Computers and Structures 322 (2026) 108121

Contents lists available at ScienceDirect

Computers and Structures

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

Efficient optimization of clutched inerter dampers using mixed Lagrangian formalism

Yixuan Zhang [a], Oren Lavan [b], Christian Málaga-Chuquitaype [a,*]

[a] Imperial College London, Department of Civil and Environmental Engineering, London, SW7 2AZ, UK
[b] Technion - Israel Institute of Technology, Faculty of Civil and Environmental Engineering, Haifa, 3200003, Israel

HIGHLIGHTS

- An efficient performance-based optimization framework for clutched inerter dampers is proposed.
- Mixed Lagrangian Formalism is adopted as the time integration scheme to obtain the adjoint sensitivity of the optimization problem.
- Two case studies reveal the parametric characteristics of clutched inerter dampers and demonstrate the computational efficiency of the proposed method.

ARTICLE INFO

Keywords:
Adjoint sensitivity
Gradient optimization
Clutched inerter damper
Mixed Lagrangian formalism

ABSTRACT

Clutched Inerter Dampers integrate inerters with one-way clutches and dampers, enabling the inerters to disengage from the host structure and dissipate energy, suppressing undesirable energy feedback and avoiding the direct alteration of the structural period, making clutched inerter dampers particularly attractive for vibration mitigation and energy harvesting applications. However, the nonlinearity and discontinuity introduced by clutch engagement-disengagement mechanisms pose significant challenges for their accurate numerical modelling and optimization-based design. Existing optimization approaches for such systems are predominantly based on simplified or linearised numerical models, or on exhaustive parameter scanning, which either fail to capture the true nonlinear behaviour or become impractical for realistic design spaces.

We present an adjoint-based, gradient-driven optimization framework for structures equipped with clutched inerter dampers in which the Mixed Lagrangian Formalism is employed as the time-integration scheme. Within the proposed framework, the nonlinear behaviour of clutched inerter dampers is fully captured, while computational efficiency and numerical robustness are achieved through the Mixed Lagrangian Formalism, which reduces the cost of individual response-history analyses and enhances stability in the presence of non-smooth dynamics. In addition, adjoint-based sensitivity analysis significantly decreases the number of simulations required during the optimization process. The framework enables efficient optimization of design parameters as demonstrated through a series of representative case studies. Our results show that, despite strong nonlinearity and discontinuous system responses, analytical gradients can be consistently derived, leading to substantial reductions in computational cost and improved optimization efficiency.

While the actual performance may be influenced by the characteristics of the design landscape and the choice of initial conditions, the proposed framework provides a robust and extensible basis for further methodological developments. It can be readily extended in future work to accommodate alternative optimization strategies or enhanced formulations.

1. Introduction

Inerters are innovative mechanical components that generate forces proportional to the relative acceleration between their terminals, distinguishing them from traditional vibration control devices such as dampers and springs [1–10]. Their unique ability to provide substantial inertial forces without a corresponding increase in mass has led to widespread interest in both structural and mechanical engineering applications, including seismic protection, vehicle suspension, and energy

* Corresponding author.
Email address: c.malaga@imperial.ac.uk (C. Málaga-Chuquitaype).

https://doi.org/10.1016/j.compstruc.2026.108121
Received 19 October 2025; Accepted 19 January 2026
Available online 27 January 2026
0045-7949/© 2026 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

harvesting systems [11–14]. However, inerters may transfer energy back to the structure, acting as a power output, thus increasing the demand under certain conditions [15,16]. Furthermore, due to the amplified inertia effect, the inerter elongates the natural period of the structure; thus, it may increase the peak deformation response of stiff structures under seismic loading.

To address these drawbacks, clutched inerter dampers (CIDs) have been proposed [15,17–19]. CIDs integrate inerters with one-way clutches and dampers (see Fig. 1(a), enabling the device to disengage from the main structure under specific conditions. This mechanism not only enhances overall energy dissipation, but also prevents the undesirable transmission of energy back into the structure, thereby improving both safety and performance. Moreover, the disengagement capability of CIDs opens new possibilities for energy harvesting and adaptive control strategies [20–22]. Despite these promising features, the introduction of clutch mechanisms leads to non-smooth and discontinuous responses, which pose significant challenges to their accurate numerical modelling and practical design [21,23].

A simplified method proposed by Makris and Kampas [15], assuming engagement whenever the driven terminal is accelerating, has been widely adopted in many studies [24–28]. However, this approach significantly overestimates the effectiveness of the inerter, as it unrealistically assumes that the device can re-engage with the structure during every loading cycle. This implies, incorrectly, that all the energy stored in the inerter can be fully dissipated immediately after disengagement [16]. In reality, the engagement efficiency, defined as the ratio of engaged time to total time for a clutched inerter, is typically very low. Other models have been proposed [9,17,18], but they all require monitoring the state of the system and manual control of the time integration process by switching between engagement and disengagement states. Furthermore, traditional numerical integration schemes struggle to accurately capture engagement and disengagement in the presence of these discontinuities, often requiring small time steps, thus increasing computational costs [23]. Therefore, there is a clear need for a numerical formulation that can robustly accommodate discontinuous switching behaviour without excessive computational cost.

In this context, the Mixed Lagrangian Formalism (MLF) has been proposed for modelling extreme nonlinear dynamic responses. MLF was first developed for the transient analysis of frames with lumped elasto-plastic hinges [29] and was later extended to progressive collapse simulations of large structures, incorporating strength degradation, elemental fracture [30], and contact [31]. Its applicability also spans multi-physics problems, including thermoelastic and thermoplastic structural responses [2,32,33]. By employing a mixed-coordinate representation that simultaneously considers global displacements and local internal forces, MLF naturally accommodates the switching behaviour of clutches without sacrificing computational efficiency or accuracy. This approach enables stable and accurate response-history analysis, even in the presence of abrupt state changes, making it particularly suitable for the simulation of CIDs. In addition, the MLF framework allows for a larger time step size for modelling non-smooth behaviour [29]. This is particularly relevant to the computationally intensive task, such as the optimal design of CIDs. Since the performance of these nonlinear and discontinuous systems depends on the careful selection of design parameters. However, MLF-based analysis alone does not address the challenge of determining optimal CID parameters for performance-driven applications.

Efficient optimization of nonlinear dynamic systems often relies on adjoint-based sensitivity analysis, which provides gradients at a cost essentially independent of the number of design variables [34]. When the design dimension is greater relative to the number of objective or constraint functions, adjoint methods offer significant computational savings over direct differentiation [35]. Hence, they have been successfully applied to hysteretic systems with nonlinear viscous dampers [36], topology optimization of energy-dissipation devices [37], and the seismic design of multi-rocking structures [38].

While the adjoint sensitivity analysis of nonlinear dynamical systems remains an active area of research, such methodologies have not been extended to the design of CIDs. Recent studies on CID optimization have primarily relied on parametric sweeps or simplified linearisation assumptions, often treating the device behaviour in an oversimplified manner with a limited number of devices [24,25,27]. Although these approaches provide insight into general performance trends, they are limited in their ability to explore high-dimensional design spaces and to account for the strongly nonlinear and non-smooth clutching behaviour inherent to CIDs. To date, no rigorous gradient-based optimization framework has been formulated that simultaneously considers the discontinuous engagement–disengagement dynamics of CIDs and the need for efficient, scalable design tools.

Our earlier work established an MLF framework capable of performing nonlinear response-history analysis of CIDs [23]; however, it did not address the performance-driven design or optimization of these devices. The present study advances the state of the art by formulating a dedicated performance-based optimization framework for CIDs. The contributions of the present research are summarised as follows:

- **A performance-based optimization framework for CID design:** This study introduces the first unified procedure that embeds CID dynamics, performance objectives, and design variables into a coherent optimization setting, enabling the systematic optimization of inertance, damping, and other device characteristics. The proposed framework can be readily extended to incorporate additional inerter-related parameters and structural performance measures, thereby offering substantial flexibility for future design studies.
- **Adjoint sensitivity formulation specialised for CIDs within the MLF analysis framework:** Although adjoint methods are widely used in structural optimization, their application to CID systems with abrupt engagement–disengagement transitions has remained unexplored. This work derives the adjoint formulation for the MLF scheme, ensuring consistent and efficient gradient computation for CID systems.
- **Validation through two optimization case studies:** The applicability and efficiency of the framework are demonstrated through two representative CID optimization problems, highlighting its ability to enhance dynamic performance while significantly reducing computational effort.

Together, these contributions advance CID research beyond simulation-oriented analysis toward a rigorous and extensible optimization paradigm, filling a critical methodological gap in the design of next-generation inertial damping devices.

2. Numerical modelling

The numerical modelling framework adopted in this study is based on the MLF, for which a detailed theoretical development has been presented in our previous work [23]. In this paper, the essential components of the MLF are restated with refined notation, streamlined derivations, and optimised organisation to facilitate their direct application to the CID system examined herein. Section 2.1 briefly revisits the core MLF concepts and governing equations relevant to nonlinear response history analysis (NRHA), while Section 2.2 extends this framework to incorporate the CID element, highlighting the modifications required to represent clutch engagement-disengagement behaviour and inertance effects within the MLF structure. This consolidated presentation ensures completeness and clarity, providing the basis for the numerical implementation used in the subsequent analyses.

2.1. Overview of MLF

The MLF is adopted here as the time integration scheme of the Nonlinear Response History Analysis (NRHA) due to its robustness and efficiency. In the classic Lagrangian formalism, the system is described by a Lagrangian $L(\mathbf{q}, \dot{\mathbf{q}}, t)$ with generalised coordinates $\mathbf{q}$ (usually

independent degrees of freedom), their derivatives $\dot{\mathbf{q}}$, and time t. Unlike the classic Lagrangian formalism, the MLF adopts two sets of state variables, including the global degrees of freedom (DOFs) $\mathbf{u}$ and the time integral of the element internal forces(impulse) $\mathbf{J}$. In MLF, the Lagrangian L and the dissipation function Φ take the following form [29]:

$$L(\mathbf{u},\dot{\mathbf{u}},\mathbf{J},\dot{\mathbf{J}}) = \frac{1}{2}\dot{\mathbf{u}}^{\top}\mathbf{M}\dot{\mathbf{u}} + \frac{1}{2}\dot{\mathbf{J}}^{\top}\mathbf{A}\dot{\mathbf{J}} + \mathbf{J}^{\top}\mathbf{B}\dot{\mathbf{u}} + \mathbf{P}^{\top}\mathbf{u} + \mathbf{1}^{\top}\mathbf{g}(\dot{\mathbf{J}}) \tag{1}$$

$$\Phi(\dot{\mathbf{u}},\dot{\mathbf{J}}) = \frac{1}{2}\dot{\mathbf{u}}^{\top}\mathbf{C}\dot{\mathbf{u}} + \mathbf{1}^{\top}\boldsymbol{\phi}(\dot{\mathbf{J}}) \tag{2}$$

Where $\mathbf{M}$ and $\mathbf{C}$ are the mass and damping matrices in global coordinates, $\mathbf{A}$ is the flexibility matrix in local coordinates, $\mathbf{B}$ is the connectivity matrix that relates local to global coordinates, and $\mathbf{P}$ is an external load vector in global coordinates. This formulation allows us to describe the system with global and local coordinates simultaneously [29]. The terms $\mathbf{1}^{\top}\boldsymbol{\phi}(\dot{\mathbf{J}})$ and $\mathbf{1}^{\top}\mathbf{g}(\dot{\mathbf{J}})$ reflect the velocity-based and displacement-based nonlinear terms, respectively. Nonlinearities, such as yielding and contact or tension-only behaviour, are incorporated through a suitable formulation of the $\boldsymbol{\phi}(\dot{\mathbf{J}})$ and $\mathbf{g}(\dot{\mathbf{J}})$ [31]. For example, gap opening/contact behaviour and elastic perfect plastic behaviour of the i^{th} local coordinate can be formulated as [39]:

$$\begin{aligned}
\mathbf{g}_i(\dot{\mathbf{J}}) &= \begin{cases} 0, & if \quad \dot{\mathbf{J}}_i - \mathbf{F}_{c,i} \le 0 \\ non-negative, & otherwise \end{cases}, \\
\boldsymbol{\phi}_i(\dot{\mathbf{J}}) &= \begin{cases} 0, & if \quad |\dot{\mathbf{J}}_i| \le \mathbf{F}_{y,i} \\ non-negative, & otherwise \end{cases}
\end{aligned} \tag{3}$$

Where $\mathbf{F}_{c,i}$ are gap opening threshold forces and $\mathbf{F}_{y,i}$ are elemental yield limits. Upon substituting the Lagrangian and dissipation functions into the Euler-Lagrange equations, the following equilibrium equation in global coordinates and compatibility equation in local coordinates are obtained:

$$\mathbf{M}\ddot{\mathbf{u}} + \mathbf{C}\dot{\mathbf{u}} + \mathbf{B}\dot{\mathbf{J}} = \mathbf{P} \tag{4}$$

$$\mathbf{A}\ddot{\mathbf{J}} + \frac{\partial\boldsymbol{\phi}(\dot{\mathbf{J}})}{\partial\dot{\mathbf{J}}} + \frac{d}{dt}\left(\frac{\partial\mathbf{g}(\dot{\mathbf{J}})}{\partial\dot{\mathbf{J}}}\right) = \mathbf{B}^{\top}\dot{\mathbf{u}} \tag{5}$$

For a lumped mass system, the global DOFs can be partitioned into three categories: 1) those with dynamic masses or damping, $\mathbf{u}_1$, 2) those without dynamic masses and damping, $\mathbf{u}_2$, 3) those with prescribed displacement or velocity constraints, $\mathbf{u}_3$. This partition is equivalent to a static condensation process, effectively resolving the singularity of the characteristic matrix for lumped mass systems [29]:

$$\begin{bmatrix} \mathbf{M} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{Bmatrix} \ddot{\mathbf{u}}_1 \\ \ddot{\mathbf{u}}_2 \\ \ddot{\mathbf{u}}_3 \end{Bmatrix} + \begin{bmatrix} \mathbf{C} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{Bmatrix} \dot{\mathbf{u}}_1 \\ \dot{\mathbf{u}}_2 \\ \dot{\mathbf{u}}_3 \end{Bmatrix} + \begin{Bmatrix} \mathbf{B}_1 \\ \mathbf{B}_2 \\ \mathbf{B}_3 \end{Bmatrix} \dot{\mathbf{J}} = \begin{Bmatrix} \mathbf{P}_1 \\ \mathbf{P}_2 \\ \mathbf{R} \end{Bmatrix} \tag{6}$$

After simplification, the governing equations of lumped mass systems become:

$$\begin{cases} \mathbf{M}\ddot{\mathbf{u}}_1 + \mathbf{C}\dot{\mathbf{u}}_1 + \mathbf{B}_1\dot{\mathbf{J}} = \mathbf{P}_1 \\ \mathbf{B}_2\dot{\mathbf{J}} = \mathbf{P}_2 \\ \mathbf{A}\ddot{\mathbf{J}} + \frac{\partial\boldsymbol{\phi}(\dot{\mathbf{J}})}{\partial\dot{\mathbf{J}}} + \frac{d}{dt}\left(\frac{\partial\mathbf{g}(\dot{\mathbf{J}})}{\partial\dot{\mathbf{J}}}\right) = \mathbf{B}_1^{\top}\dot{\mathbf{u}}_1 + \mathbf{B}_2^{\top}\dot{\mathbf{u}}_2 \end{cases} \tag{7}$$

which are discretised at $t_{n+1/2}$ using the central difference method. After some mathematical manipulation [29], the strong form of the governing equations can be converted into the following weak form optimization problem:

$$\begin{aligned}
\min \quad & \frac{1}{2}\mathbf{f}_{n+1}^{\top}\mathbf{H}\mathbf{f}_{n+1} + \bar{\mathbf{b}}_n^{\top}\mathbf{f}_{n+1} \\
s.t. \quad & \mathbf{B}_2\mathbf{f}_{n+1} = \mathbf{P}_{2,n+1} \\
& \frac{1}{2}\boldsymbol{\phi}(\mathbf{f}_{n+1}) \le 0 \\
& \frac{1}{h}\mathbf{g}(\mathbf{f}_{n+1}) \le 0
\end{aligned} \tag{8}$$

Where $\mathbf{f}_n = \dot{\mathbf{J}}_n$ are local elemental forces, $\mathbf{H}$ is the equivalent dynamic flexibility matrix, which is dual to that of the equivalent dynamic stiffness matrix of Newmark's method with $\gamma = 1/2$ and$\beta = 1/4$, and $\mathbf{b}_n$ contains the system state information from the previous time step.

This equivalent optimization problem can be interpreted through the principle of minimum incremental complementary potential energy: 'at each time step, the force vector $\mathbf{f}_{n+1}$ should minimise the incremental complementary potential energy while satisfying the compatibility constraints of non-linearities' [29]. This weak form enhances stability, allowing for larger time-step sizes compared to traditional methods, especially when sudden changes (such as clutch engagement, fracture, contact, etc.) are expected during the response of the system [40].

2.2. MLF formulation of the CID element

Recently, a CID element based on the MLF framework was proposed by Zhang et al. [23], consisting of a parallel damper and an ideal inerter connected to a one-way elastic perfect plastic element(see Fig. 1(b). This element can automatically capture the non-smooth dynamics of the clutching mechanism during time step marching while maintaining computational efficiency and numerical stability.

In the proposed CID element formulation, the inertance is implemented by including an additional energy term $\frac{1}{2}\mathbf{J}^{\top}\mathbf{m}_r\mathbf{J}$ in the Lagrangian function:

$$L(\mathbf{u},\dot{\mathbf{u}},\mathbf{J},\dot{\mathbf{J}}) = \frac{1}{2}\dot{\mathbf{u}}^{\top}\mathbf{M}\dot{\mathbf{u}} + \frac{1}{2}\dot{\mathbf{J}}^{\top}\mathbf{A}\dot{\mathbf{J}} + \mathbf{J}^{\top}\mathbf{B}\dot{\mathbf{u}} + \mathbf{P}^{\top}\mathbf{u} + \mathbf{1}^{\top}\mathbf{g}(\dot{\mathbf{J}}) + \frac{1}{2}\mathbf{J}^{\top}\mathbf{m}_r\mathbf{J} \tag{9}$$

where $\mathbf{m}_r = \mathbf{b}^{-1}$, and $\mathbf{b}$ is the inertance matrix that resembles the equivalent linear mass of the inerters in the system.

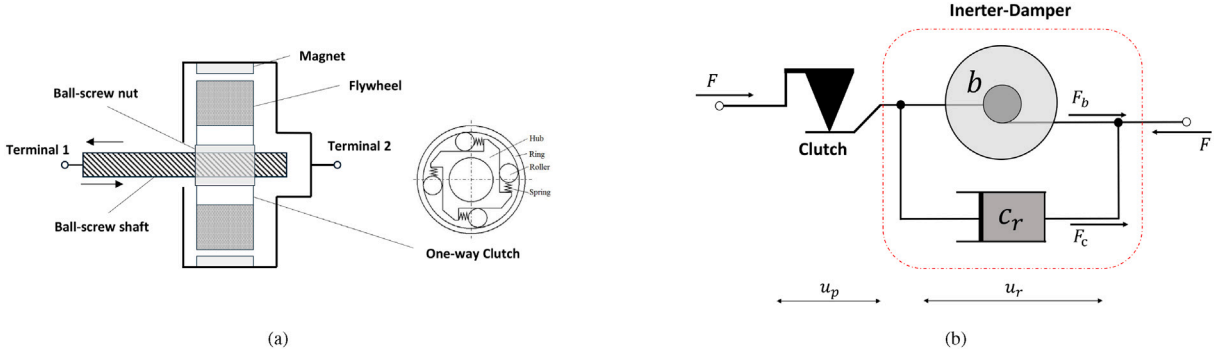


Fig. 1. (a) Schematic representation of a typical CID using a ball-screw mechanism, and (b) The MLF-based CID element proposed by Zhang et al. [23], where inertance b is the equivalent linear mass of the inerter and c_r the associate damping coefficient that reflects the viscous-like eddy current dissipation of the inerter.

While the clutching effect is modelled as a one-way elastic perfect-plastic unit, with zero yield strength in one direction and infinite yield strength in the opposite direction, hence the clutch acts as a rigid link (engagement) when the force is positive or negative but immediately yields (disengagement) once the force reverses its direction. Particularly, the motion history of the idling (disengaged) inerter is captured by the plastic deformation u_p of this unit. In this way, the engagement state switch is automatically captured by the local variable of this unit. Therefore, the clutch unit in MLF holds the following dissipation function constraint and the corresponding equivalent constraint on element force:

$$\frac{1}{2}\boldsymbol{\phi}_i(\dot{\mathbf{J}}) = \begin{pmatrix} 0, & if \ \ \mathbf{s}_{c,i}\dot{\mathbf{J}}_i \le 0 \\ non-negative, & otherwise \end{pmatrix} \le 0 \rightleftharpoons \mathbf{s}_{c,i}\mathbf{f}_i \le 0 \tag{10}$$

where $\mathbf{s}_c$ indicates whether the i^{th} inerter element is working as a tension-only ($\mathbf{s}_c = -1$) or compression-only($\mathbf{s}_c = 1$) unit. By incorporating Zhang et al.'s formulation with the governing equations (Eq. 4) above, we have:

$$\begin{cases} \mathbf{M}\ddot{\mathbf{u}}_1 + \mathbf{C}\dot{\mathbf{u}}_1 + \mathbf{B}_1\dot{\mathbf{J}} = \mathbf{P}_1 \\ \mathbf{B}_2\dot{\mathbf{J}} = \mathbf{P}_2 \\ \mathbf{A}\ddot{\mathbf{J}} + \frac{\partial \boldsymbol{\phi}(\dot{\mathbf{J}})}{\partial \dot{\mathbf{J}}} + \frac{d}{dt}\left(\frac{\partial \mathbf{g}(\dot{\mathbf{J}})}{\partial \dot{\mathbf{J}}}\right) + \dot{\mathbf{u}}_r = \mathbf{B}_1^\top \dot{\mathbf{u}}_1 + \mathbf{B}_2^\top \dot{\mathbf{u}}_2 \end{cases} \tag{11}$$

The term $\dot{\mathbf{u}}_r$ represents the equivalent linear velocity of the inerter, which incorporates the parallel damping effect in the following form [23]:

$$\dot{\mathbf{u}}_r(\dot{\mathbf{J}}, t) = e^{-\mathbf{c_r}\mathbf{m_r}t}\mathbf{m_r}\int e^{\mathbf{c_r}\mathbf{m_r}t}\dot{\mathbf{J}}(t)dt \tag{12}$$

Replacing the first derivative of global displacement with velocity $\dot{\mathbf{u}} = \mathbf{v}$ and the local impulse with force $\dot{\mathbf{J}} = \mathbf{f}$, and applying discretization using central differences at steps $n + \frac{1}{2}$, the time discretization is:

$$\begin{cases} \mathbf{v}_{1,n+1} = \mathbf{y}_1\bar{\mathbf{P}}_{1,n} + \mathbf{y}_2(\mathbf{f}_{n+1} + \mathbf{f}_n) + \mathbf{y}_3\mathbf{v}_{1,n} \\ \mathbf{v}_{r,n+1} = \mathbf{y}_4\mathbf{f}_{n+1} + \mathbf{y}_5\mathbf{v}_{r,n} + \mathbf{y}_6\mathbf{f}_n \\ \mathbf{B}_2\mathbf{f}_{n+1} = \mathbf{P}_{2,n+1} \\ \mathbf{H}\mathbf{f}_{n+1} + \frac{1}{2}\left.\frac{\partial \boldsymbol{\phi}}{\partial \mathbf{f}}\right|_{n+1} + \frac{1}{h}\left.\frac{\partial \mathbf{g}}{\partial \mathbf{f}}\right|_{n+1} - \mathbf{B}_2^\top \mathbf{v}_{2,n+1} + \bar{\mathbf{b}}_n = \mathbf{0} \end{cases} \tag{13}$$

where:

$$\bar{\mathbf{b}}_n = \bar{\mathbf{H}}\mathbf{f}_n - \frac{1}{h}\left.\frac{\partial g}{\partial \mathbf{f}}\right|_n + \frac{1}{2}\left.\frac{\partial \phi}{\partial \mathbf{f}}\right|_n + \mathbf{y}_7\bar{\mathbf{P}}_{1,n} + \mathbf{y}_8\mathbf{v}_{1,n} + \mathbf{y}_9\mathbf{v}_{r,n}$$

$$\mathbf{y}_1 = \left(\frac{\mathbf{M}}{h} + \frac{\mathbf{C}}{2}\right)^{-1}$$

$$\mathbf{y}_2 = -\left(\frac{\mathbf{M}}{h} + \frac{\mathbf{C}}{2}\right)^{-1}\frac{\mathbf{B}_1}{2}$$

$$\mathbf{y}_3 = \left(\frac{\mathbf{M}}{h} + \frac{\mathbf{C}}{2}\right)^{-1}\left(\frac{\mathbf{M}}{h} - \frac{\mathbf{C}}{2}\right)$$

$$\mathbf{y}_4 = \gamma\mathbf{m}_r h \quad , (\gamma = 0.5)$$

$$\mathbf{y}_5 = e^{-\mathbf{c}_r\mathbf{m}_r h}$$

$$\mathbf{y}_6 = (1-\gamma)e^{-\mathbf{c}_r\mathbf{m}_r h}\mathbf{m}_r h$$

$$\mathbf{y}_7 = -\frac{\mathbf{B}_1^\top}{2}\left(\frac{\mathbf{M}}{h} + \frac{\mathbf{C}}{2}\right)^{-1}$$

$$\mathbf{y}_8 = -\frac{\mathbf{B}_1^\top}{2}\left[\left(\frac{\mathbf{M}}{h} + \frac{\mathbf{C}}{2}\right)^{-1}\left(\frac{\mathbf{M}}{h} - \frac{\mathbf{C}}{2}\right) + \mathbf{I}\right]$$

$$\mathbf{y}_9 = \frac{1}{2}\left(\mathbf{I} + e^{-\mathbf{c}_r\mathbf{m}_r h}\right)$$

$$\mathbf{H} = \left[\frac{\mathbf{A}}{h} + \frac{\mathbf{B}_1^\top}{2}\left(\frac{\mathbf{M}}{h} + \frac{\mathbf{C}}{2}\right)^{-1}\frac{\mathbf{B}_1}{2} + \frac{\mathbf{y}_4}{2}\right]$$

$$\bar{\mathbf{H}} = \left[-\frac{\mathbf{A}}{h} + \frac{\mathbf{B}_1^\top}{2}\left(\frac{\mathbf{M}}{h} + \frac{\mathbf{C}}{2}\right)^{-1}\frac{\mathbf{B}_1}{2} + \frac{\mathbf{y}_6}{2}\right]$$

$$\bar{\mathbf{P}}_{1,n} = \left(\frac{\mathbf{P}_{1,n+1} + \mathbf{P}_{1,n}}{2}\right)$$

To solve the discretised governing (Eq. 13), a weak form time step optimization problem similar to Eq. (8) can be obtained. In this context, the optimization problem described above is implemented by the 'quadprog' function in MATLAB's optimization toolbox [41], which has the following structure:

$$\begin{aligned} \min \quad & \frac{1}{2}\mathbf{f}_{n+1}^\top \mathbf{H}\mathbf{f}_{n+1} + \bar{\mathbf{b}}_n^\top \mathbf{f}_{n+1} \\ s.t. \quad & \mathbf{A}_{eq}\mathbf{f}_{n+1} = \mathbf{b}_{eq} \\ & \mathbf{A}_{leq}\mathbf{f}_{n+1} \le \mathbf{b}_{leq} \end{aligned} \tag{14}$$

with velocity recursive rules:

$$\mathbf{v}_{1,n+1} = \mathbf{y}_1\bar{\mathbf{P}}_{1,n} + \mathbf{y}_2(\mathbf{f}_{n+1} + \mathbf{f}_n) + \mathbf{y}_3\mathbf{v}_{1,n} \tag{15}$$

$$\mathbf{v}_{r,n+1} = \mathbf{y}_4\mathbf{f}_{n+1} + \mathbf{y}_5\mathbf{v}_{r,n} + \mathbf{y}_6\mathbf{f}_n \tag{16}$$

Comparing the original governing equation Eq. (13), it is obvious that $\mathbf{A}_{eq} = \mathbf{B}_2$, $\mathbf{b}_{eq} = \mathbf{P}_{2,n+1}$, $\mathbf{A}_{leq} = \begin{bmatrix}\mathbf{A}_\phi \\ \mathbf{A}_g\end{bmatrix}$, $\mathbf{b}_{leq} = \begin{bmatrix}\mathbf{b}_\phi \\ \mathbf{b}_g\end{bmatrix}$, where $\mathbf{A}_\phi$, $\mathbf{b}_\phi$, and $\mathbf{A}_g$, $\mathbf{b}_g$ indicate the displacement-based constraints and the velocity-based constraints, respectively.

The arbitrary terms $\left.\frac{\partial \boldsymbol{\phi}}{\partial \mathbf{f}}\right|_n$, $\left.\frac{\partial g}{\partial \mathbf{f}}\right|_n$, and $\mathbf{v}_{2,n}$ can be calculated based on the corresponding Lagrange multipliers:

$$-\mathbf{B}_2^\top \mathbf{v}_{2,n+1} = \mathbf{A}_{eq}^\top \lambda_{eq,n+1}, \quad \frac{1}{2}\left.\frac{\partial \boldsymbol{\phi}}{\partial \mathbf{f}}\right|_{n+1} = \mathbf{A}_\phi^\top \lambda_{\phi,n+1}, \quad \frac{1}{h}\left.\frac{\partial g}{\partial \mathbf{f}}\right|_{n+1} = \mathbf{A}_g^\top \lambda_{g,n+1} \tag{17}$$

Thus, $\bar{\mathbf{b}}_n$ becomes:

$$\bar{\mathbf{b}}_n = \bar{\mathbf{H}}\mathbf{f}_n + \mathbf{A}_{leq}^\top \mathbf{s}\lambda_{leq,n} + \mathbf{y}_7\bar{\mathbf{P}}_{1,n} + \mathbf{y}_8\mathbf{v}_{1,n} + \mathbf{y}_9\mathbf{v}_{r,n} \tag{18}$$

where $\mathbf{s} = \begin{bmatrix}\mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{-1}\end{bmatrix}$, $\lambda_{leq} = \begin{bmatrix}\lambda_\phi \\ \lambda_g\end{bmatrix}$ The formulation above has been rigorously validated in terms of accuracy, stability, and efficiency in Zhang et al. [23].

3. Formulation of the optimization problem

The non-smooth dynamics of the CIDs lead to an implicit relationship between the design variables and the system response, which impedes the design process. In this section, a framework for the optimization problem is established, including design variables, objective function, and constraints. The objective is to achieve the lowest construction cost of CIDs while satisfying certain performance requirements. This framework is adaptable to various objectives and constraints (or limit states), expressed in terms of different engineering demand parameters (EDPs). The optimization problem can be solved using adjoint sensitivity analysis to efficiently compute gradients with respect to design variables.

3.1. Design variables

The design variables $\mathbf{x}$ include the inertance $\mathbf{b}$ and damping coefficient $\mathbf{c}_r$ of inerters allocated at pre-assigned potential locations:

$$\mathbf{x} = [\mathbf{b}, \mathbf{c}_r] \tag{19}$$

The former defines the magnitude of supplemental mass, and the latter defines the energy dissipation capacity. The inertance typically determines the inerter's force transmission design, e.g., the lead of the ball

screw in ball screw systems or the transmission ratio of the gearbox in rack-and-pinion systems, as well as the dimensions of the flywheel. On the other hand, the damping coefficient depends on the principles underlying the damping mechanism. For instance, in eddy current systems, the damping coefficient is defined by the characteristics of the inductor and magnet interactions [42]. Thus, both inertance and damping coefficients are positive continuous variables. A numerically small value of b and c at a given location indicates that an inerter is not required. Furthermore, the design variables can also include selected parameters of other structural components, such as cross-sectional geometry, if required [35].

3.2. Objective function

The purpose of the optimization problem is to minimise the total cost of the CID devices. However, because inerters are not yet widely implemented in the industry, empirical cost data correlating price with mechanical properties such as inertance or associated damping coefficients are still limited or unavailable [43]. Therefore, the mean of peak force output, or force capacity, of the inerter under a selection of ground motions is chosen as a proxy for estimating cost [40]. This approach is both practical and theoretically sound, as the force capacity reflects the characteristics of all units of mechanical transmission that determine their size, complexity, and performance demand. An acceptable cost model in this context can be expressed as:

$$Q_0 = \frac{1}{M}\sum_{k=1}^{K}\sum_{m=1}^{M} a\left(\gamma f_{k,peak}^{(m)}\right)^q \tag{20}$$

where $f_{k,peak}^{(m)}$ represents the peak force capacity of the k^{th} device subjected to the m^{th} earthquake record, a is a unit linear cost coefficient, q is a scaling exponent capturing the nonlinear growth of cost with increased force demands, and γ is the load safety factor.

The parameters a and q can be calibrated by drawing analogies with more established devices such as viscous and friction dampers, for which extensive cost-performance studies exist [44,45]. This cost modelling approach provides a robust and adaptable framework for incorporating inerters into optimization-based seismic design, despite the absence of direct manufacturer data.

3.3. Constraints

The optimization is subject to a set of constraints, including both design space and performance requirements. The design space χ limits the feasible range of the design variables:

$$\mathbf{x} \in \chi, \tag{21}$$

In addition, performance constraints Q_1 ensure that the optimised design satisfies safety and functionality criteria. CIDs can be installed on various types of structures, and the constraints may vary depending on the scenario. Many stringent restrictions could be considered based on code limitations and designer decisions. The most commonly adopted EDP for seismic design includes the interstorey drift ratio [46]. Herein, the mean of the maximum peak interstorey drift along the height under the selected ground motions is adopted, which is defined as:

$$Q_1 = \frac{1}{M}\sum_{m=1}^{M} \max\left(\mathbf{d}_{peak}^{(m)}\right) - d_{allow} \le 0 \tag{22}$$

where $\mathbf{d}_{peak}^{(m)}$ is the peak interstorey drift distribution along the height subjected to the m^{th} earthquake record, and d_{allow} is the allowable magnitude. Additional constraints may address specific types of structures; for example, the self-centering ratio and tendon elongation for self-centering structures. By incorporating these requirements, the optimization process ensures the CID-equipped design meets both performance and safety requirements.

3.4. Summary of the optimization problem

Considering the design variables, objective function, and constraints described above, the optimization problem is formulated as:

$$\begin{aligned} &\min\ Q_0 = \frac{1}{M}\sum_{k=1}^{K}\sum_{m=1}^{M} a\left(\gamma f_{k,peak}^{(m)}\right)^q \\ &s.t.\ Q_1 = \frac{1}{M}\sum_{m=1}^{M} \max\left(\mathbf{d}_{peak}^{(m)}\right) - d_{allow} \le 0 \\ &\mathbf{x} \in \chi \end{aligned} \tag{23}$$

where $\mathbf{d}_{peak}^{(m)}$ and $f_{k,peak}^{(m)}$ are representative peak values of the system response histories $\mathbf{f}_n$ and $\mathbf{v}_n$. Furthermore, the system response histories are determined by the series of time step optimization problems (Eq. 14) of the MLF framework at each iteration of the structural optimization problem. Thus, Eq. (23) is a nested constrained optimization problem.

4. Adjoint sensitivity analysis

For a formal optimization problem, in which a single objective and constraint depend on a large number of design variables, gradient evaluation based on Direct Differentiation Methods (DDM) becomes computationally prohibitive. By contrast, the adjoint method provides a far more efficient alternative by exploiting the structure of the implicit time-step equilibrium equations, enabling the computation of gradients with respect to all design variables at a computational cost that is largely independent of the number of variables. In this section, a concise overview of the adjoint formulation is presented.

Section 4.1 outlines the general adjoint methodology and its application to response-history-based objectives. Section 4.2 reformulates the MLF time-step optimization into a set of Karush-Kuhn-Tucker (KKT) residual equations suitable for adjoint analysis. Section 4.3 then derives the adjoint equations and the resulting gradient expressions with respect to the CID design variables. Together, these developments provide an efficient means to compute sensitivities without explicitly differentiating the nonlinear response history.

4.1. Overview of the adjoint method

For systems with implicit and non-linear relationships between design variables and objectives or constraints, the optimization landscape is often non-convex, posing convergence and solution quality challenges. Non-gradient methods, such as genetic algorithms (GA) [47] or particle swarm optimization (PSO) [48], typically require a large number of function evaluations, leading to high computational costs [49]. To improve computational efficiency, gradient-based optimization methods are preferred, as they can substantially reduce the number of function evaluations required to reach an optimal solution. However, the gradient is often difficult to implement [50], especially in the context of nested nonlinear structural response analyses. To address this challenge, an adjoint sensitivity that efficiently computes gradients from the implicit definition of objectives and constraints with respect to design variables is adopted from [39].

For a structural optimization problem, the objective or constraint function $Q_i(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n, \mathbf{x})$ can be expressed as a function of the system response history $\mathbf{u}_n$ and design variables $\mathbf{x}$, along with its gradient:

$$\frac{dQ_i}{d\mathbf{x}} = \frac{\partial Q_i}{\partial \mathbf{x}} + \sum_{n=1}^{N} \frac{\partial Q_i}{\partial \mathbf{u}_n}\frac{d\mathbf{u}_n}{d\mathbf{x}} \tag{24}$$

where N is the total number of time steps, and $\frac{\partial Q_i}{\partial \mathbf{u}_n}$ is based on the definition of Q_i, while $\frac{d\mathbf{u}_n}{d\mathbf{x}}$ relies on the implicit time step equilibrium $R(\mathbf{u}_n, x) = 0$. This implicitness makes it difficult to obtain the sensitivity $\frac{d\mathbf{u}_n}{d\mathbf{x}}$ directly. Alternatively, it is easy to determine the derivatives of $R(\mathbf{u}_n, x) = 0$ and multiply an adjoint variable $\boldsymbol{\omega}_n$ as follows:

$$\boldsymbol{\omega}_n^\top \frac{\partial R(\mathbf{u}_n, \mathbf{x})}{\partial \mathbf{u}_n}\frac{d\mathbf{u}_n}{d\mathbf{x}} + \boldsymbol{\omega}_n^\top \frac{\partial R(\mathbf{u}_n, \mathbf{x})}{\partial \mathbf{x}} = \mathbf{0} \tag{25}$$

add above expression for $\forall n \in [1, N]$ to Eq. (24) to obtain an augmented gradient expression:

$$\frac{dQ_i}{d\mathbf{x}} = \frac{\partial Q_i}{\partial \mathbf{x}} + \sum_{n=1}^{N} \left[\frac{\partial Q_i}{\partial \mathbf{u}_n} \frac{d\mathbf{u}_n}{d\mathbf{x}} + \boldsymbol{\omega}_n^\top \frac{\partial R(\mathbf{u}_n, \mathbf{x})}{\partial \mathbf{u}_n} \frac{d\mathbf{u}_n}{d\mathbf{x}} + \boldsymbol{\omega}_n^\top \frac{\partial R(\mathbf{u}_n, \mathbf{x})}{\partial \mathbf{x}} \right] \tag{26}$$

The terms related to $\frac{d\mathbf{u}_n}{d\mathbf{x}}$ can be eliminated if the adjoint variables $\boldsymbol{\omega}_n$ are correctly chosen to satisfy the adjoint equation:

$$\boldsymbol{\omega}_n^\top \frac{\partial R(\mathbf{u}_n, \mathbf{x})}{\partial \mathbf{u}_n} = -\frac{\partial Q_i}{\partial \mathbf{u}_n} \tag{27}$$

by substituting Eq. (27) into Eq. (26), an alternative expression for the sensitivity is obtained:

$$\frac{dQ_i}{dx_j} = \frac{\partial Q_i}{\partial \mathbf{x}} + \sum_{n=1}^{N} \boldsymbol{\omega}_n^\top \frac{\partial R(\mathbf{u}_n, \mathbf{x})}{\partial \mathbf{x}} \tag{28}$$

Once the adjoint variable $\boldsymbol{\omega}_n$ is obtained from the adjoint equation (Eq. 27), it can be directly employed to evaluate the adjoint sensitivities with respect to all design variables xj. This eliminates the need to compute the direct derivative $\frac{d\mathbf{u}_n}{dx_j}$ for each x_j, which is both technically challenging and computationally expensive. By exploiting the gradient information already embedded in the time-step equilibrium equation $R = 0$, the adjoint method achieves a substantial reduction in computational cost, particularly in problems involving a large number of design variables.

In practice, response history analysis often employs multi-step time integration schemes, e.g., $\mathbf{R}(\mathbf{u}_{n+1}, \mathbf{u}_n, \mathbf{x})$. These schemes introduce additional complexity into the formulation. Nevertheless, the same adjoint philosophy applies and can be extended to such cases.

4.2. KKT conditions

The first step to solve the modified nested time step optimization problem presented in Section 3.4 is to convert the optimization-based time integration scheme 14, 15 and 16 into the form $\{\mathbf{R}\} = \{\mathbf{0}\}$, complying with the following Karush-Kuhn-Tucker (KKT) conditions of the residuals:

$$\begin{cases} \mathbf{R}_{f,n+1} = \mathbf{H}\mathbf{f}_{n+1} + \mathbf{A}_{eq}^\top \boldsymbol{\lambda}_{eq,n+1} + \mathbf{A}_{leq}^\top \boldsymbol{\lambda}_{leq,n+1} + \bar{\mathbf{H}}\mathbf{f}_n + \mathbf{A}_{leq}^\top \mathbf{s}\lambda_{leq,n} + \mathbf{y}_7 \bar{\mathbf{P}}_{1,n} \\ \qquad + \mathbf{y}_8 \mathbf{v}_{1,n} + \mathbf{y}_9 \mathbf{v}_{r,n} = \mathbf{0} \\ \mathbf{R}_{eq,n+1} = \mathbf{A}_{eq}\mathbf{f}_{n+1} - \mathbf{b}_{eq} = \mathbf{0} \\ \mathbf{R}_{leq,n+1} = \mathbf{A}_{leq}\mathbf{f}_{n+1} - \mathbf{b}_{leq} + \boldsymbol{\alpha}_{n+1}^2 = \mathbf{0} \\ \mathbf{R}_{\alpha,n+1} = \boldsymbol{\lambda}_{leq,n+1}^\top \boldsymbol{\alpha}_{n+1} = \mathbf{0} \\ \mathbf{R}_{v_1,n+1} = \mathbf{v}_{1,n+1} - \mathbf{y}_1 \bar{\mathbf{P}}_{1,n} - \mathbf{y}_2(\mathbf{f}_{n+1} + \mathbf{f}_n) - \mathbf{y}_3 \mathbf{v}_{1,n} = \mathbf{0} \\ \mathbf{R}_{v_r,n+1} = \mathbf{v}_{r,n+1} - \mathbf{y}_4 \mathbf{f}_{n+1} - \mathbf{y}_5 \mathbf{v}_{r,n} - \mathbf{y}_6 \mathbf{f}_n = \mathbf{0} \end{cases} \tag{29}$$

where $\boldsymbol{\alpha}_{n+1}^2$ are slack variable vectors that convert the inequality constraints to equality constraints. This set of equations represents an equivalence of the MLF time step optimization. The variables indexed by n or $n+1$ collectively characterise a unique system state corresponding to the current and subsequent time steps, just as the variables $\mathbf{u_n}$ and$\mathbf{u_{n+1}}$ in Section 4.1.

4.3. Adjoint sensitivity

The objective or constraint function of the performance-based optimization is dependent on all system state variables appearing in the K.K.T. condition as $Q_i\left(\mathbf{f}_n, \lambda_{eq,n}, \lambda_{leq,n}, \alpha_n, \mathbf{v}_{1,n}, \mathbf{v}_{r,n}, \mathbf{x}\right)$. According to Eq. (28), the adjoint sensitivity with respect to the design variables, $\mathbf{x}$, is given by:

$$\begin{aligned} \frac{dQ_i}{d\mathbf{x}} &= \frac{\partial Q_i}{\partial \mathbf{x}} + \sum_{n=1}^{N-1} \left(\boldsymbol{\omega}_{f,n+1}^\top \frac{\partial R_{f,n+1}}{\partial x} + \omega_{eq,n+1}^\top \frac{\partial R_{eq,n+1}}{\partial x} + \boldsymbol{\omega}_{leq,n+1}^\top \frac{\partial R_{leq,n+1}}{\partial x} \right. \\ &\quad \left. + \boldsymbol{\omega}_{\alpha,n+1}^\top \frac{\partial R_{\alpha,n+1}}{\partial x} + \boldsymbol{\omega}_{v_{1,n+1}}^\top \frac{\partial R_{v_{1,n+1}}}{\partial x} + \boldsymbol{\omega}_{v_{r,n+1}}^\top \frac{\partial R_{v_{r,n+1}}}{\partial x} \right) \\ &= \frac{\partial Q_i}{\partial \mathbf{x}} + \sum_{n=1}^{N-1} \left[\boldsymbol{\omega}_{f,n+1}^\top \left(\frac{\partial \bar{\mathbf{H}}}{\partial \mathbf{x}} \mathbf{f}_n + \frac{\partial \mathbf{A}_{eq}}{\partial \mathbf{x}} \lambda_{eq,n} - \frac{\partial \mathbf{A}_{leq}}{\partial \mathbf{x}} \mathbf{s}\lambda_{leq,n} + \frac{\partial \mathbf{y}_7}{\partial \mathbf{x}} \bar{\mathbf{P}}_{1,n} \right. \right. \\ &\quad \left. \left. + \frac{\partial \mathbf{y}_8}{\partial \mathbf{x}} \mathbf{v}_{1,n} + \frac{\partial \mathbf{y}_9}{\partial \mathbf{x}} \mathbf{v}_{r,n} \right) \right] \\ &\quad + \sum_{n=1}^{N-1} \left[\boldsymbol{\omega}_{v_{1,n+1}}^\top \left(-\frac{\partial \mathbf{y}_1}{\partial \mathbf{x}} \bar{\mathbf{P}}_{1,n} - \frac{\partial \mathbf{y}_2}{\partial \mathbf{x}} \mathbf{f}_n - \frac{\partial \mathbf{y}_3}{\partial \mathbf{x}} \mathbf{v}_{1,n} \right) \right. \\ &\quad \left. - \boldsymbol{\omega}_{v_{r,n+1}}^\top \left(\frac{\partial y_5}{\partial \mathbf{x}} \mathbf{v}_{r,n} + \frac{\partial y_6}{\partial \mathbf{x}} \mathbf{f}_n \right) \right] \\ &\quad + \sum_{n=2}^{N} \left[\boldsymbol{\omega}_{f,n}^\top \left(\frac{\partial \mathbf{H}}{\partial \mathbf{x}} \mathbf{f}_n + \frac{\partial \mathbf{A}_{eq}}{\partial \mathbf{x}} \lambda_{eq,n} + \frac{\partial \mathbf{A}_{leq}}{\partial \mathbf{x}} \lambda_{leq,n} \right) \right. \\ &\quad \left. - \boldsymbol{\omega}_{v_{1,n}}^\top \frac{\partial \mathbf{y}_2}{\partial \mathbf{x}} \mathbf{f}_n - \boldsymbol{\omega}_{v_{r,n}}^\top \frac{\partial \mathbf{y}_4}{\partial \mathbf{x}} \mathbf{f}_n \right] \\ &\quad + \sum_{n=2}^{N} \left[\boldsymbol{\omega}_{eq,n}^\top \left(\frac{\partial \mathbf{A}_{eq}}{\partial \mathbf{x}} \mathbf{f}_n - \frac{\partial \mathbf{b}_{eq}}{\partial \mathbf{x}} \right) + \boldsymbol{\omega}_{leq,n}^\top \left(\frac{\partial \mathbf{A}_{leq}}{\partial \mathbf{x}} \mathbf{f}_n - \frac{\partial \mathbf{b}_{leq}}{\partial \mathbf{x}} \right) \right] \end{aligned} \tag{30}$$

Analogous to the same logic of Eq. (27), the adjoint variables are determined by solving the adjoint equation below:

$$\begin{aligned} &\begin{bmatrix} \mathbf{H} & \mathbf{A}_{eq}^\top & \mathbf{A}_{leq}^\top & 0 & -y_2^\top & -y_4^\top \\ \mathbf{A}_{eq} & 0 & 0 & 0 & 0 & 0 \\ \mathbf{A}_{leq} & 0 & 0 & D(\boldsymbol{\alpha}_n) & 0 & 0 \\ 0 & 0 & 2D(\boldsymbol{\alpha}_n) & D(\lambda_{leq,n}) & 0 & 0 \\ 0 & 0 & 0 & 0 & \mathbf{I} & 0 \\ 0 & 0 & 0 & 0 & 0 & \mathbf{I} \end{bmatrix} \begin{Bmatrix} \boldsymbol{\omega}_{f,n} \\ \boldsymbol{\omega}_{eq,n} \\ \boldsymbol{\omega}_{leq,n} \\ \boldsymbol{\omega}_{\alpha,n} \\ \boldsymbol{\omega}_{v_{1,n}} \\ \boldsymbol{\omega}_{v_{r,n}} \end{Bmatrix} \\ &= \begin{Bmatrix} -\frac{\partial Q_i}{\partial \mathbf{f}_n} - \bar{\mathbf{H}}\boldsymbol{\omega}_{f,n+1} + \mathbf{y}_2^\top \boldsymbol{\omega}_{v_{1,n+1}} + \mathbf{y}_6^\top \boldsymbol{\omega}_{v_{r,n+1}} \\ -\frac{\partial Q_i}{\partial \boldsymbol{\lambda}_{eq,n}} \\ -\frac{\partial Q_i}{\partial \boldsymbol{\lambda}_{leq,n}} - \mathbf{s}^\top \mathbf{A}_{leq} \boldsymbol{\omega}_{f,n+1} \\ -\frac{\partial Q_i}{\partial \boldsymbol{\alpha}_n} \\ -\frac{\partial Q_i}{\partial \mathbf{v}_{1,n}} - \mathbf{y}_8^\top \boldsymbol{\omega}_{f,n+1} + \mathbf{y}_3^\top \boldsymbol{\omega}_{v_{1,n+1}} \\ -\frac{\partial Q_i}{\partial \mathbf{v}_{r,n}} - \mathbf{y}_9^\top \boldsymbol{\omega}_{f,n+1} + \mathbf{y}_5^\top \boldsymbol{\omega}_{v_{r,n+1}} \end{Bmatrix} \end{aligned} \tag{31}$$

This system of adjoint equations can be solved backward in time. It is worth noting that, although the system of equations may seem large, it is very sparse. Gaussian elimination can be applied to remove unnecessary unknown multipliers while retaining only $\boldsymbol{\omega}_{f,n}$. Thus, the adjoint equation can be significantly reduced to the same size as $\mathbf{H}$, which is the number of the numerical system.

In this context, constraints or objectives Q_i are only related to the structural response components $\mathbf{f}_n$ and $\mathbf{v}_{1,n}$; thus, $\frac{\partial Q_i}{\partial \lambda_{eq,n}}$, $\frac{\partial Q_i}{\partial \lambda_{leq,n}}$, $\frac{\partial Q_i}{\partial \boldsymbol{\alpha}_n}$, and $\frac{\partial Q_i}{\partial \mathbf{v}_{r,n}}$ are all zero. Before solving the adjoint Eq. (31), $\frac{\partial Q_i}{\partial \mathbf{f}_n}$ and $\frac{\partial Q_i}{\partial \mathbf{v}_{1,n}}$ need to be specified using the chain rule based on the definition of the objective and constraint functions from Section 3.4:

$$\frac{\partial Q_0}{\partial \mathbf{f}_n} = \frac{1}{M} \sum_{k=1}^{K} \sum_{m=1}^{M} \frac{\partial Q_0}{\partial f_{k,peak}^{(m)}} \frac{\partial f_{k,peak}^{(m)}}{\partial \mathbf{f}_n} \tag{32}$$

$$\frac{\partial Q_1}{\partial \mathbf{v}_{1,n}} = \frac{1}{M} \sum_{m=1}^{M} \frac{\partial Q_1}{\partial \mathbf{d}_{peak}^{(m)}} \frac{\partial \mathbf{d}_{peak}^{(m)}}{\partial \mathbf{v}_{1,n}} \tag{33}$$

where,

$$f_{k,peak}^{(m)} = \max_t \left(\mathbf{f}_k^{(m)}\right) \approx \left[\frac{1}{N}\sum_{n=1}^{N}\left(\mathbf{f}_k^{(m)}\right)^p\right]^{\frac{1}{p}}, \quad \mathbf{f}_k^{(m)} = \mathbf{T}_k\mathbf{f}_n^{(m)} \tag{34}$$

$$\mathbf{d}_{peak}^{(m)} = \max_t \left(\mathbf{d}_n^{(m)}\right) \approx \left[\frac{1}{N}\sum_{n=1}^{N}\left(\mathbf{d}_n^{(m)}\right)^p\right]^{\frac{1}{p}}, \quad \mathbf{d}_n^{(m)} = \mathbf{T}_d\sum_{n=1}^{N}\frac{h}{2}\left(\mathbf{v}_{1,n-1}^{(m)} + \mathbf{v}_{1,n}^{(m)}\right) \tag{35}$$

where $\mathbf{T}_d$ and $\mathbf{T}_k$ are logical matrices that convert state response histories $\mathbf{f}_n$ and $\mathbf{v}_n$ to the force history of k^{th} inerter elements $\mathbf{f}_k$ and interstorey drift distribution history $\mathbf{d}_n$, respectively. N is the total number of time steps, and h is the step size. m denotes the m^{th} earthquake record. To make the terms of the max() function differentiable, the L-p norm with a large p value is adopted to approximate the peak value among the time series.

Since the design variables only involve b and c_r, the sensitivity can be simplified as:

$$\frac{d\hat{Q}_i}{db} = \sum_{n=2}^{N}\left[\boldsymbol{\omega}_{f,n}^{\top}\frac{\partial\mathbf{H}}{\partial b}\mathbf{f}_n - \boldsymbol{\omega}_{v_{r,n+1}}^{\top}\frac{\partial\mathbf{y}_4}{\partial b}\mathbf{f}_n\right] + \sum_{n=1}^{N-1}\left[\boldsymbol{\omega}_{f,n+1}^{\top}\frac{\partial\mathbf{y}_8}{\partial b}\mathbf{v}_{r,n} - \boldsymbol{\omega}_{v_{r,n+1}}^{\top}\frac{\partial\mathbf{y}_5}{\partial b}\mathbf{v}_{r,n}\right] \tag{36}$$

$$\frac{d\hat{Q}_i}{dc_r} = \sum_{n=1}^{N-1}\left[\boldsymbol{\omega}_{f,n+1}^{\top}\frac{\partial\mathbf{y}_8}{\partial c_r}\mathbf{v}_{r,n} - \boldsymbol{\omega}_{v_{r,n+1}}^{\top}\frac{\partial\mathbf{y}_5}{\partial c_r}\mathbf{v}_{r,n}\right] \tag{37}$$

During each iteration of the structural optimization, once the NRHA is complete, adjoint variables $\boldsymbol{\omega}_n$ can be computed through adjoint equations (Eq. 31) using the response history data. With the adjoint variables on hand, the adjoint sensitivity can be easily obtained from Eq. (36).

5. Case studies

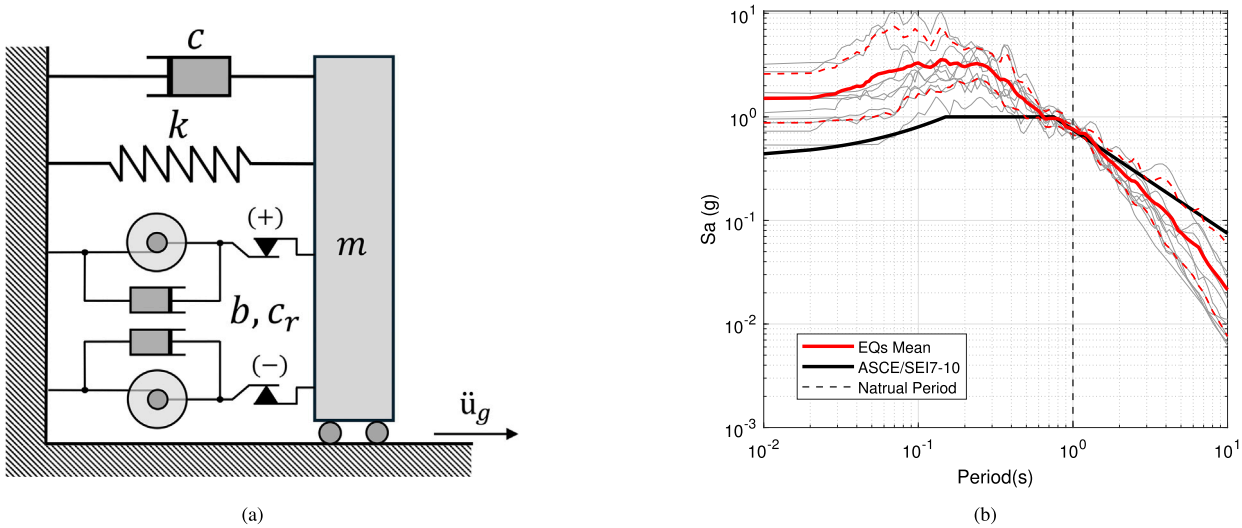
This section presents two numerical case studies to demonstrate the effectiveness of the proposed optimization framework and adjoint sensitivity formulation presented above. The first example examines a SDOF oscillator equipped with a pair of CID, providing a clear and interpretable setting for validating gradient computations and illustrating fundamental CID behaviours. The second example extends the framework to a nonlinear nine-storey benchmark building to evaluate its performance in a high-dimensional design space and under realistic structural modelling complexities. Together, these studies verify the accuracy of the adjoint gradient, illustrate the optimization landscape shaped by CID parameters, and highlight the computational advantages of the proposed method.

5.1. Optimization of a SDOF system

This example considers a SDOF oscillator subjected to a set of seismic excitations. The proposed SDOF system has a natural period of $T_n = 1\ s$ and a viscous damping ratio of $\zeta = 0.02$, with two CIDs working in opposing directions. The system is subjected to ten real earthquake records sourced from the Pacific Earthquake Engineering Research Center (PEER) database, whose response spectra are shown in Fig. 2(b). Although the addition of CIDs introduces behavioural nonlinearities, it does not significantly alter the SDOF's natural period [18]. Therefore, the ground motions are scaled to ensure consistent spectral amplitudes at $T = 1$ second. The SDOF system can represent structures that are dominated by their first mode responses, such as base-isolated structures. This simple two-dimensional optimization problem is particularly useful for validating the gradient computation method proposed above, as its low dimensionality allows for easy visualisation of results while also highlighting key characteristics of CIDs.

The inertance, b, and associated damping coefficient, c_r, of the inerters are set as the design variables, both constrained within a design space of (0,100). The objective function Q_0 is simplistically defined as the average peak clutch forces, with $a = 1$, $q = 1$, and $\gamma = 1$ in Eq. (23). The performance constraint Q_1 ensures that the average peak displacement does not exceed 0.14 m. To make the objective differentiable, the peak values are approximated by the L-p norm with a p value equal to 1e6, as stated in Eqs. (34) and (35).

The optimization problems are solved by the 'fmincon' function from MATLAB's optimization Toolbox. In each structural optimization iteration, the 10 response histories are solved using the MLF framework, which is implemented via the 'quadprog' function, and the sensitivity is calculated according to Eq. (36) based on the response history data. The sensitivity is then used by 'fmincon' to determine the next step of the structural optimization. By contrast, if the adjoint sensitivity is not provided, 'fmincon' will automatically run extra response history analyses with small perturbations on each design variable and calculate a numerical gradient based on the finite difference method. In this case study, both methods are conducted for comparison purposes.



(a) (b)

Fig. 2. (a) SDOF system equipped with twin CIDs ($m = 100\ kg, c = 8\pi\ kg/s, k = 400\pi^2\ N/m$), where (+) and (-) indicate the engagement direction of the clutch. (b) Response spectra of the ten earthquake records selected from the PEER database (grey), their associated mean (red) and mean plus/minus one standard deviation (dashed red), together with the ASCE/SEI-7–10 design spectrum (black). The vertical dashed line marks the natural period of the SDOF system ($T = 1$ s).

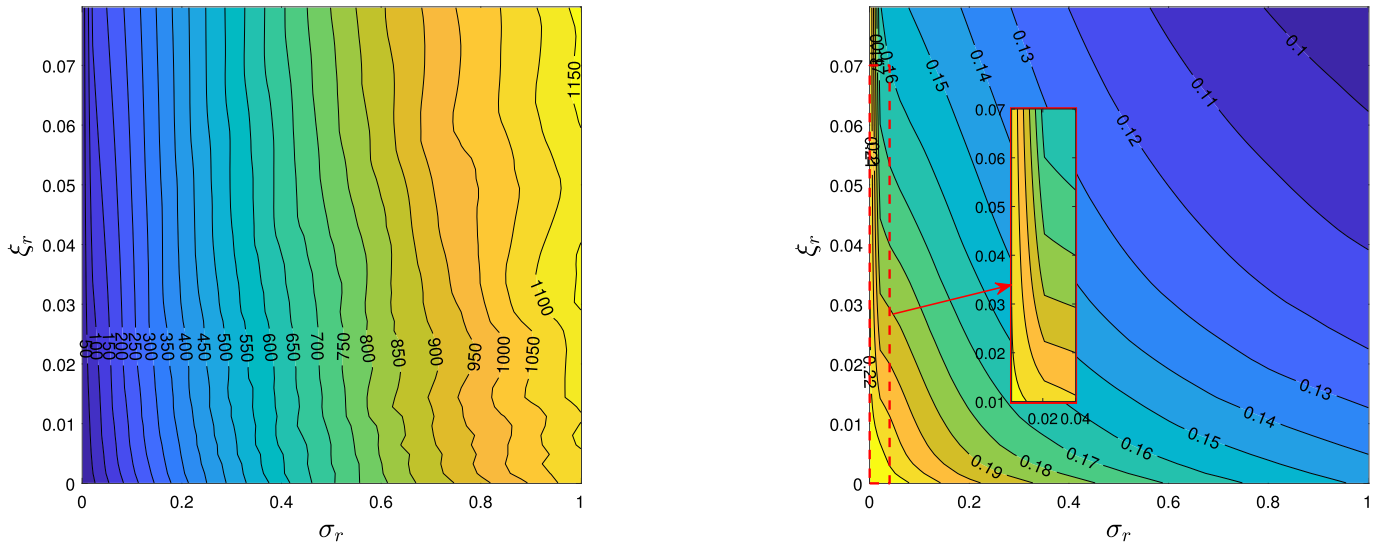


Fig. 3. (a) Mean peak clutch force over the design space defined by the non-dimensional inertance $\sigma_r = b/m$ and damping ratio $\xi_r = c_r/(2\sqrt{km})$. The contours show that clutch forces increase mainly with σ_r, while the influence of ξ_r becomes more noticeable at higher inertance levels. (b) Mean peak displacement under the ten ground motions over the same design space, including a close-up view of the low-inertance region. Displacement decreases with both σ_r and ξ_r, but exhibits a sharp transition when σ_r is very small, reflecting the near-ineffectiveness of the CID when its inertance approaches zero.

The distribution of the objective function and constraint over the design space is illustrated in Fig. 3. These help to understand the optimization landscape. The non-dimensional inertance (suppression coefficient) $\sigma_r = \frac{b}{m}$ and associated damping ratio $\zeta_r = \frac{c_r}{2\sqrt{km}}$ are adopted in this figure. Fig. 3(a) shows that the peak force response distribution is generally dominated by σ_r, especially when σ_r is low. As σ_r increases, the ζ_r contributes more effectively and the landscape becomes rugged and non-convex. The constraint (displacement) distribution is also shown in Fig. 3(b), where the displacement decreases with increasing σ_r and ζ_r in a hyperbolic pattern. The displacement distribution has a smoother terrain, except for the cliff-like region at low values of σ_r. This is explained by the fact that when the inertance is negligible, the device becomes a clutched viscous damper, which outputs zero force whether it is engaged or not. The optimal point is expected to be where the two contours are tangent to each other or on the design boundary.

Fig. 4 illustrates the optimization path, objective function magnitude, and constraint violation values throughout the optimization process, respectively. The results show that the optimization paths guided by the adjoint gradient conform to each other and converge to the same optimal point $[\sigma_r, \zeta_r] = [0.18, 0.08]$. This is consistent with the distribution of the objective function and constraint violation in Fig. 3.

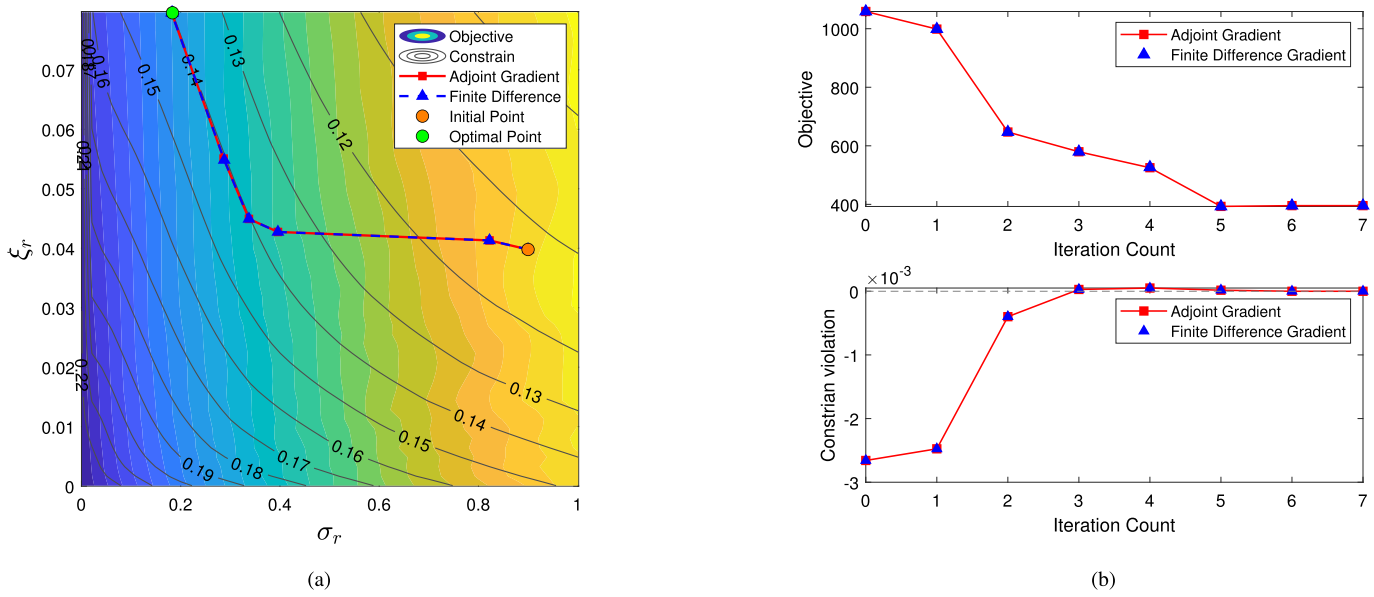


Fig. 4. (a) Optimization path in the design space defined by the non-dimensional inertance σ_r and damping ratio ξ_r, superimposed on the objective and constraint contours. The trajectories obtained using the adjoint gradient and the finite difference gradient closely coincide and converge to the same optimal point. (b) Evolution of the objective value and constraint violation over the optimization iterations. Both gradient strategies exhibit similar convergence behaviour, confirming the accuracy of the adjoint sensitivity and its consistency with the finite difference gradient.

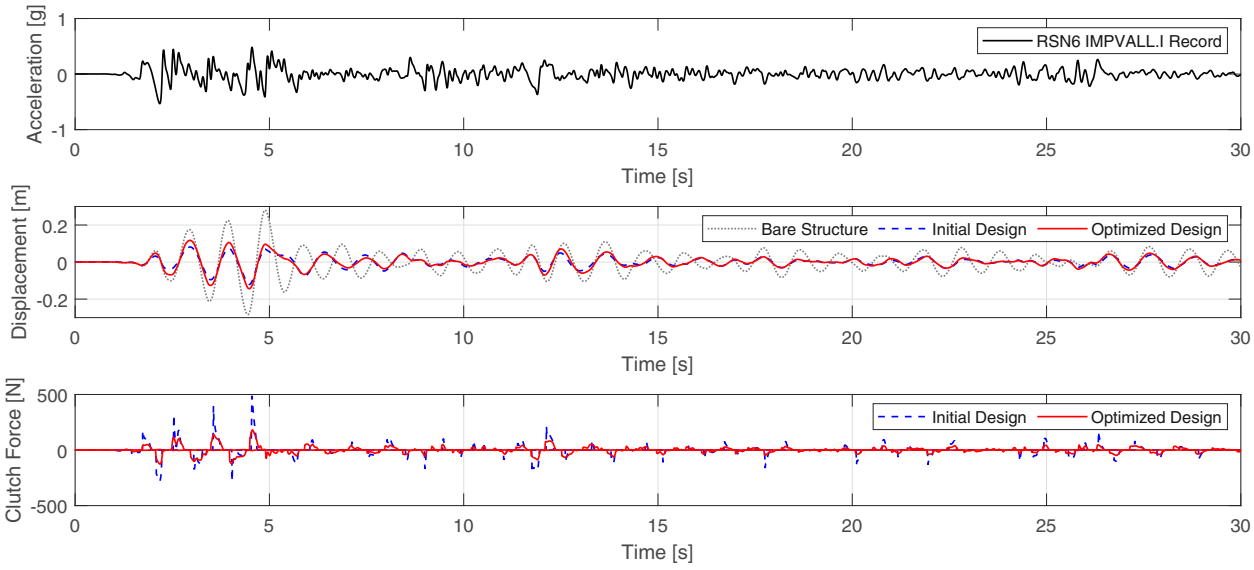


Fig. 5. Response histories of different versions of the SDOF system under the RSN6 IMPVALL.I record. The response of the bare structure, an initial CID design, and the optimised design are compared. The optimised configuration reduces clutch forces while maintaining acceptable displacement levels.

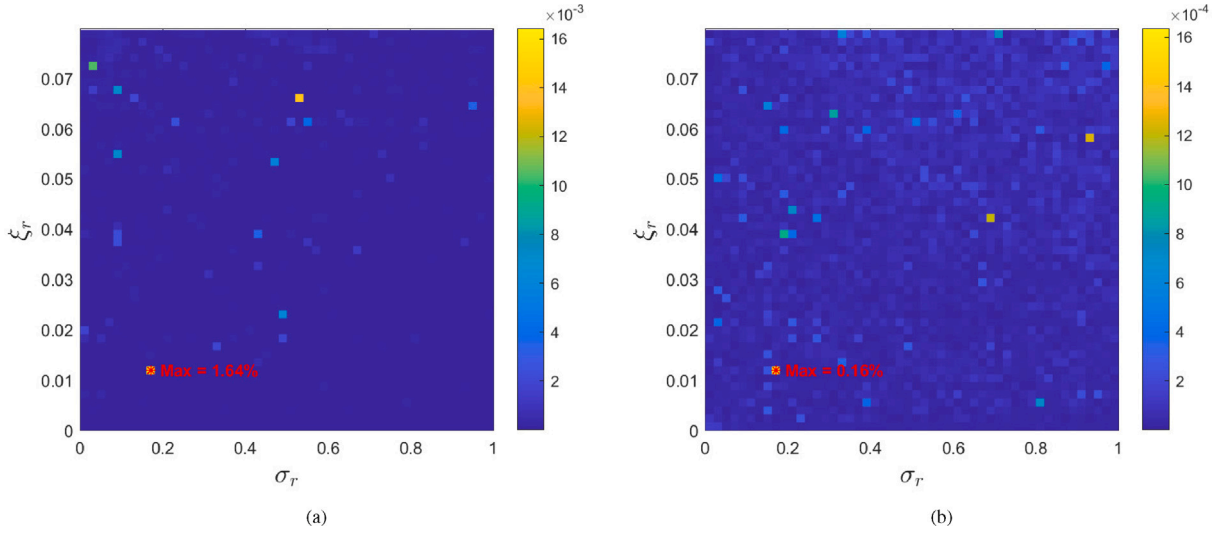


Fig. 6. Relative difference between the adjoint gradient and the finite difference gradient across the design space for (a) peak force response and (b) peak displacement response. The differences remain very small, confirming the accuracy of the adjoint sensitivity.

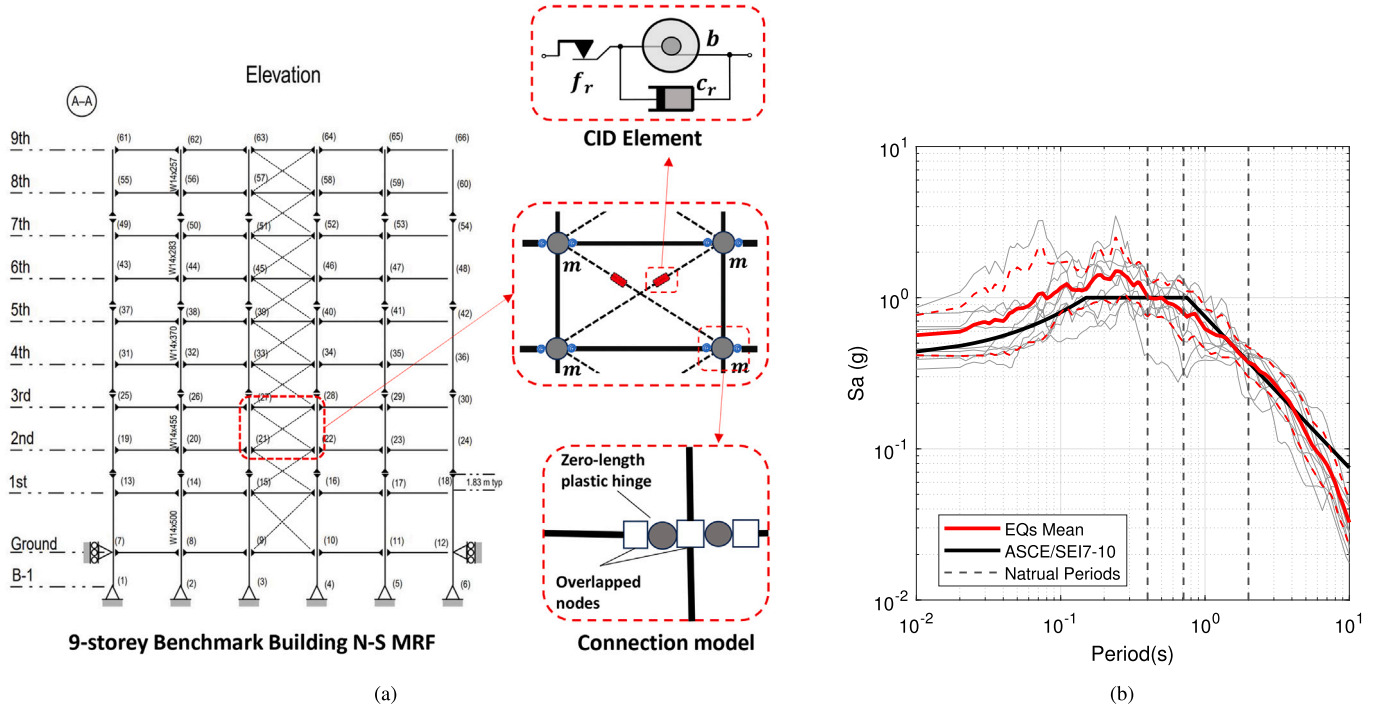
Fig. 5 shows the response histories of the SDOF system with different CID designs subjected to the RSN6 IMPVALL.I record. The response of the non-optimised system, the optimised system, and a system without CIDs are compared. The results show that the optimised system has significantly reduced clutch forces with acceptable interstorey drifts. This demonstrates the effectiveness of the proposed optimization method in controlling structural responses through the optimal design of CIDs.

To further verify the proposed adjoint gradient, Fig. 6(a) and (b) illustrate the distribution of relative differences between the adjoint gradient and the finite difference gradient in the entire design space. The finite difference gradient is computed using the forward finite difference method with an increment of 1×10^{-6}. The error distribution has no clear pattern and remains small, with maximum relative differences of 1.64% and 0.16% in force and displacement, respectively. These errors may arise from computational inaccuracies in the adjoint method, truncation errors in the finite difference approximation, and the rugged nature of the design space.

5.2. Optimization of a 9-storey benchmark building

To demonstrate the robustness of the proposed approach on more complex nonlinear structural models. The nine-storey benchmark building with a moment-resisting frame (MRF), proposed by Ohtori et al. [51], is evaluated herein (see Fig. 7). The 2-D frame is modelled using beam elements with lumped plasticity in the MLF framework [40]. The P-Delta effect is considered using the fictitious leaning column method [52]. 0.5%, 1% and 1% damping ratios are adopted for the first three modes, respectively. CIDs are deployed at each storey. The structure is subjected to a set of real earthquake records from the PEER NGA west 2 database, with response spectra shown in Fig. 7(b).

As stated in Section 3.4, the optimization aims to minimise the sum of mean peak forces in all CIDS while keeping the mean maximum relative inter-storey drifts below 2%, with $a = 1$, $q = 1$, and $\gamma = 1$. The peak values are approximated by the L-p norm among the time histories with a p value equal to 1E6, as stated in Eqs. (34) and (35). The inertance, b, and the associated damping coefficient, c_r, for each CID are limited to

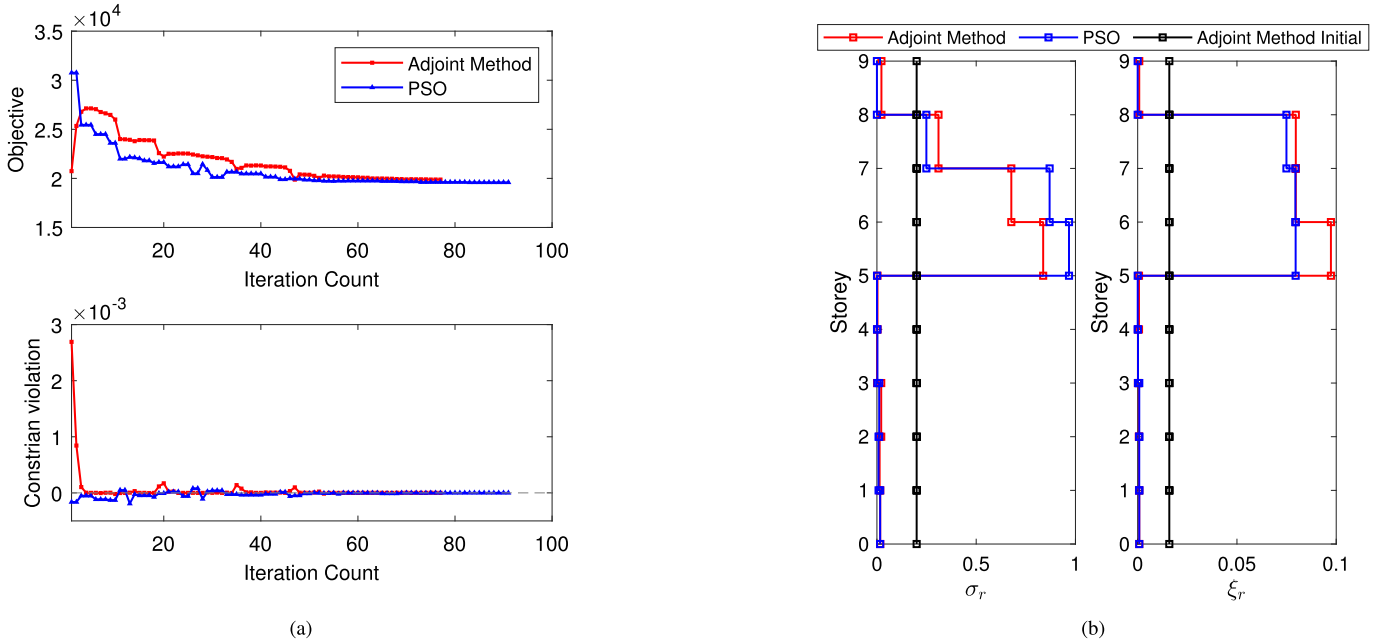


(a) (b)

Fig. 7. (a) Configuration of the 9-storey benchmark steel moment-resisting frame equipped with CIDs, including modelling details of the CID element and beam–column connections with lumped plasticity. (b) Response spectra of the ten selected input ground motions, shown together with their mean spectrum and the ASCE/SEI-7–10 design spectrum, along with vertical lines indicating the first three natural periods of the building.

5000 tonnes and 2500 tonnes · s, respectively. These values correspond to the mass of the building and a 1% equivalent critical damping ratio in the inerter.

In this example, the particle swarm optimization (PSO) algorithm [53] is also implemented for comparison purposes. To this end, the same convergence criteria are applied for the two methods, and the swarm size for PSO is set to 50. Due to the stochastic nature of PSO, multiple independent runs were carried out to obtain statistically reliable results. For the gradient-based method, multiple runs with different initial starting points were performed to avoid being trapped in local minima, which are inherent in non-convex optimization problems. Fig. 8(a) illustrates the history of the objective value and constraint violation throughout the optimization process of both methods. Fig. 8(b) shows the initial design of the adjoint method and



(a) (b)

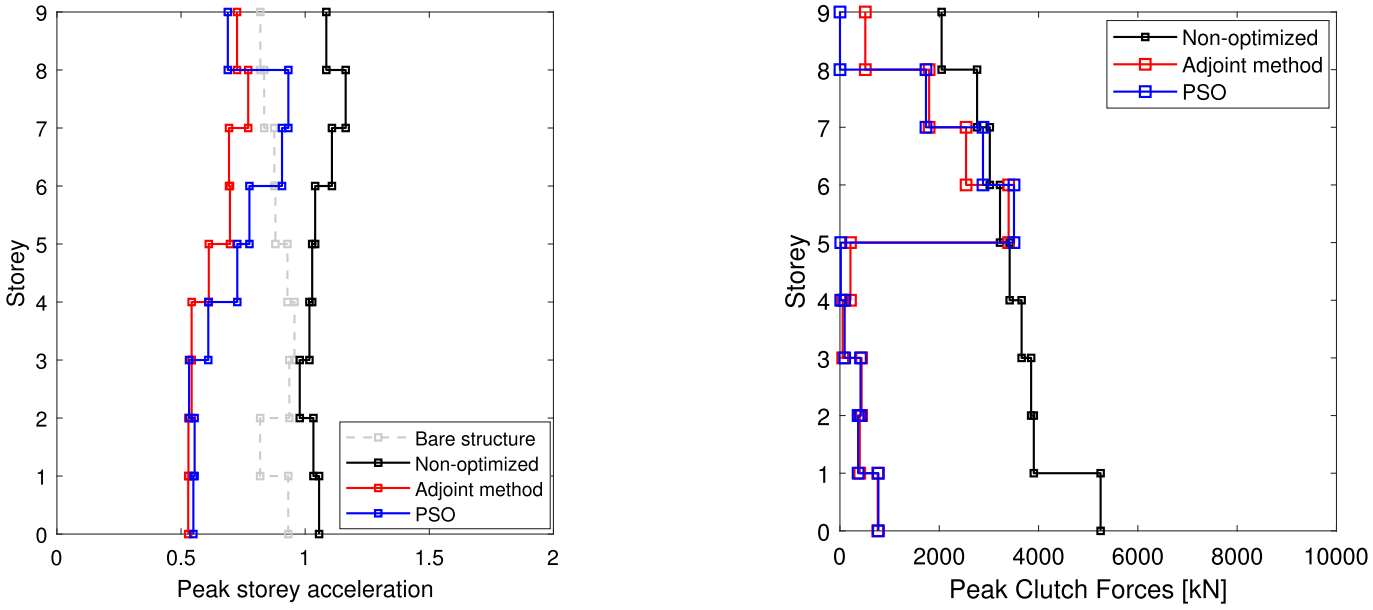
Fig. 8. Comparison of Adjoint Method and PSO for the optimization of the 9-storey building. (a) The optimization history of the objective function and the constraints violation of the two methods, and (b) the initial and optimised distribution of inertance and damping obtained by the two methods, where inertance b and damping coefficient c_r are represented by their non-dimensional values $\sigma_r = \frac{b}{m}$ and $\xi_r = \frac{c_r}{2m\omega}$ with $m = 5000$ tonnes (floor mass of the building), with $\omega = \pi$ rad/s.

Table 1
Comparison between PSO and adjoint method in terms of computational efficiency.

Method	Iterations	Number of RHAs	Gradient Calculations	Time Used (s)	Optimal Objective
Adjoint	78	94	188	5456	19,879
PSO	92	4600	0	153,441	19,595

the optimal distribution of the suppression coefficient and damping ratio obtained by both methods. It can be observed that significant inertance values are assigned primarily to the 6th, 7th, and 8th stories, while the other stories are assigned nearly zero inertance. A similar trend is reflected in the damping ratios. This observation is consistent with the SDOF example, demonstrating that when the inertance approaches zero, the associated damping has less influence on the structural response. PSO, being a global optimization algorithm, explores the design space more comprehensively but at a significantly higher computational cost. The data in Table 1 clearly demonstrate the computational advantage of the adjoint method. Although both approaches achieve similar optimal objective values, PSO requires 4600 response-history analyses compared with only 94 for the adjoint method, resulting in a total runtime nearly thirty times larger. This efficiency gain stems from the ability of the adjoint formulation to compute gradients for all design variables at once, whereas PSO must evaluate many candidate solutions without gradient information. Consequently, the adjoint method offers a far more practical optimization strategy for large-scale nonlinear structural systems.

Fig. 9(a) and (b) illustrate the storey-wise distribution of mean peak inter-storey accelerations and peak device forces for the non-optimal



(a) (b)

Fig. 9. (a) Mean peak floor accelerations for the bare structure, the non-optimised CID configuration, and the two optimised designs (adjoint method and PSO), showing the significant reduction achieved by optimization. (b) Mean peak CID clutch forces along the building height for the same configurations, illustrating how optimization greatly reduces device force demands, particularly in lower storeys.

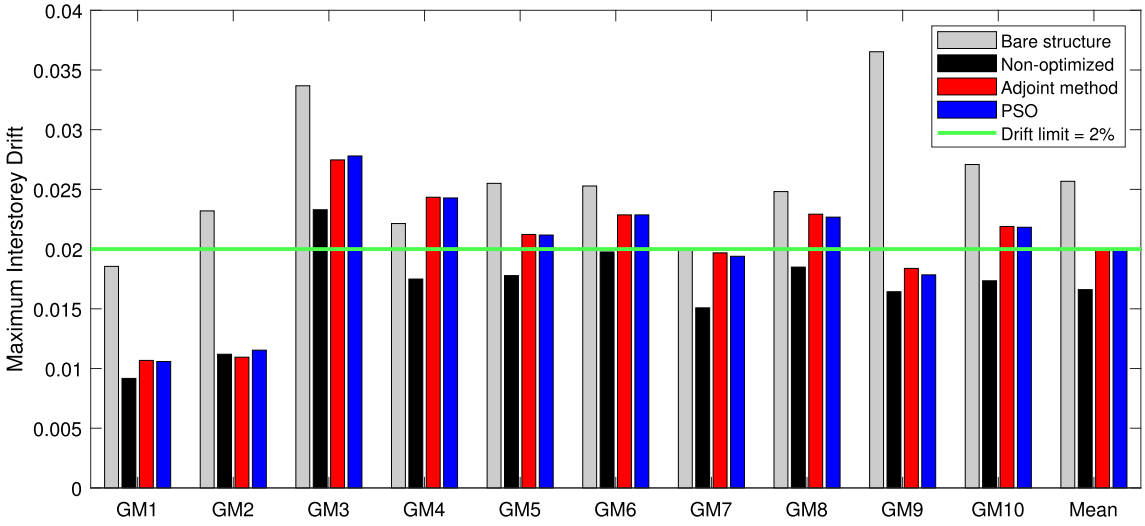


Fig. 10. Maximum interstorey drift ratios for each ground motion record and the mean value, comparing the bare structure, the non-optimised CID configuration, and the two optimised designs (adjoint method and PSO). The optimised systems keep all drift demands below the 2% limit indicated by the green line.

configuration with all design variables at their maximum values, the optimal configuration, and the bare structure, respectively. Fig. 10 shows the maximum inter-storey drift under each load case, as well as the mean value. These results indicate that the optimised system significantly reduces device peak forces and considerably suppresses peak floor accelerations while keeping inter-storey drifts within the prescribed performance limits. These findings highlight the effectiveness of optimally designed CIDs in mitigating structural responses and are consistent with previous studies [16,24,28].

6. Conclusion

This study presents an adjoint sensitivity-based optimization framework tailored for Clutched Inerter Devices (CIDs), using the Mixed Lagrangian Formalism (MLF) as the time integration scheme. The MLF scheme offers superior stability and efficiency in simulating non-smooth systems, while the adjoint method provides analytical gradients that significantly accelerate gradient-based optimization. In addition, this paper establishes a performance-based optimization framework for CIDs, where objectives and constraints are constructed from both economic considerations (e.g., device cost) and structural performance metrics (e.g., inter-story drift). This framework supports practical design decisions by linking engineering demands with device parameters in a systematic and computationally tractable manner. Two case studies serve to demonstrate the applicability and effectiveness of the proposed framework. In both examples, optimised CID configurations significantly suppress the system responses while minimising device costs that mainly depend on the device force.

Beyond the specific application to CIDs, the proposed optimization framework provides a broadly applicable and computationally efficient approach for optimising nonlinear structural control devices. Since the MLF-based time integration is formulated based on the standard form of MATLAB's `quadprog` function, the analytical sensitivity expressions derived in this study can be directly applied to any device whose dynamic behaviour can be discretised and solved in a similar quadratic programming structure, without requiring reformulation of the adjoint equations. Moreover, the adjoint method is not intrinsically bound to MLF. It can be coupled with conventional solvers, as long as the time integration can be expressed in the residual form $\boldsymbol{R} = \boldsymbol{0}$. This modularity grants the proposed methodology high flexibility and extensibility across a range of numerical schemes and optimization scenarios.

On the other hand, global optimization methods generally require a significantly larger computational effort, and this limitation becomes increasingly pronounced as the problem size grows. In large-scale applications, the computational burden associated with repeated nonlinear response-history analyses may render such methods impractical for routine engineering design. The objective of the proposed approach is therefore not to guarantee global optimality but rather to efficiently guide the search toward high-quality, practically feasible designs within a reasonable computational budget. Although the obtained solutions may not be globally optimal, our experience indicates that their performance is close to the optimum while being achieved with substantially reduced computational costs. Moreover, as problem dimensionality increases, global optimization methods often experience convergence difficulties and may fail to produce meaningful designs, further highlighting the advantage of the proposed strategy for large and computationally demanding optimization problems.

Despite its advantages, the proposed method also has certain limitations. For highly non-convex design landscapes, gradient-based solvers may become trapped in local minima, in which case global optimization techniques, such as genetic algorithms or particle swarm optimization, are generally preferable. However, as illustrated in Fig. 4(a), at least for the representative example considered in this study, the design landscape is convex and smooth. This observation increases confidence that the problem formulation is well suited for gradient-based optimization and that the obtained solutions are reliable within the scope of the present work. Furthermore, while adjoint-based sensitivity analysis is highly efficient for problems involving a large number of design variables, its computational advantage diminishes when many objectives or constraints depend on only a limited set of parameters. In such situations, the use of aggregated objective or constraint functions is recommended.

Future improvements to the proposed framework may be pursued at multiple levels. At the problem formulation level, the optimization landscape may be modified through additional smoothing or regularisation; for example, by adopting probabilistic descriptions of objective and constraint functions, which can mitigate local irregularities and reduce sensitivity to non-smooth response features. At the algorithmic level, more advanced optimization strategies may be incorporated, including sophisticated gradient-based solvers as well as adaptive mechanisms involving dynamically adjusted hyperparameters, step sizes, or move limits. Hybrid optimization schemes that combine global exploration with local refinement also represent a promising direction. In addition, robustness may be further enhanced through improved initialisation strategies, such as multi-start approaches, coarse-to-fine optimization, or warm-starting based on simplified models.

CRediT authorship contribution statement

Yixuan Zhang: Writing – original draft, Validation, Software, Investigation, Formal analysis, Conceptualization. **Oren Lavan:** Writing – review & editing, Supervision, Methodology, Conceptualization. **Christian Málaga-Chuquitaype:** Writing – review & editing, Supervision, Resources, Methodology, Conceptualization.

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] Hwang J-S, Kim J, Kim Y-M. Rotational inertia dampers with toggle bracing for vibration control of a building structure. Eng Struct 2007;29(6):1201–8. https://doi.org/10.1016/j.engstruct.2006.08.005
[2] Asai T, Araki Y, Ikago K. Energy harvesting potential of tuned inertial mass electromagnetic transducers. Mech Syst Signal Process 2017;84:659–72. https://doi.org/10.1016/j.ymssp.2016.07.048
[3] Shi X, Zhu S. Dynamic characteristics of stay cables with inerter dampers. J Sound Vib 2018;423:287–305. https://doi.org/10.1016/j.jsv.2018.02.042
[4] Saitoh M. On the performance of gyro-mass devices for displacement mitigation in base isolation systems. Struct Control Health Monit 2012;19(2):246–59. https://doi.org/10.1002/stc.419
[5] Lu L, Duan Y-F, Spencer Jr BF, Lu X, Zhou Y. Inertial mass damper for mitigating cable vibration. Struct Control Health Monit 2017;24(10):e1986. https://doi.org/10.1002/stc.1986
[6] Ma R, Bi K, Hao H. A novel rotational inertia damper for heave motion suppression of semisubmersible platform in the shallow sea. Struct Control Health Monit 2019;26(7):e2368. https://doi.org/10.1002/stc.2368
[7] Lazar IF, Neild SA, Wagg DJ. Using an inerter-based device for structural vibration suppression. Earthq Eng Struct Dyn 2014;43(8):1129–47. https://doi.org/10.1002/eqe.2390
[8] Chen MZQ, Hu Y, Huang L, Chen G. Influence of inerter on natural frequencies of vibration systems. J Sound Vib 2014;333(7):1874–87. https://doi.org/10.1016/j.jsv.2013.11.025
[9] Zhang R, Zhao Z, Pan C, Ikago K, Xue S. Damping enhancement principle of inerter system. Struct Control Health Monit 2020;27(5):e2523. https://doi.org/10.1002/stc.2523
[10] Marian L, Giaralis A. Optimal design of a novel tuned mass-damper–inerter (tmdi) passive vibration control configuration for stochastically support-excited structural systems. Probabilist Eng Mech 2014;38:156–64. https://doi.org/10.1016/j.probengmech.2014.03.007
[11] Petrini F, Giaralis A, Wang Z. Optimal tuned mass-damper-inerter (tmdi) design in wind-excited tall buildings for occupants' comfort serviceability performance and energy harvesting. Eng Struct 2020;204:109904. https://doi.org/10.1016/j.engstruct.2019.109904

[12] De Domenico D, Ricciardi G. An enhanced base isolation system equipped with optimal tuned mass damper inerter (tmdi). Earthq Eng Struct Dyn 2018;47(5):1169–92. https://doi.org/10.1002/eqe.3011
[13] Pan X, Málaga-Chuquitaype C. Seismic control of rocking structures via external resonators. Earthq Eng Struct Dyn 2020;49(12):1180–96. https://doi.org/10.1002/eqe.3284
[14] Ruiz R, Taflanidis AA, Giaralis A, Lopez-Garcia D. Risk-informed optimization of the tuned mass-damper-inerter (tmdi) for the seismic protection of multi-storey building structures. Eng Struct 2018;177:836–50. https://doi.org/10.1016/j.engstruct.2018.08.074
[15] Makris N, Kampas G. Seismic protection of structures with supplemental rotational inertia. J Eng Mech 2016;142(11):04016089. https://doi.org/10.1061/(ASCE)EM.1943-7889.0001152
[16] Zhang Y, Thiers-Moggia R, Málaga-Chuquitaype C. Impact and clutch nonlinearities in the seismic response of inerto-rocking systems. Bull Earthq Eng 2023;21(3):1713–45. https://doi.org/10.1007/s10518-022-01355-6
[17] Málaga-Chuquitaype C, Menendez-Vicente C, Thiers-Moggia R. Experimental and numerical assessment of the seismic response of steel structures with clutched inerters. Soil Dyn Earthq Eng 2019;121:200–11. https://doi.org/10.1016/j.soildyn.2019.03.016
[18] Liang Q, Li L. Theoretical and experimental analyses of structures with supplemental clutching inertia devices. Earthq Eng Struct Dyn 2023;52(3):609–23. https://doi.org/10.1002/eqe.3776
[19] Wang J, Zheng Y. Experimental validation, analytical investigation, and design method of tuned mass damper-clutching inerter for robust control of structures. Earthq Eng Struct Dyn 2023;52(2):500–23. https://doi.org/10.1002/eqe.3770
[20] Li R, Fan K, Ma X, Wen T, Liu Q, Gao X, Zhu J, Zhang Y. A rotational energy harvester with a semi-flexible one-way clutch for capturing low-frequency vibration energy. Energy 2023;281:128266. https://doi.org/10.1016/j.energy.2023.128266
[21] Thiers-Moggia R, Málaga-Chuquitaype C. Performance-based seismic design and assessment of rocking timber buildings equipped with inerters. Eng Struct 2021;248:113164. https://doi.org/10.1016/j.engstruct.2021.113164
[22] Garrido H, Curadelli O, Ambrosini D. Resettable-inertance inerter: a semi-active control device for energy absorption. Struct Control Health Monit 2019;26(11):e2415. https://doi.org/10.1002/stc.2415
[23] Zhang Y, Málaga-Chuquitaype C, Lavan O. Mixed lagrangian formulation for modeling structures with clutched inerter devices. Earthq Eng Struct Dyn 2024;53(14):4203–22. https://doi.org/10.1002/eqe.4211
[24] Jangid RS. Seismic performance assessment of clutching inerter damper for isolated bridges. Pract Period Struct Des Constr 2022;27(2):04021078. https://doi.org/10.1061/(ASCE)SC.1943-5576.0000661
[25] Thiers-Moggia R, Málaga-Chuquitaype C. Seismic protection of rocking structures with inerters. Earthq Eng Struct Dyn 2019;48(5):528–47. https://doi.org/10.1002/eqe.3147
[26] Thiers-Moggia R, Málaga-Chuquitaype C. Effect of base-level inerters on the higher mode response of uplifting structures. J Eng Mech 2021;147(8):04021041. https://doi.org/10.1061/(ASCE)EM.1943-7889.0001935
[27] Wang M, Sun F. Displacement reduction effect and simplified evaluation method for sdof systems using a clutching inerter damper. Earthq Eng Struct Dyn 2018;47(7):1651–72. https://doi.org/10.1002/eqe.3034
[28] Liang Q, Li L. Optimal design for a novel inerter-based clutching tuned mass damper system. J Vib Control 2020;26(21–22):2050–9. https://doi.org/10.1177/1077546320910532
[29] Sivaselvan MV, Reinhorn AM. Lagrangian approach to structural collapse simulation. J Eng Mech 2006;132(8):795–805. https://doi.org/10.1061/(ASCE)0733-9399(2006)132:8(795)
[30] Lavan O, Sivaselvan MV, Reinhorn AM, Dargush GF. Progressive collapse analysis through strength degradation and fracture in the mixed lagrangian formulation. Earthq Eng Struct Dyn 2009;38(13):1483–504. https://doi.org/10.1002/eqe.914
[31] Lavan O. Dynamic analysis of gap closing and contact in the mixed lagrangian framework: toward progressive collapse prediction. J Eng Mech 2010;136(8):979–86. https://doi.org/10.1061/(ASCE)EM.1943-7889.0000146
[32] Apostolakis G, Dargush GF. Mixed lagrangian formulation for linear thermoelastic response of structures. J Eng Mech 2012;138(5):508–18. https://doi.org/10.1061/(ASCE)EM.1943-7889.0000346
[33] Apostolakis G, Dargush GF. Mixed variational principles for dynamic response of thermoelastic and poroelastic continua. Int J Solids Struct 2013;50(5):642–50. https://doi.org/10.1016/j.ijsolstr.2012.10.021
[34] Giles MB, Pierce NA. An introduction to the adjoint approach to design. Flow Turbulence Combust 2000;65(3):393–415. https://doi.org/10.1023/A:1011430410075
[35] Christensen PW, Klarbring A. An introduction to structural optimization, vol. 153. Springer Science & Business Media; 2008.
[36] Pollini N, Lavan O, Amir O. Adjoint sensitivity analysis and optimization of hysteretic dynamic systems with nonlinear viscous dampers. Struct Multidiscip Optim 2018;57(6):2273–89. https://doi.org/10.1007/s00158-017-1858-2
[37] Su C, Xian J. Topology optimization of non-linear viscous dampers for energy-dissipating structures subjected to non-stationary random seismic excitation. Struct Multidiscip Optim 2022;65(7):200. https://doi.org/10.1007/s00158-022-03281-2
[38] Marzok A, Lavan O. Seismic design of multiple-rocking systems: a gradient-based optimization approach. Earthq Eng Struct Dyn 2021;50(13):3460–82. https://doi.org/10.1002/eqe.3518
[39] Lavan O. Adjoint sensitivity analysis and optimization of transient problems using the mixed lagrangian formalism as a time integration scheme. Struct Multidiscip Optim 2020;61(2):619–34. https://doi.org/10.1007/s00158-019-02383-8
[40] Sivaselvan MV, Lavan O, Dargush GF, Kurino H, Hyodo Y, Fukuda R, Sato K, Apostolakis G, Reinhorn AM. Numerical collapse simulation of large-scale structural systems using an optimization-based algorithm. Earthq Eng Struct Dyn 2009;38(5):655–77. https://doi.org/10.1002/eqe.895
[41] The MathWorks I. MATLAB optimization toolbox: quadprog function. Natick, Massachusetts, United States: The MathWorks, Inc.; 2025. https://www.mathworks.com/help/optim/ug/quadprog.html [accessed: 16 January 2025].
[42] Liang Q, Bai X, Yu D, Li L, Ou J. Dynamic performance of structures with clutching inertia-damping system: theoretical study. Earthq Eng Struct Dyn 2024;53(3):1345–62. https://doi.org/10.1002/eqe.4072
[43] Ma R, Bi K, Hao H. Inerter-based structural vibration control: a state-of-the-art review. Eng Struct 2021;243:112655. https://doi.org/10.1016/j.engstruct.2021.112655
[44] Ben-Israel NI, Lavan O. Efficient design of truss footbridges with stiffness-proportional inerters. Struct Infrastruct Eng 2024:1–14. https://doi.org/10.1080/15732479.2024.2340085
[45] Takewaki I. Building control with passive dampers: optimal performance-based design for earthquakes. John Wiley & Sons; 2009.
[46] S. Günay, K. M. Mosalam, Peer performance-based earthquake engineering methodology, revisited, Journal of Earthquake Engineering 17 (6) (2013) 829–858. https://doi.org/10.1080/13632469.2013.787377.
[47] Grefenstette JJ. Genetic algorithms and machine learning. In: Proceedings of the sixth annual conference on computational learning theory; 1993. p. 3–4.
[48] Kennedy J, Eberhart R. Particle swarm optimization. In: Proceedings of ICNN'95 - international conference on neural networks, vol. 4. 1995. p. 1942–8. https://doi.org/10.1109/ICNN.1995.488968
[49] Hare W, Nutini J, Tesfamariam S. A survey of non-gradient optimization methods in structural engineering. Adv Eng Softw 2013;59:19–28. https://doi.org/10.1016/j.advengsoft.2013.03.001
[50] Mei L, Wang Q. Structural optimization in civil engineering: a literature review. Buildings 2021;11(2). https://doi.org/10.3390/buildings11020066
[51] Ohtori Y, Christenson RE, Spencer BF, Dyke SJ. Benchmark control problems for seismically excited nonlinear buildings. J Eng Mech 2004;130(4):366–85. https://doi.org/10.1061/(ASCE)0733-9399(2004)130:4(366)
[52] Gaiotti R, Smith BS. P-delta analysis of building structures. J Struct Eng 1989;115(4):755–70. https://doi.org/10.1061/(ASCE)0733-9445(1989)115:4(755)
[53] Marini F, Walczak B. Particle swarm optimization (PSO). A tutorial. Chemom Intell Lab Syst 2015;149:153–65. https://doi.org/10.1016/j.chemolab.2015.08.020