Multi-objective parameter and topology optimization of structures equipped with tuned viscous mass dampers using automatic differentiation

Vitali Kakouka[a,*], Kohju Ikago[b]

[a]Graduate School of Engineering, Tohoku University, 6-6-11 Aramaki Aoba, 980-8572, Sendai, Japan
[b]Graduate School of Engineering, Kyoto University, C1-4-382 Katsura Nishikyo, 615-8540, Kyoto, Japan

Abstract

Traditional earthquake-resistant design often prioritizes low upfront costs by relying on structural ductility, which inherently leads to severe inelastic damage, high repair expenses, and extensive operational downtime following major seismic events. On the other hand, passive control technologies like tuned viscous mass dampers (TVMDs) offer supplemental damping without external power requirements, and allow to keep the structure essentially in the elastic range during extreme seismic events. However, optimizing both damper parameters and spatial topology within multiple-degree-of-freedom (MDOF) systems remains computationally prohibitive due to non-classical damping, modal interactions, and the non-smooth nature of real-world time-domain design objectives. To overcome these limitations, this study introduces an original formulation for the simultaneous multi-objective parameter and topology optimization of TVMDs. Furthermore, we propose a novel, high-performance computing framework that leverages hardware-accelerated automatic differentiation (AD) implemented via the PyTorch library that backs the latest developments in machine learning (ML) and artificial intelligence (AI). By treating the dynamic structural analysis program as a differentiable computation tree, exact, down to machine precision, sensitivities are systematically propagated through time-domain analyses and then used by the optimization algorithm. The proposed AD-powered framework bypasses the computational bottlenecks, truncation errors, and step-size sensitivities associated with conventional finite-difference gradient approximations. The efficiency and numerical robustness of the approach are validated using a ten-story benchmark shear building subjected to strong ground motions. The results demonstrate that the proposed method significantly accelerates the optimization convergence, lowers computational overhead, and provides structural engineers with a highly scalable tool to rapidly deliver safe, economic, and high-performance seismic designs.

Keywords: seismic control, tuned viscous mass damper, optimal design, automatic differentiation, topology optimization

1. Introduction

Traditionally, earthquake-resistant design minimizes upfront costs by relying on the ductility of lateral load-resisting systems. This approach dissipates seismic energy through inelastic (hysteretic) behavior [1, 2]. However, this strategy often results in significant structural damage, leading to high repair costs and prolonged downtime in the aftermath of severe earthquakes. To address these vulnerabilities, extensive research over the past decades has focused on the development and application of passive viscous and viscoelastic dampers. These systems provide supplemental damping to minimize seismic losses, ideally maintaining a linear structural response even during major earthquake events. Passive control technologies are widely used in civil engineering owing to their simplicity and independence from external power supplies [3–7]. However, conventional dampers often exhibit insufficient control forces under long-period ground motions, which compromises their seismic control effectiveness [8, 9]. To overcome this limitation, inerter-based vibration control systems (IVCSs) have attracted significant research attention due to their capability to enhance energy dissipation and improve structural performance [6, 10–14].

*Corresponding author
Email address: kakouka.vitali.q5@dc.tohoku.ac.jp, v.kakouka@outlook.com (Vitali Kakouka)

The inerter [15] is a two-terminal mechanical element that generates a force proportional to the relative acceleration between its terminals, completing the force-current analogy between mechanical and electrical networks. Various physical realizations have been developed, including ball-screw [5, 16], rack-and-pinion [15], hydraulics [17], fluid [7], electromagnetics [18] and eddy current [19]. Through mass amplification that allows to avoid the weight penalty on structure as in case of traditional tuned mass dampers (TMD) [20], inerters enable engineers to effectively control the structural response by placing them in multiple locations of a structure.

Among representative IVCSs, Ikago et al. [10] proposed the tuned viscous mass damper (TVMD), which achieves a large apparent mass with limited physical mass and has been implemented in practical structures in Japan [21]. Lazar et al. [6] introduced the tuned inerter damper (TID). Marian et al. [22] proposed the tuned mass damper inerter (TMDI). The damping enhancement mechanism and its relationship to response reduction have also been theoretically clarified [12, 23]. Kang et al. [24] demonstrated the practical application of TVMDs spanning three stories. Xie et al. [13] studied the topology of viscoelastically supported viscous mass damper (VeVMD) [25] that combines the TVMD and TID configurations, and concluded that the acceleration response of structure controlled with TVMDs could be significantly improved by slightly compromising the displacement control capacity.

Closed-form optimal design solutions have been established for single degree of freedom (SDOF) systems under H_∞ and H_2 criteria [10, 22]. Such formulations are mathematically convenient, have high interoperability of underlying physics and often admit compact analytical expression. However, extending these results to multiple degree of freedom (MDOF) structures remains challenging due to non-classical damping and modal interactions. Approximate analytical methods [11] and numerical optimization approaches [12, 26] have therefore been developed to determine optimal parameters and device distributions. Practical engineering design is typically concerned with time-domain performance, and in real-world applications engineers aim to directly control peak structural responses, such as limiting maximum displacements or inter-story drifts (displacement control), maximum accelerations (acceleration control), or a combination of both within a multi-objective optimization problem, and minimization of transfer functions in frequency domain does not always lead to the optimal and most economic design for a particular site, and thus parameters acquired as a result of such optimization often require further adjusting. These performance design metrics are highly dependent on ground motion characteristics at a specific site, and their calculation involves usage of non-smooth operators, such as maximum, mean, and absolute value, which naturally arise when setting peak or averaged response limits.

A major challenge in time-domain optimization lies in the computation of gradients of the objective functions and constraints with respect to design variables. For realistic structural models, especially those involving transient dynamic analysis, the governing equations are implemented as complex computer programs. The resulting objective and constraint functions, that represent target performance, generally do not admit closed-form symbolic derivatives. Consequently, widely used constrained nonlinear optimization algorithms, such as sequential quadratic programming (SQP) [27], sequential least squares programming (SLSQP) [28], or bound-constrained quasi-Newton methods like limited-memory Broyden-Fletcher-Goldfarb-Shanno with bound constraints method (L-BFGS-B) [29] often approximate Hessians numerically using finite differences, effectively treating the structural analysis code as a "black box" instead of extracting useful features that could help to evaluate gradients in a more robust way. This approach suffers from several drawbacks: finite-difference gradients are sensitive to step-size selection, introduce truncation and round-off errors, and require multiple function evaluations per design variable [30]. For large-scale or computationally intensive dynamic simulations, this leads to substantial computational cost associated with repetitive inversion of system matrices and potentially degraded convergence behavior. This challenge can be addressed by method backing the recent major advance in artificial intelligence (AI) and deep neural networks (DNN) – automatic differentiation (AD).

AD has emerged as a cornerstone of computational science, providing an efficient mechanism for evaluating the derivatives of functions defined through computer programs [31]. Unlike numerical differentiation, which suffers from truncation errors, or symbolic differentiation, which can lead to "expression swell," when intermediate mathematical expressions grow exponentially in number of terms or complexity during derivation or simplification, even if the final result is relatively simple, AD decomposes complex programs into sequences of elementary operations. By systematically applying the chain rule, AD computes exact derivatives to machine precision [30].

The versatility of AD lead to breakthroughs in diverse domains, ranging from machine learning [32] and molecular dynamics [33] to computational fluid dynamics [34]. Within structural optimization, the adoption of AD traces back to Ozaki et al. [35], who utilized Fortran-based analysis frameworks for mechanical systems. This was followed by specialized applications in shell optimization using Non-Uniform Rational B-Splines (NURBS) [36] and topology optimization [37]. Wu et al. [38] recently developed an open source structural topology optimization framework JAX-SSO based on JAX [39] library and demonstrated the potential and robustness of applying AD for real-world structural engineering problems.

Despite these foundational steps, AD is not yet commonly applied in structural topology and parameters optimization. Historically, this stagnation was attributed to the computational overhead and complexity of legacy AD packages [40], as well as limited availability of commercial-grade hardware or cloud computing. Consequently, the power of modern software libraries, such as PyTorch [41] or JAX, and parallel hardware and graphics processing units (GPUs) acceleration has not been fully exploited for structural optimization problems.

This study proposes original formulation of multi-objective TVMD optimal design problem for simultaneous parameter and topology (distribution along the structure) optimization. To overcome the limitations of current optimization approaches and make the stated problem tractable, this study also proposes a novel framework for damping optimization based on GPU-accelerated AD using PyTorch library. By leveraging recent advances in differentiable programming, high-performance computing, and hardware acceleration, gradients of time-domain performance measures can be computed efficiently and accurately. The rapid increase in computational power, together with the widespread availability of high-performance GPUs and open-source software libraries, has made AD practically accessible to engineers without requiring the explicit development of specialized differentiation code or advanced programming skills.

Importantly, AD can handle complex program structures, including loops, conditional statements, and non-smooth yet piecewise-differentiable operators such as maximum, mean and absolute ($|\cdot|$) value, which are common in engineering optimization problems and that introduce numerical instabilities when evaluating the gradients numerically. In addition, AD allows to effectively propagate the gradients through Fast Fourier Transform (FFT) and Inverse Fast Fourier Transform (iFFT) [42] operations, making it highly attractive for dynamic analysis of structures. As a result, the proposed method enables accurate sensitivity analysis for time-domain objectives and constraints, leading to improved convergence properties, enhanced numerical stability, and significantly reduced computational cost compared to conventional black-box gradient approximations.

By combining the recent advances in high-performance computing and modern machine learning frameworks, and the extensive theory of damper optimization, this study aims to provide engineers with valuable information and practical examples that could help them to significantly reduce the cycle of iterative design and thus deliver safe design more quickly, which is crucial in the competitive field of structural engineering. The rest of the paper is organized as follows: in Section 2 the problem of optimal passive control of single degree of freedom (SDOF) and multiple degree of freedom (MDOF) structures using tuned viscous mass dampers (TVMD) is presented. The section concludes with the definition of optimal damping design problem by setting the structural performance goals, then extends it for the problem of optimal damper distribution, and discusses the advantages and disadvantages of current optimization algorithms Section 3 first gives a brief introduction to AD and discusses its attractiveness for the problem of optimal structural seismic control, and then introduces the proposed AD-backed optimization framework. Finally, in Section 4 the proposed approach is applied for optimization of damper parameters and topology optimization of TVMDs incorporated into a ten-story benchmark shear building subjected to strong ground motions, with further discussion on its efficiency and accuracy.

2. Optimal passive control of structures using tuned viscous mass dampers (TVMD)

2.1. Optimal control of single degree of freedom (SDOF) structures

Figure 2.1 shows an undamped SDOF structure incorporated with a TVMD. The mass and stiffness of the primary structure are m and k respectively, and the additional mass, stiffness of the supporting spring and the damping coefficient are m_r, k_b and c_d, respectively. u and u_d denote the relative displacement of the primary system and the inerter, and u_g is the ground displacement.
The equation of motion of such a system can be written as:

$$\begin{cases} m\ddot{u}(t) + ku(t) + P(t) = -m\ddot{u}_g(t) \\ P(t) = m_r\ddot{u}_d(t) + c_d\dot{u}_d(t) = k_b(u(t) - u_d(t)) \end{cases} \tag{1}$$

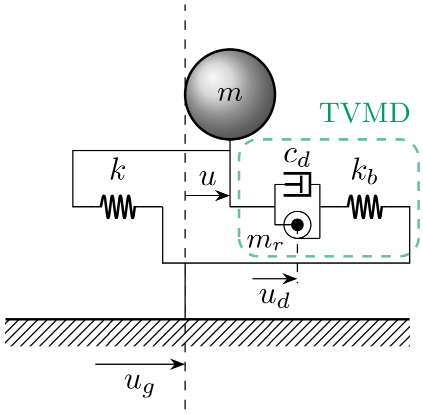


Figure 2.1: Undamped SDOF structure equipped with a TVMD.

where $P(t)$ is the damper force ignoring friction, and the over-dot $(\dot{\ })$ denotes differentiation with respect to time t

It can be noted, that unlike the TMD, TVMD is activated by the inter-story drift and the spring and dashpot are arranged in series. This difference causes a difference in the resonance curves and the optimal designs for the TVMD and TMD. Nonetheless, the fixed point method is applicable to a seismic control system with the TVMD [11]. According to the fixed point method, the optimum angular frequency of the additional vibration system ω_r^{opt}, which is to minimize the peak value of the dynamic amplification factor (DAF), is derived as [11]:

$$\omega_r^{\text{opt}} = \frac{1}{\sqrt{1-\mu}}\,\omega_0 \tag{2}$$

where μ is the mass ratio - the ratio of TVMD inertance to the primary structure mass:

$$\mu = \frac{m_r}{m} \tag{3}$$

Therefore, the optimal design of an undamped SDOF structure serves as a foundation that extends to MDOF systems, as detailed in the next section.

2.2. *Optimal displacement and acceleration control of multiple degree of freedom (MDOF) structures*

2.2.1. *Seismic response of a MDOF structure controlled with TVMDs*

Consider a damped multiple-degree-of-freedom (MDOF) structure incorporated with TVMDs, one device per degree of freedom, n TVMDs in total, shown in Figure 2.1. The equation of motion of such a model can be defined as:

$$\mathbf{M}\ddot{\mathbf{u}}(t) + \mathbf{C}\dot{\mathbf{u}}(t) + \mathbf{K}\mathbf{u}(t) = -\mathbf{M}\boldsymbol{\iota}\ddot{u}_g(t) \tag{4}$$

where $\ddot{u}_g(t)$ is ground acceleration, $\mathbf{u}(t)$ is the $2n$ relative displacements vector, that consists of primary system displacements vector $\mathbf{u}_p$ and damping system displacements vector $\mathbf{u}_d$:

$$\mathbf{u} = \begin{bmatrix}\mathbf{u}_p\\ \mathbf{u}_d\end{bmatrix}_{2n} \tag{5}$$

where

$$\mathbf{u}_p = \{u_{p1}, u_{p2}, \cdots, u_{pn}\}_n^T \tag{6}$$
$$\mathbf{u}_d = \{u_{d1}, u_{d2}, \cdots, u_{dn}\}_n^T \tag{7}$$

$\boldsymbol{\iota}$ is the influence coefficient vector, that distributes ground motion DOF to mass DOFs, and consists of an all-ones vector $\mathbf{1}_n$ and an all-zeros vector $\mathbf{0}_n$, as the inerters are activated not by the ground motion, but by inter-story drifts:

$$\boldsymbol{\iota} = \begin{bmatrix}\mathbf{1}_n\\ \mathbf{0}_n\end{bmatrix}_{2n} \tag{8}$$

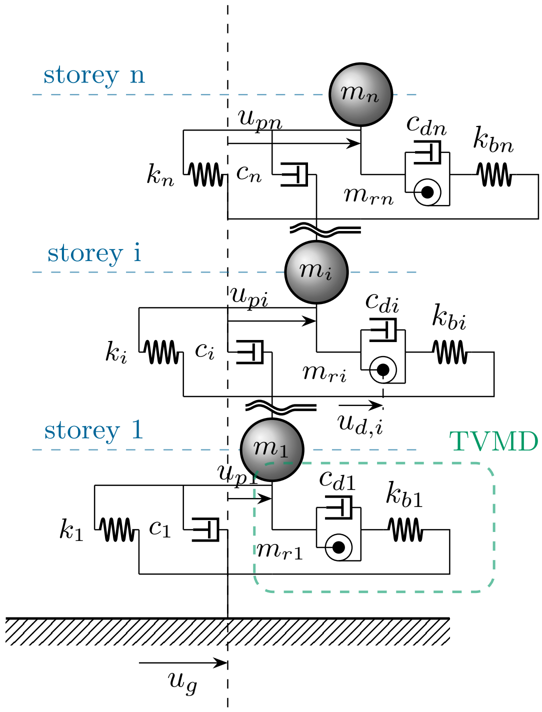


Figure 2.2: Analytical model of a damped MDOF structure equipped with TVMDs.

$\mathbf{M}$, $\mathbf{C}$ and $\mathbf{K}$ are global mass, damping and stiffness matrices of the system, respectively:

$$\mathbf{M} = \begin{bmatrix} \mathbf{M}_p & \mathbf{0} \\ \mathbf{0} & \mathbf{M}_r \end{bmatrix} \tag{9}$$

$$\mathbf{C} = \begin{bmatrix} \mathbf{C}_p & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_d \end{bmatrix} \tag{10}$$

$$\mathbf{K} = \begin{bmatrix} \mathbf{K}_p + \mathbf{K}_{b11} & \mathbf{K}_{b12} \\ \mathbf{K}_{b21} & \mathbf{K}_{b22} \end{bmatrix} \tag{11}$$

where $\mathbf{M}_p$, $\mathbf{C}_p$ and $\mathbf{K}_p$ are mass, inherent damping and stiffness matrices of the primary system:

$$\mathbf{M}_p = diag(\mathbf{m}) = \begin{bmatrix} m_1 & 0 & \cdots & 0 \\ 0 & m_2 & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ 0 & \cdots & 0 & m_n \end{bmatrix} \tag{12}$$

$$\mathbf{C}_p = \begin{bmatrix} c_1 + c_2 & -c_2 & 0 & 0 \\ -c_2 & c_2 + c_3 & \ddots & \vdots \\ 0 & \ddots & -c_{n-1} + c_n & -c_n \\ 0 & \cdots & -c_n & c_n \end{bmatrix} \tag{13}$$

$$\mathbf{K}_p = \begin{bmatrix} k_1 + k_2 & -k_2 & 0 & 0 \\ -k_2 & k_2 + k_3 & \ddots & \vdots \\ 0 & \ddots & -k_{n-1} + k_n & -k_n \\ 0 & \cdots & -k_n & k_n \end{bmatrix} \tag{14}$$

and matrices of the supplementary damping system $\mathbf{M}_r$, $\mathbf{C}_d$ and $\mathbf{K}_b$ are defined as:

$$\mathbf{M}_r = diag(\mathbf{m}_r) = \begin{bmatrix} m_{r1} & 0 & \cdots & 0 \\ 0 & m_{r2} & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ 0 & \cdots & 0 & m_{rn} \end{bmatrix} \tag{15}$$

$$\mathbf{C}_d = diag(\mathbf{c}_d) = \begin{bmatrix} c_{d1} & 0 & \cdots & 0 \\ 0 & c_{d2} & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ 0 & \cdots & 0 & c_{dn} \end{bmatrix} \tag{16}$$

$$\mathbf{K_{b11}} = \begin{bmatrix} k_{b1}+k_{b2} & -k_{b2} & \cdots & 0 \\ -k_{b2} & k_{b2}+k_{b3} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & k_{bn} \end{bmatrix} \tag{17}$$

$$\mathbf{K}_{b12} = \begin{bmatrix} -k_{b1} & k_{b2} & \cdots & 0 & 0 \\ 0 & -k_{b2} & k_{b3} & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & -k_{b,n-1} & k_{bn} \\ 0 & 0 & 0 & 0 & -k_{bn} \end{bmatrix} \tag{18}$$

$$\mathbf{K}_{b21} = \mathbf{K}_{b12}^T \tag{19}$$

$$\mathbf{K}_{b22} = diag(\mathbf{k}_b) = \begin{bmatrix} k_{b1} & 0 & \cdots & 0 \\ 0 & k_{b2} & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ 0 & \cdots & 0 & k_{bn} \end{bmatrix} \tag{20}$$

By performing the Fourier transformation of Equation (4), substituting $\dot{\mathbf{U}}(\omega) = i\omega\mathbf{U}(\omega)$ and $\ddot{\mathbf{U}}(\omega) = -\omega^2\mathbf{U}(\omega)$ and grouping the terms, we obtain the governing equation of motion of the controlled system in frequency domain:

$$(\mathbf{K} + i\omega\mathbf{C} - \omega^2\mathbf{M})\mathbf{U}(\omega) = -\mathbf{M}\boldsymbol{\iota}\ddot{U}_g(\omega) \tag{21}$$

where $\mathbf{U}(\omega)$ and $\ddot{U}_g(\omega)$ are Fourier transforms of displacements of the system and ground acceleration, respectively:

$$\mathbf{U}(\omega) = \mathcal{F}\{\mathbf{u}(t)\} \tag{22}$$
$$\ddot{U}_g(\omega) = \mathcal{F}\{\ddot{u}_g(t)\} \tag{23}$$

For convenience, we will define the system matrix $\mathbf{A}(\omega)$ as:

$$\mathbf{A}(\omega) \equiv \mathbf{K} + i\omega\mathbf{C} - \omega^2\mathbf{M} \tag{24}$$

Then, relative displacement response $\mathbf{U}(\omega)$ and total (absolute) acceleration response $\ddot{\mathbf{U}}_t(\omega)$ in frequency domain can be expressed as:

$$\mathbf{U}(\omega) = \begin{bmatrix} \mathbf{U}_p(\omega) \\ \mathbf{U}_d(\omega) \end{bmatrix} = -\mathbf{M}\boldsymbol{\iota}\ddot{U}_g(\omega)\mathbf{A}(\omega)^{-1} \tag{25}$$

$$\ddot{\mathbf{U}}_t(\omega) = \begin{bmatrix} \ddot{\mathbf{U}}_{p,t}(\omega) \\ \ddot{\mathbf{U}}_{d,t}(\omega) \end{bmatrix} = \ddot{\mathbf{U}}(\omega) + \ddot{\mathbf{U}}_\mathbf{g}(\omega) = -\omega^2\mathbf{U}(\omega) + \boldsymbol{\iota}\ddot{U}_g(\omega) \tag{26}$$

An inter-story drifts response vector $\mathbf{\Delta_p}(\omega)$ in frequency domain is calculated as a product of transformation matrix $\mathbf{T}_r$ and relative story displacements vector of the primary system $\mathbf{U}_p(\omega)$:

$$\mathbf{\Delta_p}(\omega) = \mathbf{T}_r \mathbf{U}_p(\omega) \tag{27}$$

where $\mathbf{T}_r$ is the $n \times n$ matrix that transforms story displacements to inter-story drifts:

$$\mathbf{T}_r = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ -1 & 1 & \ddots & \vdots & \vdots \\ 0 & \ddots & \ddots & 0 & 0 \\ \vdots & \cdots & -1 & 1 & 0 \\ 0 & \cdots & 0 & -1 & 1 \end{bmatrix} \tag{28}$$

Finally, to obtain the inter-story drifts $\boldsymbol{\delta}_p(t)$ and total acceleration response $\ddot{\mathbf{u}}_{p,t}(t)$ time histories of the primary structure in time domain we perform inverse Fourier transform of Equations (26) and (27):

$$\boldsymbol{\delta}_p(t) = \mathcal{F}^{-1}\{\boldsymbol{\delta}_p(\omega)\} \tag{29}$$

$$\ddot{\mathbf{u}}_{p,t}(t) = \mathcal{F}^{-1}\{\ddot{\mathbf{U}}_{p,t}(\omega)\} \tag{30}$$

This establishes the foundation for defining the TVMD optimal design problem.

2.2.2. TVMD parameter optimization problem

For optimal structural design, the primary objective function and the constraints typically rely on peak response values across all degrees of freedom to serve as performance metrics. Consequently, the maximum absolute inter-story drift angle $\delta_{p,\max}$ and the maximum absolute total acceleration $\ddot{u}_{p,\max}$ of the primary structure are defined as:

$$\delta_p^{max} = \max_{t,i=1,\ldots,n} |\delta_{p,i}(t)/h_i| \tag{31}$$

$$\ddot{u}_p^{max} = ||\ddot{\mathbf{u}}_{p,t}||_\infty \tag{32}$$

where $||\cdot||_\infty$ operator denotes the infinity norm, and h_i is the height of the ith storey.

In this way, for a set of k selected ground motions $\ddot{\mathbf{u}}_g = \{\ddot{u}_{g,1}, \cdots, \ddot{u}_{g,k}\}$, representative of the site, a general multi-objective inter-story drift and acceleration response minimization problem can be defined as:

$$\underset{\mathbf{X}}{\text{minimize}} \quad \{\delta_p^{max}(\mathbf{X}, \ddot{\mathbf{u}}_g), \ddot{u}_p^{max}(\mathbf{X}, \ddot{\mathbf{u}}_g)\} \tag{33}$$

where $\mathbf{X}$ is the parameter vector of the supplementary damping system:

$$\mathbf{X} = \{\mathbf{m}_r, \mathbf{c}_d, \mathbf{k}_b\}^T \in \mathbb{R}^{3n} \tag{34}$$

The solution to this problem by itself does not necessarily result in an economical design. In practical structural engineering applications, design objectives typically involve constraining maximum structural responses, such as peak displacements or inter-storey drifts (displacement control), peak accelerations (acceleration control), or a combination of both, in order to satisfy the prescribed performance level. Displacement and acceleration controls, however, generally represent competing design paradigms, and thus an appropriate trade-off needs to be introduced. The TVMD was originally developed with displacement control as the primary objective, with optimal design achieved through minimization of the displacement transfer function [11]. Nevertheless, as demonstrated by Xie et al. [13], acceleration control can still be attained in structures equipped with TVMDs, albeit with slight compromise in the displacement control performance.

It can be noted from Equation (34) that for an untuned configuration, there are $3n$ independent design parameters. However, TVMDs achieve optimal seismic response reduction when tuned to the dominant modes of a MDOF structure. Although implementing multiple independent TVMD types tuned to each corresponding mode can mitigate higher-order modal responses, this study focuses exclusively on the first-mode control.

Similarly to the Equation (2) for a SDOF system, the optimum angular frequency of the additional vibration system for the first mode of vibration of a MDOF structure can be expressed as:

$$\omega_r = \frac{1}{\sqrt{1-\mu}}\omega_1 \tag{35}$$

with the modal mass ratio μ defined as:

$$\mu = \frac{M_{r,1}}{M_{p,1}} = \frac{\phi_1^T \mathbf{T}_r^T \mathbf{M}_r \mathbf{T}_r \phi_1}{\phi_1^T \mathbf{M}_p \phi_1} \tag{36}$$

where ϕ_1 is the first modal shape for the uncontrolled undamped structure and M_r and M_p are the mass matrices of the distributed TVMD and the primary structure, respectively.

And the rest of the optimal damper parameters – the TVMD spring stiffnesses vector $\mathbf{k}_b$ and damping coefficients vector $\mathbf{c}_d$ – are defined as:

$$\mathbf{k}_b = \omega_r^2 \mathbf{m}_r \tag{37}$$

$$\mathbf{c}_d = 2\zeta_d \omega_r \mathbf{m}_r \tag{38}$$

where ζ_d is the modal damping ratio.

Therefore, after eliminating the dependent parameters, parameter vector $\mathbf{X}$ reduces to:

$$\mathbf{X} = \{\mathbf{m}_r, \zeta_d\}^T \in \mathbb{R}^{n+1} \tag{39}$$

The multi-objective optimization problem defined in Equation (33) can be converted to a single-objective problem by introducing the weighted sum of the objectives δ_p^{max} and $\ddot{u}_p^{max}$. After defining the objective $\mathbf{Y} = \{\delta_p^{max}, \ddot{u}_p^{max}\}^T$, weight $\mathbf{\Theta} = \{\theta_1, \theta_2\}^T$ and scale $\mathbf{\Xi} = \{\xi_1, \xi_2\}^T$ vectors, response limit constraints and positive and upper-bound parameter constraints, the single-objective TVMD optimal design problem can be stated as:

$$\begin{aligned} \underset{\mathbf{X}}{\text{minimize}} \quad & y = \theta_1 \xi_1 \delta_p^{max} + \theta_2 \xi_2 \ddot{u}_p^{max} \\ \text{subject to} \quad & \delta_p^{max} - \delta_{lim} \leq 0, \\ & \ddot{u}_p^{max} - \ddot{u}_{lim} \leq 0, \\ & m_{r,i} - m_{r,lim} \leq 0, \\ & -m_{r,i} \leq 0, \\ & -\zeta_d \leq 0 \end{aligned} \tag{40}$$

where δ_{lim} and $\ddot{u}_{lim}$ are the upper limits of inter-story drift angle and absolute acceleration, respectively, $m_{r,lim}$ is the upper limit of available inertance for a single device, θ_1 and θ_2 are the weight factors, indicating the contribution of each term, and ξ_1 and ξ_2 are the factors, introduced to bring the terms to the same unitless scale.

Two edge cases of the values of the weight vector $\mathbf{\Theta}$ can be noted. Setting $\mathbf{\Theta} = \{1, 0\}^T$ reduces the optimization formulation in Equation (40) to a pure displacement control problem. Conversely, selecting $\mathbf{\Theta} = \{0, 1\}^T$ transforms the objective into a pure acceleration control problem.

It can be observed, that the optimal solution to the above-stated problem yields a set of n distinct damper inertances $m_{r,i}$ alongside the modal damping ratio ζ_d. However, deploying a unique damper specification on every storey introduces significant practical and economic challenges during fabrication and installation. From a structural engineering perspective, standardizing a limited number of device types across multiple stories is vastly preferred to minimize costs and simplify quality control. As a result, the engineering objective must be not only to identify the optimal damper parameters, but also address the optimal grouping of these devices into a number of types and determine their spatial distribution throughout the structure.

2.2.3. Proposed simulteneous damper parameter and distribution (topology) optimization problem formulation

In order to solve the damper distribution optimization problem, current study suggests introducing an additional objective term in a form of K-means clustering [43] loss L_c, also called the inertia or within-cluster sum of squares

(WCSS), to the Equation (40). That would stimulate the intertances to group into the specified p inertance types (or clusters) $m_{c,j}$, $j \in \{1, \ldots, p\}$ by creating the "attraction force", allowing for simultaneous parameter and topology optimization.

Thus, the primary objective of the clustering task is to partition the n-dimensional vector of inertances, $\mathbf{m}_r = \{m_{r,1}, m_{r,2}, \ldots, m_{r,n}\}^T$, into p discrete device types represented by the parameter vector of cluster centers, $\mathbf{m}_c = \{m_{c,1}, m_{c,2}, \ldots, m_{c,p}\}^T$. The loss term that minimizes the total within-cluster variance by mapping damper intertance in each storey to its nearest type (or cluster center), is defined as:

$$L_c(\mathbf{m}_r, \mathbf{m}_c) = \arg\min_{\mathbf{S}} \sum_{k=1}^{p} \sum_{m_{r,i} \in S_k} (m_{r,i} - m_{c,k})^2 \tag{41}$$

where S_k is the subset of inertances assigned to a kth damper type.

Therefore, after adding the clustering loss term to the objective function, the objective function, weight and scale vectors are redefined as, respectively:

$$\mathbf{Y} = \{\delta_p^{max}, \ddot{u}_p^{max}, L_c\}^T \tag{42}$$

$$\mathbf{\Theta} = \{\theta_1, \theta_2, \theta_3\}^T \tag{43}$$

$$\mathbf{\Xi} = \{\xi_1, \xi_2, \xi_3\}^T \tag{44}$$

Because the clustering formulation introduces p additional parameters to define the damper types, the design vector $\mathbf{X}$ from Equation (39) is augmented as:

$$\mathbf{X} = \{\mathbf{m}_r, \mathbf{m}_c, \zeta_d\}^T \in \mathbb{R}^{n+p+1} \tag{45}$$

Finally, the parameter optimization problem from Equation (40) can be rewritten as a simultaneous parameter and topology optimal design problem in a vector form:

$$\begin{aligned} \underset{\mathbf{X}}{\text{minimize}} \quad & y(\mathbf{X}, \ddot{\mathbf{u}}_g) = \mathbf{1}^T(\mathbf{\Theta} \odot \mathbf{\Xi} \odot \mathbf{Y}(\mathbf{X}, \ddot{\mathbf{u}}_g)) \\ \text{subject to} \quad & \delta_p^{max} - \delta_{lim} \leq 0, \\ & \ddot{u}_p^{max} - \ddot{u}_{lim} \leq 0, \\ & m_{r,i} - m_{r,lim} \leq 0, \\ & -m_{r,i} \leq 0 \quad \forall i \in \{1, \ldots, n\}, \\ & -m_{c,j} \leq 0 \quad \forall j \in \{1, \ldots, p\}, \\ & -\zeta_d \leq 0 \end{aligned} \tag{46}$$

where $\odot$ denotes Hadamard product (element-wise vector multiplication), and $\mathbf{1} = \{1, 1, 1\}^T$ is the vector of ones used for the overall summation of scaled and weighted objective terms.

Solving this optimization problem simultaneously identifies the optimal continuous damper parameters and their distribution along the structure. By incorporating the "attraction force", the damper inertance in each story is driven toward its corresponding cluster centroid, representing an intertance type. This penalization framework effectively moves an otherwise discrete combinatorial problem into a continuous, differentiable optimization space, thereby enabling the usage of efficient gradient-based optimization methods. However, despite this relaxation, the high nonlinearity of the objective function and the nature of dynamic response evaluation introduce several computational challenges that impede the convergence of traditional optimization techniques.

2.3. Challenges associated with the classical numerical optimization methods

The formulation presented above introduces several mathematical and computational challenges that complicate its solution via traditional gradient-based optimization techniques.

First, as formulated in Equation (25), evaluating the dynamic structural response requires the inversion of the system matrix $\mathbf{A}$, an operation with an algorithmic complexity of $O(N^3)$, where $N = 2n$. Furthermore, traditional constrained optimization techniques rely on calculating the sensitivities (or gradients) of the objective function and constraints with respect to the design variables via finite difference methods. This gradient approximation scales linearly with the dimensionality of the design space, yielding a complexity of $O(S)$, where $S = n + p + 1$ denotes the length of the design vector $\mathbf{X}$, leading to S evaluations of the dynamic structural response. Consequently, relying on numerical gradient approximations introduces a prohibitive computational overhead, making the optimization framework difficult to scale and numerically intractable for the simultaneous parameter and topology optimization problem.

Moreover, evaluation of the objective function terms and constraints involves non-smooth operators such as maximum, minimum, mean and absolute value, as well as inverse fast Fourier transform (iFFT) [42], which makes the numerical evaluation of gradients highly unstable.

To address these challenges, previous approaches for damping optimization had to rely on approximations of structural response (e.g. using probabilistic methods like complete quadratic combination (CQC) [11]) to avoid recurring inversion of the system matrix and relied on numerical approximation of the gradients. This made the optimization process less straightforward and accurate, and introduced numerical instability in more complex cases. Another possible solution is to replace the numerical structural finite element (FE) model with a meta (or surrogate) machine learning model [44–48], that enables a highly efficient approximation of the structural response, especially when utilizing modern GPUs and hardware acceleration. Although this approach allows to approximate the response of linear and, to an extent, non-linear systems in some small range of parameters, this approach introduces new challenges associated with proper training of these models, generation of the training data and development of an effective neural network architecture that is capable of capturing the complex structural phenomena, that are currently not completely addressed.

Therefore, in order to make the problem formulated in Equation (46) tractable, without introducing inaccuracy and excessive complexity, the present study proposes an optimization framework that utilizes *automatic differentiation (AD)*, the technique backing the recent advances in machine learning and artificial intelligence.

3. Proposed multi-objective TVMD optimization framework using automatic differentiation

3.1. Overview of automatic differentiation (AD) for the gradient-based optimization

Automatic Differentiation (AD), also referred to as algorithmic differentiation, is a computational technique for evaluating derivatives of functions defined by computer programs. Unlike symbolic differentiation, which manipulates mathematical expressions analytically, or numerical differentiation, which treats function as a "black box" and approximates derivatives using finite differences, AD computes derivatives exactly up to machine precision by systematically applying the chain rule of calculus to elementary operations [30, 32].

Consider a function $f : \mathbb{R}^n \to \mathbb{R}^m$ implemented as a sequence of elementary operations:

$$\mathbf{v}_k = \phi_k(\mathbf{v}_{i_1}, \dots, \mathbf{v}_{i_p}), \quad k = 1, \dots, L, \tag{47}$$

where ϕ_k denotes an elementary operation (e.g., addition, multiplication, trigonometric functions, etc), and the final outputs are $\mathbf{y} = \mathbf{v}_L$. An arbitrary AD computation tree is depicted in Figure 3.1.
The Jacobian matrix $J = \partial \mathbf{y}/\partial \mathbf{x}$ can be computed by propagating derivatives alongside the evaluation of the function itself, using the chain rule:

$$\frac{\partial \mathbf{v}_k}{\partial \mathbf{x}} = \sum_j \frac{\partial \phi_k}{\partial \mathbf{v}_j} \frac{\partial \mathbf{v}_j}{\partial \mathbf{x}}. \tag{48}$$

There are two principal modes of AD that are commonly employed: *forward mode* and *reverse mode*. The former one propagates derivatives from inputs to outputs, and is computationally efficient when the number of input variables is small relative to the number of outputs. The latter one, in turn, propagates adjoint sensitivities from outputs back to inputs, and its computational cost is largely independent of the number of input variables and scales primarily with the number of outputs. Therefore, for scalar-valued objective functions ($m = 1$), reverse mode allows computation of the full gradient $\nabla f \in \mathbb{R}^n$ at a cost comparable to a small multiple of one function evaluation [30].

Although the theoretical foundations of AD were established several decades ago [30, 49], its widespread adoption has occurred only recently. This delay can be attributed to the substantial memory and computational requirements

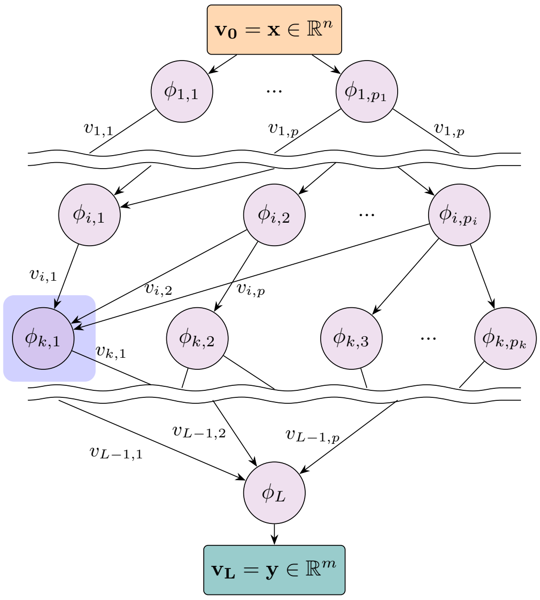


Figure 3.1: An arbitrary automatic differentiation computation tree for a function $f : \mathbb{R}^n \to \mathbb{R}^m$. Nodes represent elementary operations, and edges are inputs/outputs.

of reverse-mode AD, which requires storing intermediate variables during the forward pass for subsequent use in the backward pass. The rapid increase in computational power, availability of consumer-grade high-performance GPUs and the emergence of high-level AD-enabled software frameworks such as PyTorch [41] and JAX [39], that handle the creation of computational graph automatically and seamlessly from the user perspective, have made large-scale application of AD feasible in scientific computing and engineering optimization.

3.2. Advantages of AD for the damper optimal design problem

The use of AD provides several significant advantages over numerical differentiation approaches commonly employed in algorithms such as sequential quadratic programming (SQP) [27], sequential least squares programming (SLSQP) [28], or limited-memory Broyden-Fletcher-Goldfarb-Shanno with bound constraints method (L-BFGS-B) [29].

For the optimization problem formulated in Equation (46), with the objective $y = f(\mathbf{X})$, gradient-based optimization algorithms require evaluation of the gradient $\nabla f(\mathbf{X})$, the Jacobian matrices for the inequality constraints, and, if possible, Hessian matrices or Hessian-vector products to provide the algorithm with information about the surface curvature to improve convergence.

Traditional implementations though often rely on numerical differentiation, such as forward or central finite differences:

$$\frac{\partial f}{\partial x_i} \approx \frac{f(\mathbf{X} + \varepsilon \mathbf{e}_i) - f(\mathbf{X})}{\varepsilon}, \tag{49}$$

where ε is a small perturbation parameter. While straightforward to implement, such approximations suffer from truncation and round-off errors and would require $O(n+p+1)$ additional function evaluations per gradient computation.

In contrast, AD computes exact derivatives (up to machine precision) with no need for perturbation of the parameters. For a scalar objective function, like the one suggested in the current study, reverse-mode AD computes the full gradient at a cost typically within a constant factor of $4 \sim 6$ of a single function evaluation, i.e with algorithmic complexity $O(4 \sim 6)$ [30], which is crucial when dealing with such a computationally demanding objective function and enables the introduction of large number of parameters without hindering the performance. This advantage, enhanced by utilization of modern graphic processing units (GPU), makes AD particularly attractive for such a high-dimensional design optimization problem.

Additionally, AD eliminates truncation errors inherent in finite-difference schemes and avoids sensitivity to step-size selection. Since derivatives are generated directly from the implemented computational graph, inconsistencies between analytical derivations and numerical implementation are avoided, significantly improving the robustness and convergence.

And finally, since AD operates on the computational graph and accumulates the gradients by chain rule, it allows to evaluate the formulated objective function and constraints that contain operations like maximum, minimum, absolute, mean and inverse Fourier transformation, without introducing numerical instability and loss in precision.

Consequently, the usage of automatic differentiation not only makes the proposed damper optimal design problem practically tractable, but also computationally efficient on standard consumer-grade hardware. Such efficiency is of high importance for structural engineering practitioners, as it facilitates the rapid exploration of the design space within practical project timelines.

3.3. Proposed framework for the TVMD optimal design

To solve the optimal design problem defined in Equation (46), we employ a primal-dual optimization framework based on the method of Lagrange multipliers. This approach converts the constrained optimization problem into an unconstrained saddle-point problem, allowing for the simultaneous update of the design variables and the constraint penalties.

Let $\mathbf{g}(\mathbf{X}, \ddot{\mathbf{u}}_g) \leq \mathbf{0}$ represent the vector of inequality constraints from Equation (46), which includes the inter-storey drift and absolute acceleration performance limits, inertance limit and positivity side constraints. We introduce a vector of non-negative Lagrange multipliers $\boldsymbol{\lambda} = \{\lambda_1, \lambda_2, \ldots, \lambda_{3+2n+p}\}^T$. The augmented objective function, or the Lagrangian $\mathcal{L}(\mathbf{X}, \boldsymbol{\lambda})$, is formulated as follows:

$$\mathcal{L}(\mathbf{X}, \boldsymbol{\lambda}) = \mathbf{1}^T (\boldsymbol{\Theta} \odot \boldsymbol{\Xi} \odot \mathbf{Y}) + \boldsymbol{\lambda}^T \mathbf{g} \tag{50}$$

It can be noted, that $n + p + 1$ of the Lagrange multipliers correspond to the positivity constraints. In order to reduce the dimensionality of the $\boldsymbol{\lambda}$ vector and thus improve robustness and convergence of the optimization process, we will apply reparametrization by introducing the supplementary parameter vector $\bar{\mathbf{X}} \in \mathbb{R}^{n+p+1}$, and map it to the vector $\mathbf{X}$ using Softplus function [50]:

$$\mathbf{X} = \ln(1 + \exp(\bar{\mathbf{X}})) = \mathrm{Softplus}(\bar{\mathbf{X}}) \tag{51}$$

The Softplus function was chosen as a smooth, non-monotonic approximation of the linear rectifier, ensuring continuous differentiability at zero. However, in modern deep learning frameworks such as PyTorch, the standard Softplus function is extended to include a scaling parameter β and a numerical stability threshold ξ. The mapping $\mathbf{X} = \mathrm{Softplus}(\bar{\mathbf{X}})$ is applied element-wise, and each output element x_i is defined by the following piecewise equation:

$$x_i = \begin{cases} \frac{1}{\beta} \ln\left(1 + \exp(\beta \cdot \bar{x}_i)\right) & \text{if } \bar{x}_i \cdot \beta \leq \xi \\ \bar{x}_i & \text{if } \bar{x}_i \cdot \beta > \xi \end{cases} \tag{52}$$

This reparameterization guarantees the positivity of vector $\mathbf{X}$. As a result, the positivity constraints can be eliminated, significantly reducing the dimensionality of the Lagrange multiplier vector to $\boldsymbol{\lambda} = \{\lambda_1, \lambda_2, \ldots, \lambda_{n+2}\}^T$.

Therefore, the optimization goal is to find a saddle point $(\bar{\mathbf{X}}^*, \boldsymbol{\lambda}^*)$ that satisfies the minimax condition:

$$\max_{\boldsymbol{\lambda} \geq \mathbf{0}} \min_{\bar{\mathbf{X}}} \mathcal{L}(\bar{\mathbf{X}}, \boldsymbol{\lambda}) \tag{53}$$

As discussed earlier in Subsection 3.2, in order to evaluate the complex objective function and constraints, exact gradient information is required, and to achieve that, in this work we utilize automatic differentiation (AD). The entire forward-pass, from the creation of system matrices to the evaluation of dynamic response described in Equations (9) to (32), and constraint evaluation pipeline are implemented using PyTorch [41] library, that implicitly handles the creation of AD computation tree. While alternative frameworks like JAX [39] offer competitive performance, PyTorch is selected due to its mature ecosystem, extensive documentation, and overall ease of use. These attributes are highly critical in the field of practical structural engineering where robust software maintenance and rapid deployment are prioritized.

The optimization problem is solved using a dual-optimizer scheme utilizing the AD-generated gradients. The framework splits the updates into two distinct execution paths within a unified loop. First, a primal optimizer minimizes the Lagrangian $\mathcal{L}$ with respect to $\bar{\mathbf{X}}$ using *gradient descent* (also referred to as *steepest descent*) driving the system toward the optimum. Simultaneously, a dual optimizer maximizes the Lagrangian $\mathcal{L}$ with respect to the multipliers $\boldsymbol{\lambda}$ via *gradient ascent*, ensuring that constraint violations are penalized and thus the constraint function values are driven to zero. By alternating updates between these two optimizers, the algorithm robustly converges to the optimal design that satisfies all the specified criteria. After the optimization is complete, the value of the optimal parameter vector $\mathbf{X}^*$, which is of the primary interest, is calculated from the optimal supplementary parameter vector $\bar{\mathbf{X}}^*$ using Equation (52).

In order to get both optimal values of damper parameters and optimal distribution of the damper types (cluster centers) along the structure, it is proposed to conduct the optimization in two steps:

1. *Optimization without clustering (with the clustering loss weight factor $\theta_3 = 0$)*
   As the result of this phase we get optimal inertance values $\mathbf{m}_c^*$, unique in each storey;
2. *Fine-tuning with clustering ($\theta_3 > 0$)*
   Equation (41) is used to find the initial values of cluster centers $\mathbf{m}_c$, and optimization results in updated optimal inertances $\mathbf{m}_r^*$ and optimal intertance types $\mathbf{m}_c^*$, and the distribution is inferred from how close the values of $\mathbf{m}_r^*$ are to these centers.

The procedures described in this section establish foundation for the further application of the proposed method for practical design problems. In order to validate the effectiveness of the proposed optimization framework, in the next section it will be applied to the problem of optimal seismic control of a benchmark tall structure, following the procedures of the Japanese code and practice.

4. Numerical evaluation of the proposed method on a benchmark structure

4.1. Configuration of the target structure

With the aim of evaluating the performance and efficiency of the proposed optimization algorithm on a multi-story structure, we consider optimal control problem of the ten-story benchmark shear building model from the Japan Society of Seismic Isolation [51] employing TVMDs. Current study references procedures of the Japanese code and standard practice.

4.1.1. Properties of the benchmark structure

The shear lumped mass model of the benchmark structure is presented in Figure 4.1, and its properties and modal characteristics are listed in Tables 4.1 and 4.2, respectively.

Table 4.1: Specification of the lumped mass shear model of the benchmark structure [51]

Storey	Mass m_i [ton]	Stiffness k_i [kN/m]	Height [m]
10	875.41	158550.00	4
9	649.49	180110.00	4
8	656.22	220250.00	4
7	660.20	244790.00	4
6	667.24	291890.00	4
5	670.10	306160.00	4
4	675.71	328260.00	4
3	680.00	383020.00	4
2	681.63	383550.00	4
1	699.90	279960.00	6

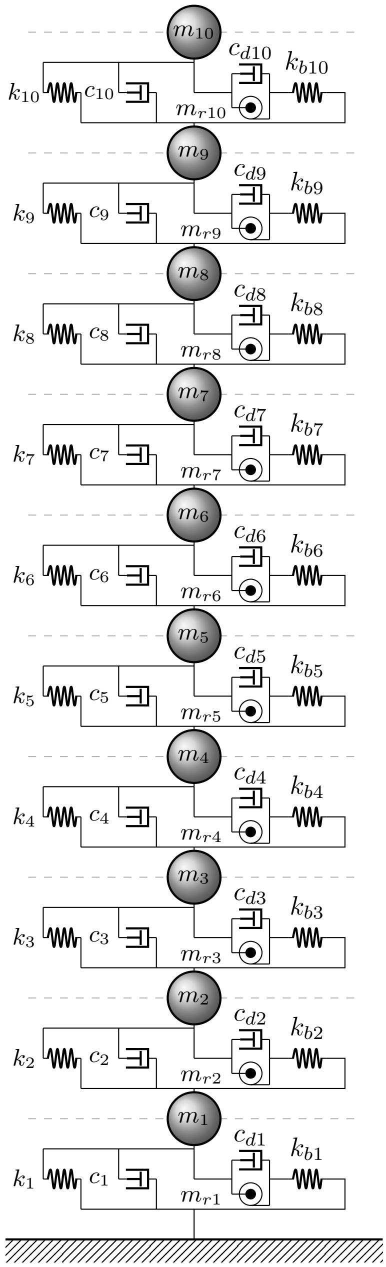


Figure 4.1: Ten-storey benchmark shear structure [51] equipped with TVMDs.

Table 4.2: Modal characteristics of the model

Mode	Period [s]	Angular frequency [rad/s]
1st	2.01	3.13
2nd	0.76	8.27
3rd	0.46	13.70

4.1.2. Inherent structural damping

In order to simulate inherent energy dissipation mechanisms, such as material friction and joint slippage, which limit unrealistic resonant vibrations during dynamic loading, and thus formulate the primary system damping matrix $\mathbf{C}_p$, we need to assign structural damping. In this study, a stiffness-proportional Rayleigh damping formulation is adopted, where the damping matrix is modeled as directly proportional to the structural stiffness matrix. In this way, by applying a stiffness-proportional formulation and setting the primary structure's first-mode damping ratio to $\zeta_1 = 0.02$ (2%), we obtain the following expression:

$$\mathbf{C}_p = \frac{2\zeta_1}{\omega_1}\mathbf{K}_p = \frac{2 \times 0.02}{3.13}\mathbf{K}_p \tag{54}$$

where $\omega_1 = 3.13\, rad/s$ is the first fundamental circular frequency of the undamped primary structure.

4.2. Implementation and validation of the dynamic analysis computation tree

4.2.1. Creation of the automatic differentiation computation tree using PyTorch

Following the procedures discussed in Subsection 3.3, the computational tree for dynamic response evaluation, from the creation of system mass, stiffness and damping matrices to the calculation of the peak response values and constraints, described in Equations (9) to (32) is implemented using PyTorch [41] library. Creation of the AD computation tree is handled completely automatically, as far as the whole procedure can be presented as a chain of elementary operations, supported by the framework. The generated computation tree for inter-storey drifts history evaluation is shown in Figure 4.2. As discussed in Subsection 3.1, every node in the tree represent an elementary operation, and the forward pass flows from the top to the bottom, and input is the parameter vector $\mathbf{X}$. For instance, node 'FftC2RBackward0' represents inverse Fourier transform from Equation (29), used to convert the inter-storey drifts response from frequency to time domain.

4.2.2. Validation of the computation tree's forward-pass

Accuracy of sensitivities evaluated during the backward pass of the computation tree primarily depends on the accuracy of the forward pass, which is used to calculate dynamic structural response time histories. Therefore, in this section the accuracy of the forward pass is validated by comparing it against the corresponding finite element (FE) model. The reference FE model is constructed as a 1D assembly within the OpenSees [52] framework using the OpenSeesPy [53] Python wrapper, and the artificial earthquake record BCJ-L2 from the Building Center of Japan (BCJ) [54] is employed as the validation input ground motion.

The parameters of the TVMD used in the model are listed in Table 4.3. Within the FE model, the components of the TVMD are modeled as: an *Inerter* element represents the inerter, a *ZeroLength* element with *Elastic* uniaxial material models the spring, and a *ZeroLength* element incorporating *ViscousDamper* material models the dashpot.

Table 4.3: TVMD parameters and distribution in the validation model

Stories	$m_r[t]$	$\omega_r[rad/s]$	ζ_d
1-3	5960.00		
4-6	4950.00	4.01	0.83
7-10	4270.00		

Figure 4.3 compares the time-history responses of the benchmark structure's top node computed via the proposed approach and the OpenSees Newmark-beta integration method. Only minor discrepancies are observed at the beginning and the end of the time history, which are attributed to link effect. The magnitude of these discrepancies depends on the number of samples, which is determined by the ground motion recording's duration and sampling rate. Because the proposed optimization algorithm relies exclusively on peak response values, these localized discrepancies are expected to introduce no significant error. Consequently, the implemented procedure is validated as accurate, and thus can be used in the proposed optimization algorithm to calculate the structural performance metrics.

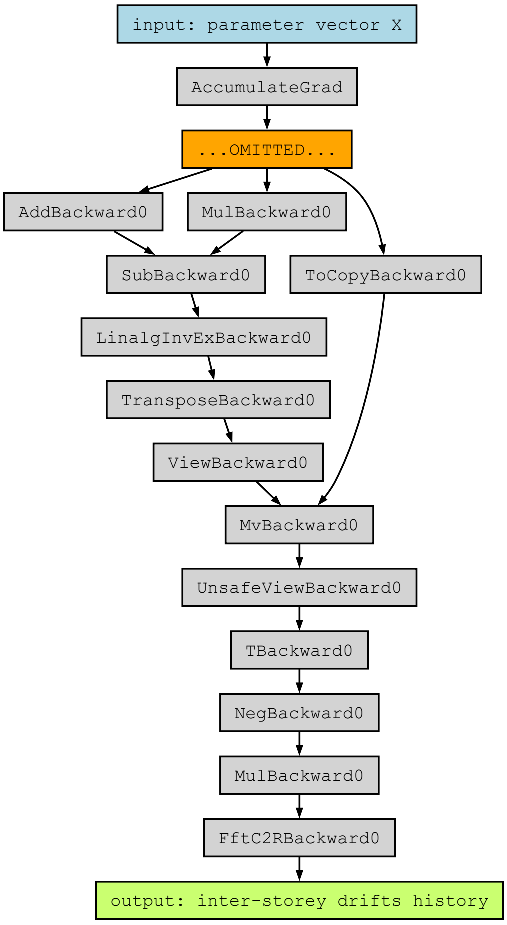


Figure 4.2: Simplified automatic differentiation computation tree for inter-storey drifts history calculation automatically generated by PyTorch. The forward pass goes from top to bottom, and the gradients are evaluated during the backward pass.

4.3. Ground motions selection and preprocessing

In order to evaluate structural performance metrics in terms of peak inter-story drifts and absolute accelerations, a suite of site-specific ground motions was selected and preprocessed for the the further use by the optimization framework.

4.3.1. Selected ground motions

Following the procedures of the Japanese code and practice, the total of six ground motion recordings were selected for the optimization process, three of which were scaled to peak ground velocity (PGA) equal to 50 kine (0.5 m/s) for the extremely rare earthquake (ERE, 500 years return period). The information about the scaled ground motions is summarized in Table 4.4.

Table 4.4: Design ground motions: recorded ground motions scaled to PGV=0.5 m/s

No.	Event name	Year	Station	Component
1	Kern County Earthquake, CA, USA	1952	Taft	E-W
2	Imperial Valley Earthquake, CA, USA	1940	El Centro	N-S
3	Tokachi-oki Earthquake, Japan	1968	Hachinohe Harbor	N-S

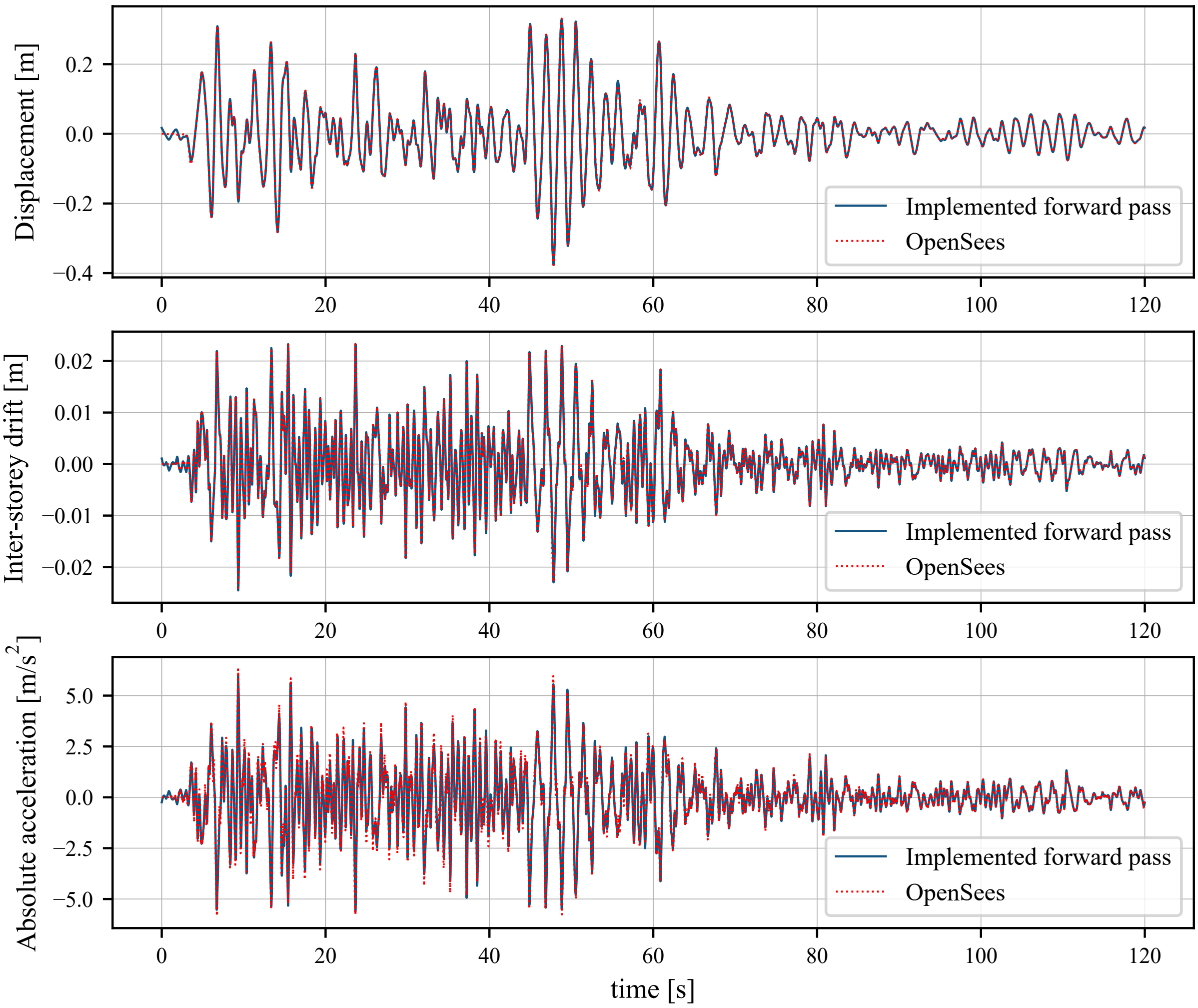


Figure 4.3: Comparison of the top-node time-history responses computed via the implemented procedure and OpenSees.

The remaining three ground motions were matched to the target design spectrum for Soil Type 1 using amplitude scaling procedures in accordance with the Building Center of Japan guidelines [55]. The summary of the seed ground motions and their post-matching PGVs is given in Table 4.5.

Table 4.5: Design ground motions: seed ground motions used for spectral matching to the target spectrum

No.	Event name	Year	Station	Comp.	PGV [m/s]
4	Kern County Earthquake, CA, USA	1952	Taft	E-W	0.81
5	Imperial Valley Earthquake, CA, USA	1940	El Centro	N-S	0.86
6	Kobe Earthquake, Japan	1995	Japan Meteorological Agency	E-W	0.79

To enable efficient GPU-accelerated matrix operations, as a final part of the preprocessing stage the selected scaled and spectrally compatible ground motions were resampled to a uniform time step of dt = 0.02 s. Additionally, the records were cropped or zero-padded to a standardized duration of 60 s. This uniform formatting allows for parallelized computation of structural responses across the entire ground motion ensemble, significantly enhancing

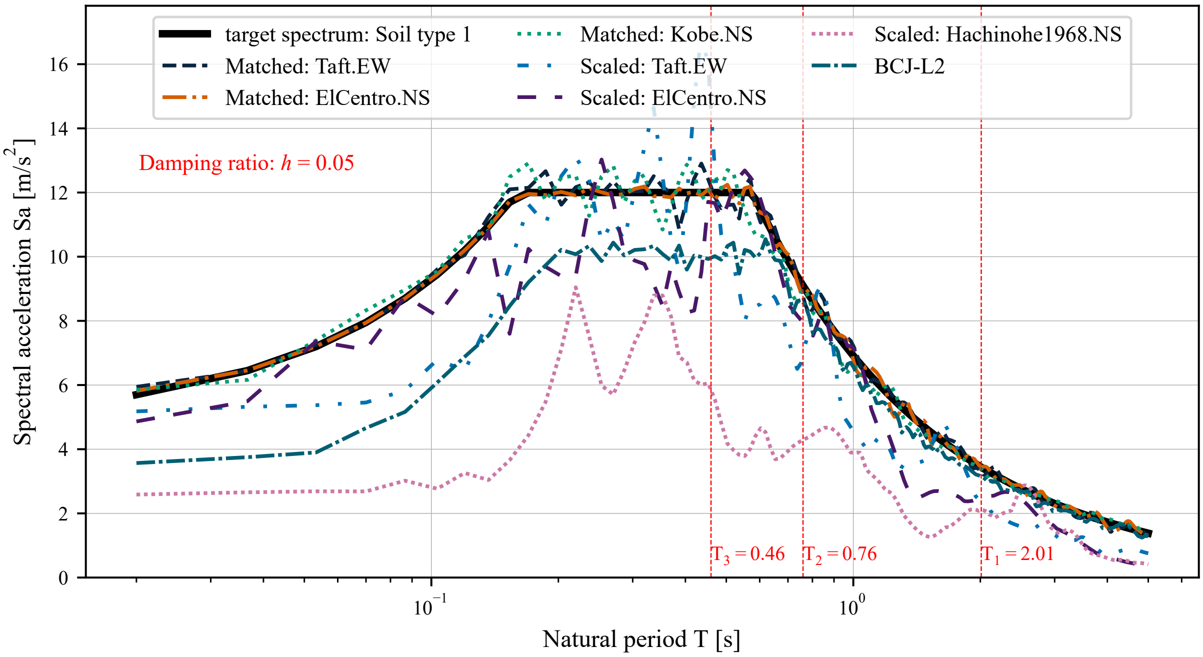


Figure 4.4: Acceleration response spectra of the ensemble of selected scaled and spectrally compatible ground motions (h = 0.05)

computational speed and enabling the algorithm to effectively identify the most critical ground motion for the TVMD configuration at the particular optimization step. The acceleration response spectra for the preprocessed ground motion ensemble and BCJ-L2 record, used for validation, are plotted in Figure 4.4 together with the target spectrum.

4.4. Optimization process and results

Upon the successful implementation and validation of the suggested response evaluation pipeline, the optimization framework detailed in Subsection 3.3 is applied to achieve the optimal design of TVMDs for the benchmark structure. This section examines two distinct case studies: a pure displacement control problem, which exclusively optimizes the maximum inter-story drift angle, and a balanced displacement and acceleration control problem. The latter case aims to simultaneously suppress the maximum inter-story drift below the designated performance limit and minimize structural acceleration. In both cases, the optimization yields not only the optimal damper parameters, but also the optimal distribution and grouping into damper typologies to make the design more practical.

4.4.1. Implementation of the optimization algorithm

The proposed optimization framework is implemented by solving the dual problem formulation derived in Subsection 3.3 via a primal-dual gradient approach. To execute this scheme, the algorithmic pipeline is built within the PyTorch ecosystem utilizing Cooper [56], an open-source Python library designed for solving non-convex continuous constrained optimization problems. During each optimization step, the framework simultaneously updates the design parameters (primal variables) to minimize the objective function and the Lagrange multipliers (dual variables) to maximize constraint fulfillment. Both the primal and dual problems are optimized using independent instances of the root mean square propagation (RMSProp) [57] solver. RMSProp is an adaptive optimization algorithm that dynamically scales the learning rate for each individual parameter by dividing it by an exponentially decaying average of the parameter's recent squared gradients. This mechanism ensures that parameters with large, volatile gradients experience smaller, controlled step sizes, while parameters with sparse or small gradients receive larger steps to accelerate convergence. This adaptive scaling prevents gradient vanishing or explosion, ensuring stable convergence behavior when navigating the complex minimax surface of a primal-dual formulation we are dealing with in the current study. In order to further enhance the convergence, the scale vector $\boldsymbol{\Xi}$ described in Equation (44) is calculated at the beginning of the optimization based on the initial unweighted values of the the objective terms. The optimizer properties are listed in Table 4.6.

Table 4.6: Properties of the Primal and Dual Solvers

Property	Primal	Dual
Algorithm	RMSProp [57]	RMSProp
Method	Minimization	Maximization
Initial learning rate, lr_{init}	10^{-5}	10^{-5}
Momentum	0.40	0.40
Decay	0.00	0.00
Target parameters	$\bar{\mathbf{X}}$	$\boldsymbol{\lambda}$

4.4.2. Displacement control

For the pure displacement control case, the weight coefficient θ_1 is set to 1, and θ_2 is set to 0, as described in Subsection 2.2.2, and the inertance limit is set as $m_{r,lim} = 10000\,ton$. In this way, the optimization starts with the first step of the algorithm by setting the weight vector $\mathbf{\Theta} = \{1, 0, 0\}^T$, and optimization is conducted until the inter-storey drifts performance threshold $\delta_{lim} = 0.01$ rad is reached. Optimization history is plotted in Figure 4.5, and as a result we get a distribution with unique dampers in each storey. To make the design practical, in this study TVMDs will be grouped into $p = 3$ damper typologies. To achieve that without the loss in structural performance, optimization moves to the second step. At first, Equation (41) is used to initialize the cluster (group) centers, and then the fine-tuning optimization is conducted with $\mathbf{\Theta} = \{1, 0, 1\}^T$ until structural performance is restored. The optimization history of the second step is depicted in Figure 4.6. The optimal distribution and achieved structural performance is showed in Figure 4.6, with damper parameters and grouping summarized in Table 4.7. It can be noted, that algorithm was able to identify the most critical excitations from the ensemble and fit very closely to the corresponding responses, which is achieved by efficiently calculating all the response histories in parallel at every step.

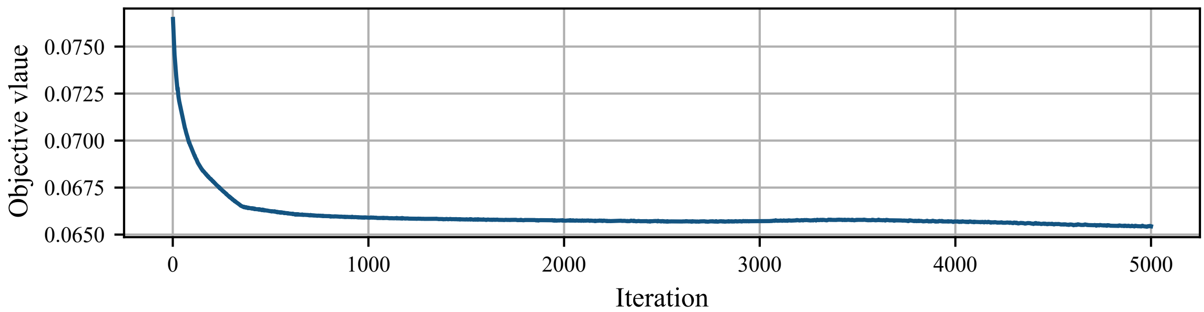


Figure 4.5: Optimization history for the displacement control case, optimization step 1 (without clustering loss)

4.4.3. Simulteneous displacement and acceleration control

For the balanced displacement and acceleration control problem, the procedures described in the previous section are repeated, with the value of θ_2 changing its value from 0 to 1, and the increased upper inertance limit $m_{r,lim} = 15000$ ton. As in the previous case, for the both steps, optimization was conducted until the maximum inter-storey drifts value is less than the performance threshold $\delta_{lim} = 0.01$ rad. The optimization history for the first and the second optimization steps are shown in Figures 4.8 and 4.9, respectively. The final TVMD distribution and performance is shown in Figure 4.10, and optimal damper parameters and grouping are summarized in Table 4.8. A significant increase in the allocated inertance at the top storey in comparison to the pure displacement control case can be noted. This increase is attributed to significantly reduced acceleration response of the top floor.

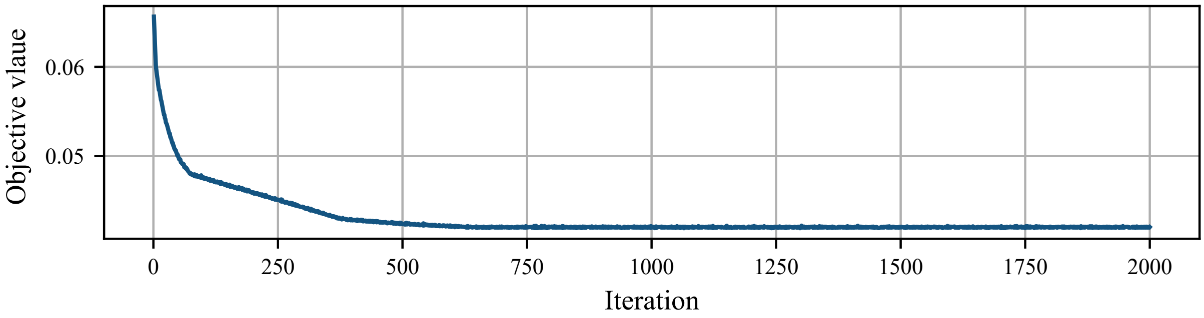
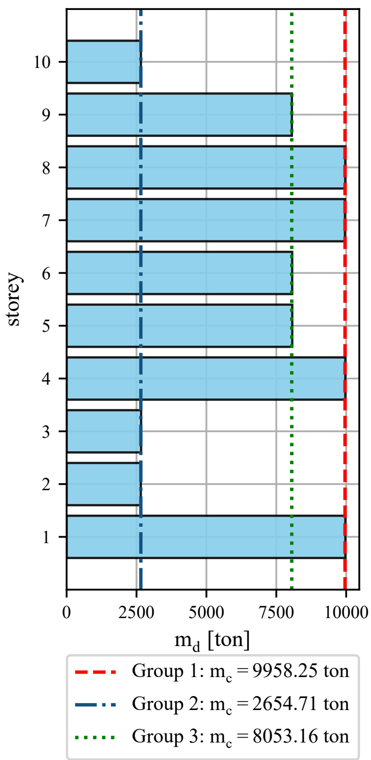
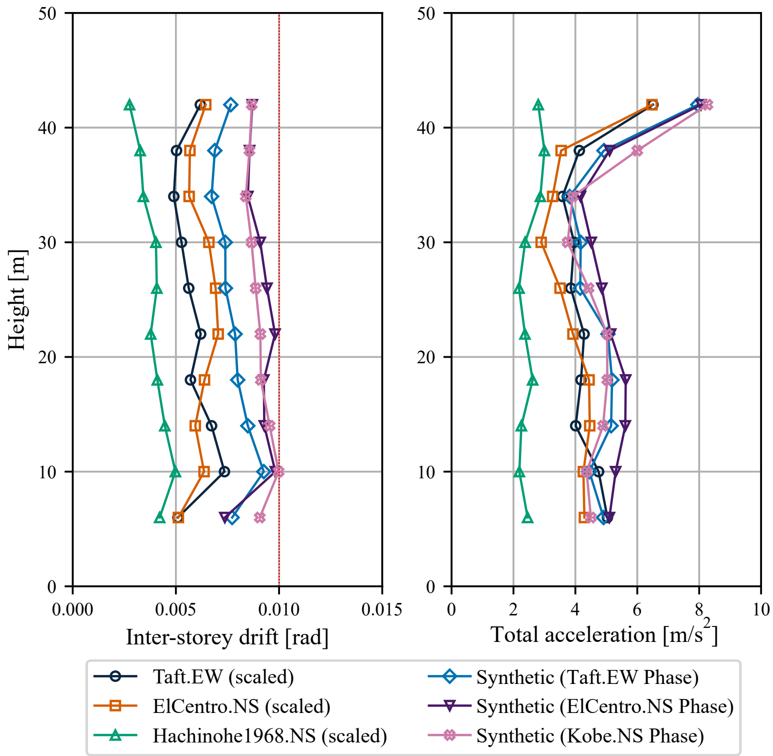


Figure 4.6: Optimization history for the displacement control case, optimization step 2 (with clustering loss)



(a) Optimal inertance distribution



(b) Height-wise maximum inter-storey drift and total absolute acceleration distributions

Figure 4.7: Results of the displacement control optimization using the proposed algorithm. Achieved performance: maximum inter-storey drift angle $\delta_p^{max} = 0.01 rad$, maximum absolute total acceleration $\ddot{u}_p^{max} = 8.27 m/s^2$

5. Conclusion and future studies

This study introduced original formulation of multi-objective optimal control problem for simultaneous parameter and topology optimization of multiple-degree-of-freedom (MDOF) structures equipped with tuned viscous mass dampers (TVMDs). Furthermore, to overcome the limitations of current optimization approaches and propose an optimization algorithm for the above-stated problem, a novel, high-performance computational framework was successfully developed and validated. By merging latest developments in structural control of inerter-based vibration control systems (IVCSs) with modern differentiable programming paradigms, this research addresses a long-standing bottleneck in time-domain optimal design. Based on the theoretical derivations, algorithmic implementation, and

Table 4.7: Optimal TVMD parameters distribution and grouping for the pure displacement control case

Group	Storey	m_r [ton]	k_b [kN/m]	c_d [kN · s/m]
1	1 4 7 8	9958	131027	17080
2	2 3 10	2655	34930	4553
3	5 6 9	8053	105961	13812

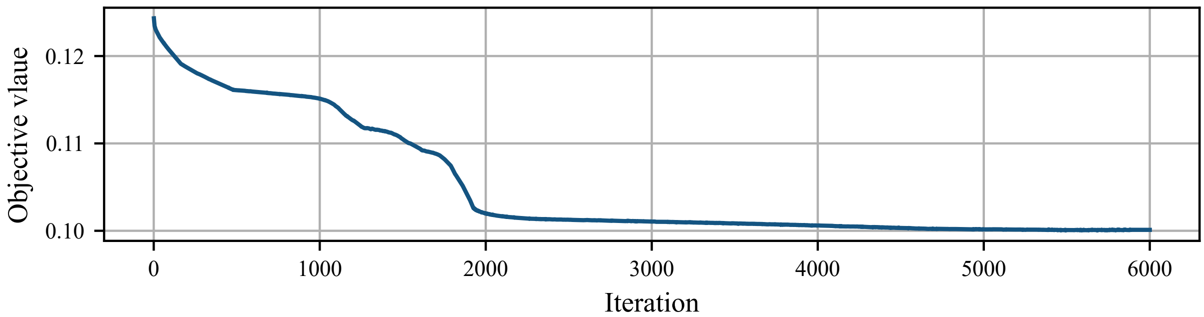


Figure 4.8: Optimization history for the displacement and acceleration control case, optimization step 1 (without clustering loss)

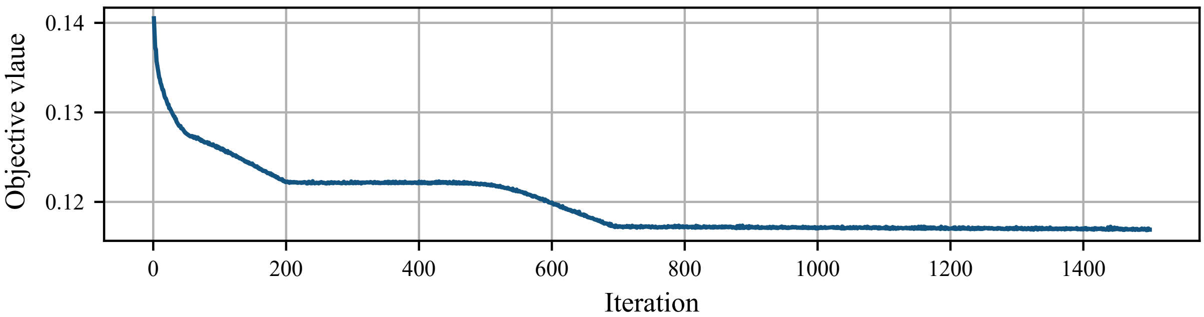
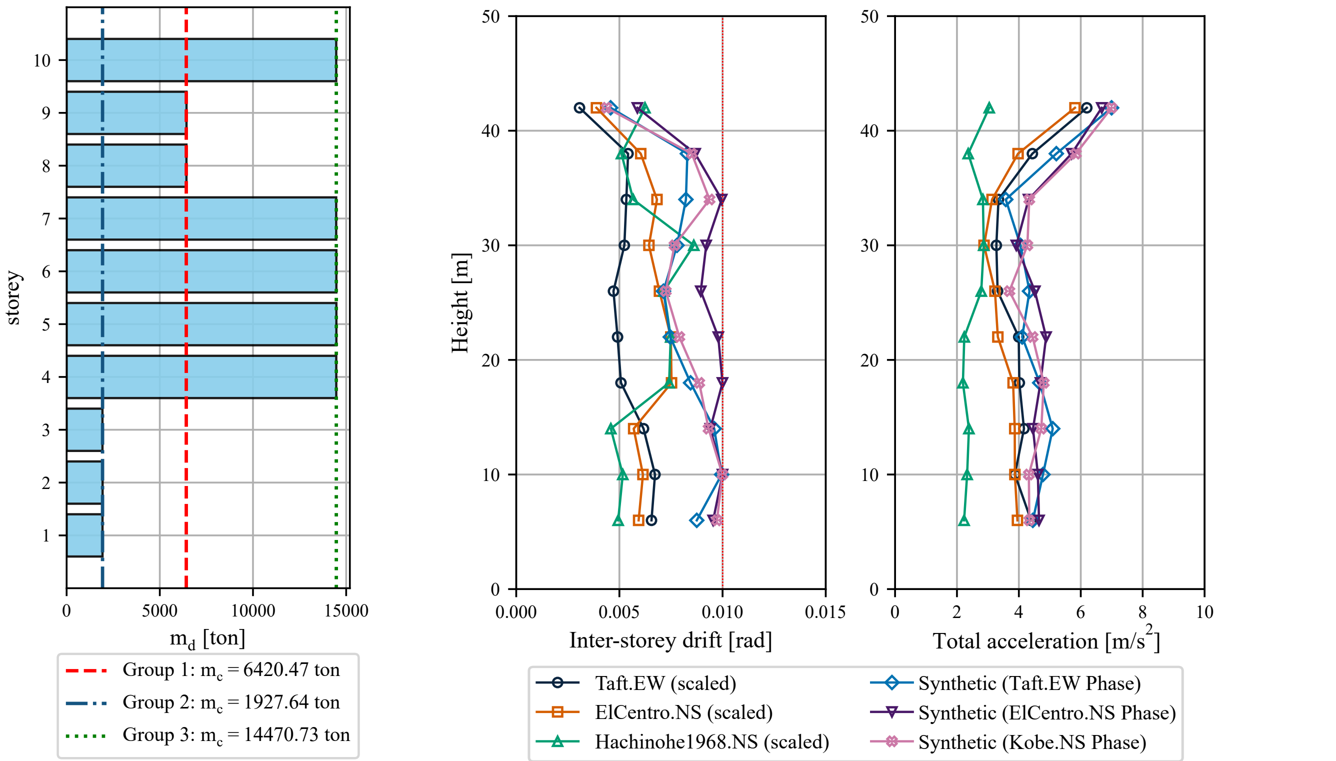


Figure 4.9: Optimization history for the displacement and acceleration control case, optimization step 2 (with clustering loss)

numerical verification on a ten-story benchmark shear building, suggested optimization framework was able to not only successfully identify the optimal parameter value, but also topology and grouping of the devices. The PyTorch-based implementation achieved stable and efficient convergence on standard consumer-grade computing hardware, demonstrating its practical viability for practicing structural engineers with minimal coding experience.

In summary, the proposed framework provides structural engineers with an accessible and robust tool that significantly shortens the iterative design cycle. By removing the need for hand-derived analytical sensitivities or expensive numerical approximations, this work leverages advances in mainstream deep learning to achieve economic structural engineering design.

However, the proposed approach does not account for critical engineering constraints such as peak damper force



(a) Optimal inertance distribution

(b) Height-wise maximum inter-storey drift and total absolute acceleration distributions

Figure 4.10: Results of the simultaneous displacement and acceleration control optimization using the proposed algorithm. Achieved performance: maximum inter-storey drift angle $\delta_p^{max} = 0.01rad$, maximum absolute total acceleration $\ddot{u}_p^{max} = 6.99m/s^2$

Table 4.8: Optimal TVMD parameters distribution and grouping for the simultaneous displacement and acceleration control case

Group	Storey	m_r [ton]	k_b [kN/m]	c_d [kN · s/m]
1	5 6 9	6420	86282	10907
2	2 3 10	1928	25905	3275
3	1 4 7 8	14471	194466	24582

or total damper inertance. Because these factors significantly influence connection design details, they directly govern the practical applicability of the resulting optimal design. Addressing these limitations will be the focus of future studies, which will evaluate more comprehensive and practical structural design scenarios.

CRediT authorship contribution statement

Vitali Kakouka: Writing - original draft, Conceptualization, Investigation, Validation, Visualization, Methodology, Software. **Kohju Ikago**: Supervision, Methodology, Writing - review & editing,

Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

During the preparation of this work, the author(s) used generative AI technologies, such as ChatGPT and Gemini, solely for the purpose of editing, grammar correction, and improving the quality of the English language. After using these tools, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the final contents of the publication. No generative AI tools were used to create any figures or tables in this paper.

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.

Acknowledgements

The first author expresses their sincere gratitude to the Ministry of Education, Culture, Sports, Science and Technology (MEXT) of Japan for providing the educational scholarship.

Data availability statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

[1] A. Chopra, Dynamics of Structures, Prentice-Hall international series in civil engineering and engineering mechanics, Pearson Education, 2007.

[2] C. Christopoulos, A. Filiatrault, Principles of Passive Supplemental Damping and Seismic Isolation, IUSS Press, 2006.

[3] G. W. Housner, L. A. Bergman, T. K. Caughey, A. G. Chassiakos, R. O. Claus, S. F. Masri, R. E. Skelton, T. T. Soong, B. F. Spencer, J. T. P. Yao, Structural control: Past, present, and future, Journal of Engineering Mechanics 123 (9) (1997) 897–971. doi:10.1061/(ASCE)0733-9399(1997)123:9(897).

[4] J. Kim, J. Ryu, L. Chung, Seismic performance of structures connected by viscoelastic dampers, Engineering Structures 28 (2) (2006) 183–195. doi:https://doi.org/10.1016/j.engstruct.2005.05.014.

[5] Y. Watanabe, K. Ikago, N. Inoue, H. Kida, S. Nakaminami, H. Tanaka, Y. Sugimura, K. Saito, Full-scale dynamic tests and analytical verification of a force-restricted tuned viscous mass damper, in: 15th World conference on earthquake engineering. Lisbon, Portugal, 2012.

[6] I. F. Lazar, S. Neild, D. Wagg, Using an inerter-based device for structural vibration suppression, Earthquake Engineering & Structural Dynamics 43 (8) (2014) 1129–1147. doi:https://doi.org/10.1002/eqe.2390.

[7] D. De Domenico, G. Ricciardi, I. Takewaki, Design strategies of viscous dampers for seismic protection of building structures: A review, Soil Dynamics and Earthquake Engineering 118 (2019) 144–165. doi:https://doi.org/10.1016/j.soildyn.2018.12.024.

[8] A. Keivan, B. M. Phillips, M. Ikenaga, K. Ikago, Causal realization of rate-independent linear damping for the protection of low-frequency structures, Journal of Engineering Mechanics 143 (9) (2017) 04017058. doi:10.1061/(ASCE)EM.1943-7889.0001259.

[9] H. Luo, K. Ikago, C. Chong, A. Keivan, B. Phillips, Performance of low-frequency structures incorporated with rate-independent linear damping, Engineering Structures 181 (2019) 324–335. doi:https://doi.org/10.1016/j.engstruct.2018.12.022.

[10] K. Ikago, Y. Sugimura, K. Saito, N. Inoue, Seismic displacement control of multiple-degree-of-freedom structures using tuned viscous mass dampers, in: G. Lombaert, G. Muller, G. {De Roeck}, G. Degrande (Eds.), Proceedings of the 8th International Conference on Structural Dynamics, EURODYN 2011, Proceedings of the 8th International Conference on Structural Dynamics, EURODYN 2011, University of Southampton, Institute of Sound Vibration and Research, 2011, pp. 1800–1807, 8th International Conference on Structural Dynamics, EURODYN 2011 ; Conference date: 04-07-2011 Through 06-07-2011.

[11] K. Ikago, K. Saito, N. Inoue, Seismic control of single-degree-of-freedom structure using tuned viscous mass damper, Earthquake Engineering & Structural Dynamics 41 (3) (2012) 453–474. doi:https://doi.org/10.1002/eqe.1138.

[12] C. Pan, R. Zhang, Design of structure with inerter system based on stochastic response mitigation ratio, Structural Control and Health Monitoring 25 (6) (2018) e2169. doi:https://doi.org/10.1002/stc.2169.

[13] R. Xie, K. Ikago, Device topology optimization for an inerter-based structural dynamic vibration absorber, Frontiers in Built Environment Volume 10 - 2024 (2024). doi:10.3389/fbuil.2024.1508190.

[14] J. Kang, Z. Zhao, L. Xie, C. Wan, S. Xue, Proportional distribution pattern and modal principle of tuned viscous mass dampers for multi-degree-of-freedom structures, Engineering Structures 322 (2025) 119175. doi:https://doi.org/10.1016/j.engstruct.2024.119175.

[15] M. Smith, Synthesis of mechanical networks: the inerter, IEEE Transactions on Automatic Control 47 (10) (2002) 1648–1662. doi:10.1109/TAC.2002.803532.

[16] T. Arakaki, H. Kuroda, F. Arima, Y. Inoue, K. Baba, Development of seismic devices applied to ball screw : Part 1 basic performance test of rd-series, AIJ Journal of Technology and Design 5 (8) (1999) 239–244. doi:10.3130/aijt.5.239_1.

[17] Y. Nakamura, H. Watanabe, S. Kawamata, Seismic response control of structures by accelerated liquid mass damper, in: Proceedings of the 9th World Conference on Earthquake Engineering (9WCEE), Vol. V, Tokyo-Kyoto, Japan, 1988, pp. 785–790.

[18] Y. Nakamura, A. Fukukita, K. Tamura, I. Yamazaki, T. Matsuoka, K. Hiramoto, K. Sunakoda, Seismic response control using electromagnetic inertial mass dampers, Earthquake Engineering & Structural Dynamics 43 (4) (2014) 507–527. doi:https://doi.org/10.1002/eqe.2355.

[19] D. Li, K. Ikago, A. Yin, Structural dynamic vibration absorber using a tuned inerter eddy current damper, Mechanical Systems and Signal Processing 186 (2023) 109915. doi:https://doi.org/10.1016/j.ymssp.2022.109915.

[20] J. Ormondroyd, J. P. Den Hartog, The theory of the dynamic vibration absorber, Transactions of the American Society of Mechanical Engineers 49-50 (2) (2023) 021007. doi:10.1115/1.4058553.

[21] Y. Sugimura, K. Ikago, K. Saito, N. Inoue, Response control effect of steel building structure using tuned viscous mass damper, in: Proceedings of the 15th World Conference on Earthquake Engineering (15WCEE), Lisbon, Portugal, 2012, p. Paper ID 0138, accessed via NICEE WCEE Archive.

[22] L. Marian, A. Giaralis, Optimal design of a novel tuned mass-damper–inerter (tmdi) passive vibration control configuration for stochastically support-excited structural systems, Probabilistic Engineering Mechanics 38 (2014) 156–164. doi:https://doi.org/10.1016/j.probengmech.2014.03.007.

[23] R. Zhang, Z. Zhao, C. Pan, K. Ikago, S. Xue, Damping enhancement principle of inerter system, Structural Control and Health Monitoring (2020) e2523doi:DOI:10.1002/stc.2523.

[24] J. Kang, K. Ikago, Seismic control of multidegree-of-freedom structures using a concentratedly arranged tuned viscous mass damper, Earthquake Engineering & Structural Dynamics 52 (14) (2023) 4708–4732. doi:https://doi.org/10.1002/eqe.3977.

[25] Z. Chen, K. Junya, I. Masahiro, I. Kohju, I. Norio, Viscoelastically supported viscous mass damper incorporated into a seismic isolation system, Journal of Earthquake and Tsunami 10 (03) (2016) 1640009. doi:10.1142/S1793431116400091.

[26] A. Taflanidis, A. Giaralis, D. Patsialis, Multi-objective optimal design of inerter-based vibration absorbers for earthquake protection of multi-storey building structures, Journal of the Franklin Institute 356 (14) (2019) 7754–7784, special Issue on Inerter-based Systems.

[27] P. T. Boggs, J. W. Tolle, Sequential quadratic programming, Acta Numerica 4 (1995) 1–51. doi:10.1017/S0962492900002518.

[28] D. Kraft, A Software Package for Sequential Quadratic Programming, Deutsche Forschungs- und Versuchsanstalt für Luft- und Raumfahrt Köln: Forschungsbericht, Wiss. Berichtswesen d. DFVLR, 1988.

[29] R. H. Byrd, P. Lu, J. Nocedal, C. Zhu, A limited memory algorithm for bound constrained optimization, SIAM Journal on Scientific Computing 16 (5) (1995) 1190–1208. doi:10.1137/0916069.

[30] A. Griewank, A. Walther, Evaluating Derivatives: Principles and Techniques of Algorithmic Differentiation, 2nd Edition, Society for Industrial and Applied Mathematics, Philadelphia, PA, 2008. doi:10.1137/1.9780898717761.

[31] H.-C. Fischer, Automatic differentiation and applications, in: E. Adams, U. Kulisch (Eds.), Scientific Computing with Automatic Result Verification, Vol. 189 of Mathematics in Science and Engineering, Elsevier, 1993, pp. 105–142. doi:https://doi.org/10.1016/S0076-5392(08)62844-1.

[32] A. G. Baydin, B. A. Pearlmutter, A. Radul, J. M. Siskind, Automatic differentiation in machine learning: a survey, J. Mach. Learn. Res. 18 (2015) 153:1–153:43.

[33] S. S. Schoenholz, E. D. Cubuk, Jax, m.d. a framework for differentiable physics*, Journal of Statistical Mechanics: Theory and Experiment 2021 (12) (2021) 124016. doi:10.1088/1742-5468/ac3ae9.

[34] D. A. Bezgin, A. B. Buhendwa, N. A. Adams, Jax-fluids: A fully-differentiable high-order computational fluid dynamics solver for compressible two-phase flows, Computer Physics Communications 282 (2023) 108527. doi:https://doi.org/10.1016/j.cpc.2022.108527.

[35] I. Ozaki, F. Kimura, M. Berz, Higher-order sensitivity analysis of finite element method by automatic differentiation, Computational Mechanics 16 (4) (1995) 223–234. doi:10.1007/BF00369867.

[36] L. F. R. Espath, R. V. Linn, A. M. Awruch, Shape optimization of shell structures based on nurbs description using automatic differentiation, International Journal for Numerical Methods in Engineering 88 (7) (2011) 613–636. doi:https://doi.org/10.1002/nme.3183.

[37] S. A. Nørgaard, M. Sagebaum, N. R. Gauger, B. S. Lazarov, Applications of automatic differentiation in topology optimization, Structural and Multidisciplinary Optimization 56 (5) (2017) 1135–1146. doi:10.1007/s00158-017-1708-2.

[38] G. Wu, A framework for structural shape optimization based on automatic differentiation, the adjoint method and accelerated linear algebra, Structural and Multidisciplinary Optimization 66 (7) (Jun. 2023). doi:10.1007/s00158-023-03601-0.

[39] J. Bradbury, R. Frostig, P. Hawkins, M. J. Johnson, C. Leary, D. Maclaurin, G. Necula, A. Paszke, J. VanderPlas, S. Wanderman-Milne, Q. Zhang, JAX: composable transformations of Python+NumPy programs (2018).

[40] F. van Keulen, R. Haftka, N. Kim, Review of options for structural design sensitivity analysis. part 1: Linear systems, Computer Methods in Applied Mechanics and Engineering 194 (30) (2005) 3213–3243, structural and Design Optimization. doi:https://doi.org/10.1016/j.cma.2005.02.002.

[41] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Köpf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, S. Chintala, PyTorch: an imperative style, high-performance deep learning library, Curran Associates Inc., Red Hook, NY, USA, 2019.

[42] J. W. Cooley, J. W. Tukey, An algorithm for the machine calculation of complex fourier series, Mathematics of Computation 19 (90) (1965) 297–301.

[43] X. Jin, J. Han, K-Means Clustering, Springer US, Boston, MA, 2010, pp. 563–564. doi:10.1007/978-0-387-30164-8_425.

[44] R. Zhang, Y. Liu, H. Sun, Physics-guided convolutional neural network (phycnn) for data-driven seismic response modeling, Engineering Structures 215 (2020) 110704. doi:https://doi.org/10.1016/j.engstruct.2020.110704.

[45] R. Zhang, Y. Liu, H. Sun, Physics-informed multi-lstm networks for metamodeling of nonlinear structures, Computer Methods in Applied Mechanics and Engineering 369 (2020) 113226. doi:https://doi.org/10.1016/j.cma.2020.113226.

[46] Y.-T. Chou, W.-T. Chang, J. G. Jean, K.-H. Chang, Y.-N. Huang, C.-S. Chen, Structgnn: An efficient graph neural network framework for static structural analysis, Comput. Struct. 299 (C) (Aug. 2024). doi:10.1016/j.compstruc.2024.107385.

[47] Y.-T. Chou, P.-C. Kuo, K.-Y. Li, W.-T. Chang, Y.-N. Huang, C.-S. Chen, Inductive graph-based long short-term memory network for the prediction of nonlinear floor responses and member forces of steel buildings subjected to orthogonal horizontal ground motions, Earthquake Engineering & Structural Dynamics 54 (2) (2025) 491–507. doi:https://doi.org/10.1002/eqe.4264.

[48] H. Peng, W. Zheng, Y. Yu, Y. Luo, Rapid seismic response prediction of steel frames based on graph attention network (gat) with mamba blocks, Structures 79 (2025) 109390. doi:https://doi.org/10.1016/j.istruc.2025.109390.

[49] R. E. Wengert, A simple automatic derivative evaluation program, Commun. ACM 7 (8) (1964) 463–464. doi:10.1145/355586.364791.

[50] C. Dugas, Y. Bengio, F. Bélisle, C. Nadeau, R. Garcia, Incorporating second-order functional knowledge for better option pricing, in: T. Leen, T. Dietterich, V. Tresp (Eds.), Advances in Neural Information Processing Systems, Vol. 13, MIT Press, 2000.

[51] Japan Society of Seismic Isolation, Tokyo, Japan, Manual for design and fabrication for passive vibration control of structures [in Japanese], 3rd Edition (2013).

[52] F. McKenna, Opensees: A framework for earthquake engineering simulation, Computing in Science & Engineering 13 (4) (2011) 58–66. doi:10.1109/MCSE.2011.66.

[53] M. Zhu, F. McKenna, M. H. Scott, Openseespy: Python library for the opensees finite element framework, SoftwareX 7 (2018) 6–11. doi:https://doi.org/10.1016/j.softx.2017.10.009.

[54] B. C. of Japan, Guidelines for Generating Simulated Earthquake Ground Motions used for Aseismic Design, Building Center of Japan, 1992.

[55] E. C. for "Structural Design Practice of Super High-Rise Buildings", Structural Design Practice of Super High-Rise Buildings based on Performance Evaluation: A Compendium of Studies and Judgments by Practicing Structural Designers of Super High-Rise and Base-Isolated Buildings [in Japanese], The Building Center of Japan (BCJ), Tokyo, Japan, 2019.

[56] J. Gallego-Posada, J. Ramirez, M. Hashemizadeh, S. Lacoste-Julien, Cooper: A Library for Constrained Optimization in Deep Learning (2025). arXiv:2504.01212.

[57] S. Ruder, An overview of gradient descent optimization algorithms (2017). arXiv:1609.04747.