA-IIIB scheme for case 1 and 2.

Moreover, assume that $\boldsymbol{x}_n$ and $\boldsymbol{x}_n^{\text{ref}}$ approximate the values of $\boldsymbol{x}$ at the times $t_n = n\Delta t$. The time evolution of the error between the simulated and the reference solution is therefore $\boldsymbol{e}_n^{\boldsymbol{x}} = \boldsymbol{x}_n - \boldsymbol{x}_n^{\text{ref}}$ and we define the overall simulation error as

$$\|\boldsymbol{e}^{\boldsymbol{x}}\|_1 = \Delta t \sum_{n=1}^{N} \|\boldsymbol{e}_n^{\boldsymbol{x}}\|_1 = \Delta t \sum_{n=1}^{N} \sum_{k=1}^{n_x} |e_n^{\boldsymbol{x},k}|, \tag{74}$$

where $e_n^{\boldsymbol{x},k}$ is the kth component of the vector $\boldsymbol{e}_n^{\boldsymbol{x}}$.

The convergence of the presented Lobatto IIIA-IIIB scheme is shown in Fig. 9. For all simulations used in this convergence study, the absolute tolerance used for solving the time steps with the semismooth Newton method was 10^{-14}. Moreover, in both cases, the time span $I = [0, 1.6]$ is analyzed only, i.e., in both cases the slip-stick transition is excluded from the convergence study. With this choice, the motion in case 1 is smooth and corresponds to the motion of a point mass, which moves along the slope and is subjected to sliding friction described by a single-valued force law. For case 1, the convergence rate $2s - 2$ known for the Lobatto IIIA-IIIB schemes [3] for bilaterally constrained systems is observed in Fig. 9. This is no surprise, as for case 1, the contact is always closed and the discrete contact law is consistent with the discretization of bilateral constraints, i.e., the contact law reduces to the discretization of a bilateral constraint. This showcases the importance of discretizing the normal contact law consistently with the bilateral constraints. Otherwise, the good convergence behavior could be lost.

For case 2, Fig. 9 shows a convergence behavior that is not monotonous and appears to be of first order. This convergence behavior is in line with the results for other event-capturing schemes [9,27]. As all event-capturing schemes, the presented methods smear the impact law over a whole time interval and use the velocity at the beginning and the end of the time step as pre- and post-impact velocities, respectively. It is clear that this kind of discretization becomes exact for $\Delta t \to 0$. Hence, as observed, the presented scheme is convergent. The apparent first order convergence might be explained by the fact that the just mentioned smearing of the impact law resolves the exact impact time only to the order of the time step.

7. Conclusion

We presented a family of Lobatto IIIA-IIIB methods for the simulation of mechanical systems with frictional contact. All introduced schemes treat the nonsmoothness arising from the frictional contact in an event-capturing way. For the derivation of the schemes, we have stated a generalized version of the principle of virtual action. Subsequently, the principle of virtual action

was discretized using a Lobatto quadrature and the schemes resulted as the necessary and sufficient conditions of the variational principle. For the discretization of the frictional contact law, the unilateral constraints were discretized consistently with the bilateral constraints. Additionally, the Newton-type (frictional) impact law has been discretized by demanding the impact law integrally over a whole time step.

For the derivation of the schemes it was crucial to formulate a version of the principle of virtual action that treats the generalized coordinates, the velocities and the momenta as a priori separated functions of time. By doing so, their interrelation is encoded by the functional defining the virtual action. Hence, discretizing the principle of virtual action leads to discretized versions of the interrelations between the positions, velocities and momenta.

After deriving the presented schemes, these have been numerically validated using benchmark examples. The bouncing ball example shows that the presented Lobatto IIIA-IIIB family is indeed event-capturing. All schemes can overcome accumulation points and the discretization of the contact law correctly reproduces the analytical solutions in case of frictional impact and slip-stick transitions. With the slider-crank mechanism, we could showcase that the presented schemes can cope with systems with multiple contact points that close simultaneously and with systems with high frequency contact patterns. Finally, the point mass on a slope benchmark was used to showcase the drop in convergence order for discontinuous motions, which is typical for event-capturing schemes. Moreover, with the same benchmark, we highlighted the importance of having a discrete contact law that is consistent with the discretization of the bilateral constraints. As a consequence of this consistency, the convergence order $2s-2$ of the Lobatto IIIA-IIIB schemes for bilaterally constrained systems is retrieved for motions with persistent contact. This result is of major importance, as many engineering applications have motions with persistent contact rather than impacting motions, e.g., rolling contact in vehicle dynamics simulations.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

[1] J.E. Marsden, M. West, Discrete mechanics and variational integrators, Acta Numerica 2001 10 (2001) 357–514.
[2] L. Jay, Symplectic partitioned Runge–Kutta methods for constrained Hamiltonian systems, SIAM J. Numer. Anal. 33 (1) (1996) 368–387.
[3] R. Sato Martín de Almagro, Discrete Mechanics for Forced and Constrained Systems (Ph.D. thesis), Universidad Complutense Madrid, 2019.
[4] C. Glocker, Set-Valued Force Laws, Springer Berlin Heidelberg, Berlin, Heidelberg, 2001.
[5] B. Brogliato, Nonsmooth Mechanics: Models, Dynamics and Control, Springer, 2016.
[6] J.J. Moreau, Unilateral contact and dry friction in finite freedom dynamics, in: J.J. Moreau, P.D. Panagiotopoulos (Eds.), Non-Smooth Mechanics and Applications. CISM Courses and Lectures, Springer, Wien, 1988, pp. 1–82.
[7] V. Acary, M. Brémond, O. Huber, On solving contact problems with coulomb friction: formulations and numerical comparisons, in: R. Leine, V. Acary, O. Brüls (Eds.), Advanced Topics in Nonsmooth Dynamics: Transactions of the European Network for Nonsmooth Dynamics, Springer International Publishing, Cham, 2018, pp. 375–457.
[8] R.C. Fetecau, J.E. Marsden, M. Ortiz, M. West, Nonsmooth Lagrangian mechanics and variational collision integrators, SIAM J. Appl. Dyn. Syst. 2 (3) (2003) 381–416.
[9] C. Studer, Numerics of Unilateral Contacts and Friction: Modeling and Numerical Time Integration in Non-Smooth Dynamics, in: Lecture Notes in Applied and Computational Mechanics, vol. 47, Springer, Berlin, Heidelberg, 2009.
[10] G. Capobianco, S.R. Eugster, Time finite element based Moreau-type integrators, Internat. J. Numer. Methods Engrg. 114 (3) (2018) 215–231.
[11] V. Acary, Projected event-capturing time-stepping schemes for Nonsmooth mechanical systems with unilateral contact and Coulomb's friction, Comput. Methods Appl. Mech. Engrg. 256 (2013) 224–250.
[12] J. Galvez, F.J. Cavalieri, A. Cosimo, O. Brüls, A. Cardona, A nonsmooth frictional contact formulation for multibody system dynamics, Internat. J. Numer. Methods Engrg. 121 (16) (2020) 3584–3609.
[13] G. Capobianco, J. Harsch, S.R. Eugster, R.I. Leine, A nonsmooth generalized-alpha method for mechanical systems with frictional contact, Internat. J. Numer. Methods Engrg. 122 (22) (2021) 6497–6526.
[14] J. Harsch, G. Capobianco, S.R. Eugster, R.I. Leine, An extension to the RATTLE integrator for mechanical systems with frictional contact, Nonlinear Anal. Hybrid Syst. (submitted 2023).
[15] T. Schindler, V. Acary, Timestepping schemes for nonsmooth dynamics based on discontinuous Galerkin methods: Definition and outlook, Math. Comput. Simulation 95 (2014) 180–199.
[16] J.J. Moreau, Bounded variation in time, in: J.J. Moreau, P.D. Panagiotopoulos, G. Strang (Eds.), Topics in Nonsmooth Mechanics, Birkhäuser Verlag, Basel, 1988, pp. 1–74.
[17] E. De Giorgi, L. Ambrosio, Un Nuovo Tipo Di Funzionale Del Calcolo Delle Variazioni, Atti della Accademia Nazionale dei Lincei. Classe di Scienze Fisiche, Matematiche e Naturali. Rendiconti Lincei. Matematica e Applicazioni 82 (2) (1988) 199–210.
[18] C. Glocker, Simulation of Hard Contacts with Friction: An Iterative Projection Method, in: A. Johann, H.-P. Kruse, F. Rupp, S. Schmitz (Eds.), Recent Trends in Dynamical Systems, Vol. 35, Springer, Basel, 2013, pp. 493–515, Series Title: Springer Proceedings in Mathematics & Statistics.
[19] C. Glocker, Energetic consistency conditions for standard impacts. Part I: Newton-type inequality impact laws and Kane's example, Multibody Syst. Dyn. 29 (1) (2013) 77–117, Publisher: Springer Netherlands.
[20] T. Winandy, M. Baumann, R.I. Leine, Variational analysis of inequality impact laws for perfect unilateral constraints, in: Advanced Topics in Nonsmooth Dynamics, Springer, 2018, pp. 47–92.
[21] L.A. Pars, A Treatise on Analytical Dynamics, Heinemann, 1965.

[22] N. Bou-Rabee, J.E. Marsden, Hamilton–Pontryagin integrators on Lie groups part I: Introduction and structure-preserving properties, Found. Comput. Math. 9 (2) (2009) 197–219.
[23] L.O. Jay, Lobatto methods, in: B. Engquist (Ed.), Encyclopedia of Applied and Computational Mathematics, Springer, Berlin, Heidelberg, 2015, pp. 817–826.
[24] E. Hairer, G. Wanner, Solving Ordinary Differential Equations II, second ed., Springer, 2002.
[25] P. Alart, A. Curnier, A mixed formulation for frictional contact problems prone to Newton like solution methods, Comput. Methods Appl. Mech. Engrg. 92 (1991) 353–375.
[26] R.I. Leine, H. Nijmeijer, Dynamics and Bifurcations of Non-Smooth Mechanical Systems, in: Lecture Notes in Applied and Computational Mechanics, vol. 18, Springer, Berlin, Heidelberg, 2004.
[27] O. Brüls, V. Acary, A. Cardona, Simultaneous enforcement of constraints at position and velocity levels in the nonsmooth generalized-α scheme, Comput. Methods Appl. Mech. Engrg. 281 (2014) 131–161.