Electronic Research Archive

https://www.aimspress.com/journal/era

ERA, 34(1): 433–462.
DOI: 10.3934/era.2026021
Received: 27 October 2025
Revised: 22 December 2025
Accepted: 04 January 2026
Published: 15 January 2026

Research article

Model order reduction using dual ν-gap metrics: A multi-objective optimization approach

Mingyu Kim[1], Sukyung Seo[1], Suhwan Choi[1], Yeongjae Kim[1], Yeongmi Kim[2,*] and Tae-Hyoung Kim[1,*]

[1] Department of Mechanical Engineering, College of Engineering, Chung-Ang University, Seoul 06974, Republic of Korea

[2] Department of Medical Technologies, MCI The Entrepreneurial School, Innsbruck 6020, Austria

* **Correspondence:** Email: yeongmi.kim@mci.edu, kimth@cau.ac.kr.

Abstract: This study investigated the problem of model order reduction by employing the ν-gap metric and introduced a metaheuristic optimization framework to address its inherent nonconvexity. The ν-gap metric quantifies the closed-loop distance between two systems, making it useful for generating low-order models that preserve closed-loop stability characteristics. However, minimizing the conventional ν-gap alone may result in poor dynamic fidelity. To address this limitation, a modified ν-gap metric based on frequency response matching was developed to explicitly capture frequency-wise discrepancies associated with time-domain behavior. By jointly considering the conventional and modified ν-gap metrics, a dual ν-gap–based multi-objective model reduction framework was formulated, in which stability-related and time-domain fidelity objectives are treated in a complementary manner. The resulting optimization problem is highly nonconvex and constrained due to the winding number condition. A multi-objective particle swarm optimization framework was therefore employed as a numerical tool to generate Pareto-optimal reduced-order models. Benchmark studies on large-scale systems demonstrated that the proposed framework enables a systematic exploration of trade-offs between stability preservation and time-domain fidelity, yielding practically meaningful reduced-order models with tunable performance characteristics.

Keywords: model order reduction; ν-gap metric; parameter identification; particle swarm optimization; multi-objective optimization

1. Introduction

In the engineering field, the model order reduction (MOR) technique has proven to be a valuable approach addressing the challenges posed by complexities of large-scale systems [1]. This approach

simplifies analysis and control design tasks by providing a low-dimensional model that preserves the essential characteristics of the original system, making it more manageable for analysis, design, and implementation [2, 3]. Standard measures of model fidelity, represented by the approximation error between the original and reduced-order models, are commonly expressed using H_∞ or H_2 norms [4–6]. These metrics not only guide MOR effectiveness assessment but also contribute to the development of diverse reduction techniques within the field of engineering, including truncation methods [7] and Krylov-interpolation-based methods [8]. These MOR methods have advanced across various engineering disciplines. Nevertheless, each method exhibits unique strengths and limitations as per the application context. In particular, the truncation method stands out for its ability to provide global error bounds while maintaining system stability. However, its computational complexity often proves excessively costly, particularly when dealing with large-scale systems. By contrast, the Krylov-interpolation-based method proficiently handles large dimensions in deriving the low-dimensional model but frequently fails to ensure asymptotic stability and lacks efficiently computable global error bounds [9]. Notably, the aforementioned MOR techniques, which rely on H_2/H_∞ metrics, quantify the distance between the original model and its reduced representation (i.e., the approximation error) within an open-loop setting. However, these metrics generally fail to capture discrepancies relevant to closed-loop behavior.

The gap and graph metrics were developed as closed-loop–oriented measures that are more suitable than conventional H_2/H_∞-type criteria when the objective is to compare systems in terms of their closed-loop behavior under perturbations and performance variations [10–12]. Within this line of research, the ν-gap metric has been established as a practically computable distance measure that captures closed-loop stability-related characteristics in a less conservative and computationally simpler manner than alternative closed-loop metrics [13]. Accordingly, several ν-gap–based model reduction methods have been proposed using state-space representations together with iterative linear matrix inequality (LMI)-type procedures [14–16]. While these approaches provide systematic optimization procedures, their performance can be sensitive to the feasibility of intermediate LMI steps and to the choice of initial candidates, particularly when the feasible set is narrow due to topological constraints. To mitigate initialization sensitivity, a homotopy-based parameter search scheme was also introduced to guide the search toward an appropriate reduced model [17, 18]; however, such continuation-based strategies still depend on a carefully selected search path and may become restrictive when the admissible solution region is highly nonconvex. More recently, a semidefinite programming (SDP)-based approach using frequency-domain data was presented in [19], which reformulates the original nonconvexity into a quasi-convex form by introducing a central transfer function so as to account for the winding number condition (WNC). Despite this advantage, practical difficulties remain, including a restricted search space and the challenge of selecting a suitable initial solution set. In contrast, [20] formulated the ν-gap model reduction problem as a bilinear matrix inequality (BMI) and employed SDP optimization; however, this formulation does not explicitly enforce the WNC associated with the Nyquist-based interpretation of the ν-gap metric, and its solution quality may heavily depend on the initial set of candidate solutions.

The present study investigates two methodological approaches to MOR, both framed as optimization problems. In both approaches, a modified ν-gap functions as a pivotal metric for assessing model fidelity, providing a standardized evaluation criterion. In MOR, minimizing the response error between lower- and higher-order models is crucial, as it signifies the precision of the

low-order model approximation to the original model. However, relying solely on the conventional ν-gap metric for MOR may cause unexpected response errors between the original and reduced-order models. This phenomenon arises from the explicit limitation of the conventional ν-gap metric, which only provides worst-case bound-related information on the distance between the original and reduced-order models within the closed-loop framework. To address this challenge, the present study proposes a modified ν-gap metric. This modified metric leverages a dataset spanning a wider frequency domain to enhance the response-matching-related characteristics between the original and reduced-order models, thereby mitigating discrepancies and improving overall accuracy. Subsequently, our study utilizes this modified ν-gap metric to develop MOR techniques from two distinct perspectives: 1) The identification of a reduced-order model through such techniques is formulated as a set of optimization problems. 2) These problems are then tackled using a variant of the metaheuristic particle swarm optimization (PSO), which is effective in solving a wide range of nonconvex optimization problems.

In the first technique, the MOR issue is formulated as a single-objective optimization problem. Here, the proposed modified ν-gap metric acts as the cost function to be minimized, while the WNC serves as a crucial constraint condition during the exploration for available parameters of the reduced-order model. To deal with this single-objective optimization problem effectively and guarantee the WNC constraint, a metaheuristic solver is employed. This solver, primarily derived from a modification of the PSO algorithm featuring a distributed cyclic neighborhood search mechanism [21], is designed to directly identify the unknown model parameters while guaranteeing the WNC constraint. The challenges arising from the inability to integrate the WNC [20] and the limitations of the search space caused by the indirect resolution of the WNC [19] can be effectively managed by the proposed metaheuristic solver because of its capability to directly handle the nonconvex WNC.

In the second technique, the MOR problem is addressed by formulating a multi-objective optimization framework. The conventional and modified ν-gap metrics provide complementary perspectives in the model reduction process. Specifically, the modified ν-gap metric emphasizes response matching between the original and reduced-order models, whereas the conventional ν-gap metric characterizes the weakest topology that preserves closed-loop stability [22]. However, minimizing the conventional ν-gap alone does not directly account for response-matching accuracy, and achieving a small modified ν-gap value does not necessarily guarantee a small conventional ν-gap value. This intrinsic discrepancy indicates a trade-off between stability-related and response-matching objectives, which cannot be adequately captured within a single-objective formulation. Despite this observation, existing ν-gap–based MOR studies have predominantly relied on single-objective optimization frameworks [19, 22, 23]. While multi-objective optimization has been widely studied in a general context to handle competing objectives [24, 25], its application to ν-gap–based model reduction problems involving topological constraints such as the WNC has not been systematically explored, to the best of the authors' knowledge. Motivated by this gap, the present study formulates the MOR problem as a constrained multi-objective optimization task, treating the conventional and modified ν-gap metrics as independent objective functions. This formulation enables the generation of a Pareto-optimal set of reduced-order models, representing a family of candidate solutions that explicitly captures the trade-off between closed-loop stability preservation and response-matching accuracy. The resulting problem is solved using a multi-objective particle swarm optimization

(MOPSO) algorithm as a numerical tool for generating such Pareto-optimal reduced-order models.

The remainder of this paper is organized as follows. Section 2 provides a comprehensive review of the conventional ν-gap metric and its correlation with the stability margin concept; then, it introduces a modified version that better captures frequency-domain discrepancies. Section 3 presents the formulation of the MOR problem as a single-objective optimization with a WNC constraint and then outlines how such a nonconvex NP-hard optimization problem is solved using a metaheuristic solver. Numerical simulations are presented to evaluate the effectiveness of the proposed method compared with existing ν-gap-based approaches. Section 4 describes the motivation behind the need for multi-objective optimization in model reduction, formulates the corresponding problem, and explains how the MOPSO algorithm is applied to derive a Pareto-optimal set of reduced-order models. Finally, Section 5 concludes the paper.

2. The ν-gap metric framework for reduced-order modeling

2.1. The ν-gap metric and stability margin: Preliminaries and key properties

The ν-gap metric is utilized to evaluate the similarity between two dynamical systems when both are integrated within a feedback loop alongside an identical controller. In this work, finite-dimensional, continuous-time, real-rational linear time-invariant (LTI) transfer functions, denoted as $P_1(s)$ and $P_2(s)$, are considered. These models are assumed to be proper (or strictly proper) with minimal state-space realizations, and are either internally stable or stabilizable under the unity-feedback configuration. These standard assumptions ensure the existence of normalized coprime factorizations, which are fundamental to defining the ν-gap metric. Accordingly, a normalized right coprime factorization of P_i is given by $P_i = N_i M_i^{-1}$, while a normalized left coprime factorization of P_i is expressed as $P_i = \tilde{M}_i^{-1}\tilde{N}_i$, where $\tilde{N}_i$ and $\tilde{M}_i$ denote the adjoints of the linear operators N_i and M_i, respectively. Additionally, let $G_i = \begin{bmatrix} N_i \\ M_i \end{bmatrix}$ be the normalized right graph symbol of P_i and $\tilde{G}_i = [-\tilde{M}_i \ \tilde{N}_i]$ be the normalized left graph symbol of P_i. For clarity and ease of reference, Table 1 summarizes the main mathematical symbols and notations used throughout the paper together with their definitions. The computation of the ν-gap metric comprises two steps: verifying a WNC and calculating a frequency-domain norm. First, the WNC is defined as follows: for $\forall \omega \in \mathbb{R}$,

$$\det\left((G_2^* G_1)(j\omega)\right) \neq 0 \text{ and } \operatorname{wno}\left(\det(G_2^* G_1)\right) = 0, \tag{2.1}$$

where $*$ denotes the complex conjugate transpose as $G^*(s) := G^T(-s)$, and $\operatorname{wno}(g(s))$ represents the winding number about the origin of $g(s)$ when s follows the standard Nyquist D-contour. Second, the ν-gap between $P_1(s)$ and $P_2(s)$ is defined as follows:

$$\delta_\nu(P_1, P_2) = \begin{cases} \|\tilde{G}_2 G_1\|_\infty, & \text{if the WNC in (2.1) holds,} \\ 1, & \text{otherwise,} \end{cases} \tag{2.2}$$

where $\|\cdot\|_\infty$ denotes the H_∞-norm of a dynamical system. Here, checking the WNC is based on the normalized right coprime factorizations $P_1 = N_1 M_1^{-1}$ and $P_2 = N_2 M_2^{-1}$. Defining the graph symbols as $G_1 = \begin{bmatrix} N_1 \\ M_1 \end{bmatrix}$ and $G_2^* = [N_2^* \ M_2^*]$ yields $G_2^* G_1 = N_2^* N_1 + M_2^* M_1$, which implies checking whether this satisfies the condition in (2.1). Furthermore, note that $0 \leq \delta_\nu(P_1, P_2) \leq 1$, and $\delta_\nu(P_1, P_2) = 0$ if and

only if $P_1(s) = P_2(s)$. By contrast, the stability margin, $b_{P,C}$, for the closed-loop system involving the plant $P(s)$ and the controller $C(s)$ is defined as follows:

$$b_{P,C} = \begin{cases} \left\| \begin{bmatrix} I \\ C \end{bmatrix} (I+PC)^{-1} \begin{bmatrix} I & P \end{bmatrix} \right\|_\infty^{-1}, & \text{if } C \text{ stabilizes } P, \\ 0, & \text{otherwise.} \end{cases} \tag{2.3}$$

Then, several formalizable properties associated with robustness and performance can be stated as follows (see Table 1) [26]:

Table 1. List of symbols and notations.

Symbol	Description
$\|H\|_\infty$	H_∞ norm of a transfer matrix $H(s)$
$\det(H)$	Determinant of a transfer matrix $H(s)$
$H^*(s)$	Complex conjugate transpose of a transfer matrix $H(s)$; for real-rational systems considered in this study, it is defined as $H^*(s) := H^T(-s)$
$P = NM^{-1} = \tilde{M}^{-1}\tilde{N}$	Normalized right and left coprime factorizations of $P(s)$
$G = \begin{bmatrix} N \\ M \end{bmatrix}$	Normalized right graph symbol of a system P
$\tilde{G} = [-\tilde{M}\ \tilde{N}]$	Normalized left graph symbol of a system P
δ_ν	ν-gap metric
ν_{mod}	Modified ν-gap metric
$b_{P,C}$	Closed-loop stability margin for the loop formed by P and C
$\text{wno}(g)$	Winding number of $g(s)$ along an LHP-indented Nyquist contour (avoiding imaginary-axis poles)
$\bar{\sigma}(\cdot)$	Maximum (largest) singular value of a matrix

Proposition 1. *Let $[P, C]$ denote the standard feedback configuration. Given a nominal transfer function $P_1(s)$ and a perturbed function $P_2(s)$ along with a compensator $C(s)$:*

(i) $[P_2, C]$ is stable for all perturbed transfer functions P_2 satisfying $\delta_\nu(P_1, P_2) \leq \beta$, if and only if $b_{P_1,C} > \beta$,

(ii) $[P_2, C]$ is stable for all compensators C satisfying $b_{P_1,C} > \beta$, if and only if $\delta_\nu(P_1, P_2) \leq \beta$,

(iii) $\delta_\nu(P_1, P_2) \leq \|H(P_1, C) - H(P_2, C)\|_\infty \leq \delta_\nu(P_1, P_2)(b_{P_1,C} b_{P_2,C})^{-1}$, where $H(P, C)$ denotes the closed-loop transfer function as $H(P, C) := \begin{bmatrix} P \\ I \end{bmatrix} (I - CP)^{-1} \begin{bmatrix} -C & I \end{bmatrix}$.

This proposition implies that the stability of a closed-loop system, when a nominal plant $P_1(s)$ undergoes modification or perturbation to $P_2(s)$, can be efficiently assessed by utilizing both the ν-gap metric and the stability margin [19]. Moreover, this proposition highlights that when provided with a controller $C(s)$ and plants $P_1(s)$ and $P_2(s)$, the ν-gap metric and stability margin serve as key quantitative tools for estimating bounds on closed-loop performance.

2.2. Reduced-order model approximation evaluated by the ν-gap metric

A reduced-order model is typically employed to accelerate simulations and analyses of higher-order large-scale systems, as well as to simplify control design procedures, particularly when computational resources are limited. Thus, it is crucial for such a reduced-order model to closely capture both the steady-state and transient behaviors of the underlying large-scale system. As described in Section 2.1, the ν-gap metric provides a quantitative measure of the closed-loop discrepancy between two models, offering upper and lower bounds on the performance when used in conjunction with the stability margin. Therefore, $\delta_\nu(P_1, P_2)$ in (2.2) provides a reliable measure of the discrepancy between $P_1(s)$ and $P_2(s)$, both of which are subject to control by the same feedback compensator with nearly unity gain [26].

From this perspective, constructing a reduced-order model that closely approximates the original model in the ν-gap sense appears feasible. However, in the ν-gap metric, the discrepancy between the two systems is evaluated solely through the H_∞ norm. This highlights a crucial limitation of the conventional ν-gap model reduction, as it focuses exclusively on the maximal deviation over the entire frequency spectrum, without accounting for variations across different frequency bands. An illustrative example of this issue is presented below.

Example 1. *Consider a plant characterized by the following fourth-order transfer function [17]:*

$$P(s) = \frac{64.297(s^2 + 2.1488s + 400.95)}{s(s + 0.066411)(s^2 + 3.9386s + 4661.5)}. \tag{2.4}$$

The system is inherently unstable; however, it can be effectively stabilized by implementation of a simple unity negative feedback control. Given this setup, the ν-gap metric, which evaluates the distance between two systems under closed-loop conditions, appears to be a promising tool for assessing the accuracy of reduced-order models. Now, consider reducing the plant in (2.4) to the following second-order models:

$$R_1(s) = \frac{9.307}{s^2 + 1.569s - 2.035}, \tag{2.5}$$

$$R_2(s) = \frac{5.356}{s^2 + 0.0756s + 0.0182}. \tag{2.6}$$

Notably, the two reduced-order models possess distinct system parameters. However, the ν-gaps between $P(s)$ and the two reduced-order models, $R_1(s)$ and $R_2(s)$, are both computed to be approximately 0.2136; i.e., $\delta_\nu(P(s), R_1(s)) \approx \delta_\nu(P(s), R_2(s)) \approx 0.2136$. This result suggests that both $R_1(s)$ and $R_2(s)$ are expected to yield closed-loop performances similar to that of $P(s)$. In practice, however, the step response of $R_2(s)$ under unity negative feedback closely resembles the response of $P(s)$, whereas $R_1(s)$ exhibits significantly different behavior, as illustrated in Figure 1. This example demonstrates that relying solely on the ν-gap metric for MOR may not guarantee minimal response discrepancy between the original and reduced-order models.

These observations reveal that although the conventional ν-gap metric can effectively assess certain closed-loop properties, it has a limitation in accurately capturing the time-domain behavior, particularly regarding the transient response. This limitation underscores the need for a refined metric that explicitly accounts for these critical dynamics. To address this issue, a modified ν-gap metric is proposed in

Section 3, aiming to better reflect the time-domain characteristics of the system and ensure a closer match between the original and reduced-order models.

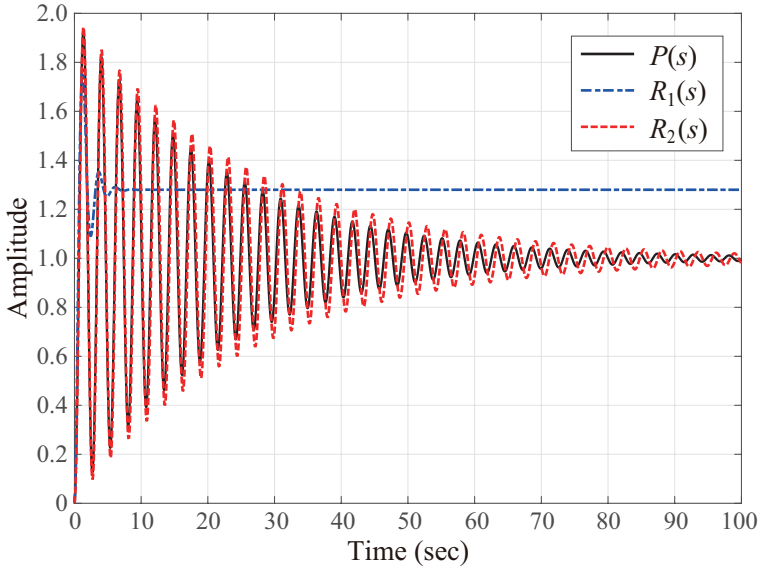


Figure 1. Step responses of the original model $P(s)$ and the reduced-order models $R_1(s)$ and $R_2(s)$ under unity feedback control.

3. Modified ν-gap metric and model reduction framework

3.1. Modified ν-gap metric: Formulation and validation

The conventional ν-gap metric has limited ability to capture time-domain characteristics between original and reduced-order models, highlighting the need for a refined approach. To overcome this limitation, this section introduces a modified ν-gap metric, designed to capture time-domain dynamics more effectively and improve the fidelity of reduced-order models in closed-loop performance.

In the frequency domain, the quantity $\tilde{G}_2 G_1$ involving the operators M_1, N_1, $\tilde{M}_2$, and $\tilde{N}_2$ can be expressed as follows:

$$\tilde{G}_2 G_1 = \begin{bmatrix} -\tilde{M}_2 & \tilde{N}_2 \end{bmatrix} \begin{bmatrix} N_1 \\ M_1 \end{bmatrix} = -\tilde{M}_2 N_1 + \tilde{N}_2 M_1 = \tilde{M}_2(-N_1 M_1^{-1} + \tilde{M}_2^{-1}\tilde{N}_2)M_1 = \tilde{M}_2(-P_1 + P_2)M_1. \quad (3.1)$$

Given matrices A and B, if B is right-invertible, then $\bar{\sigma}(AB) = \bar{\sigma}(A(BB^*)^{1/2})$. Similarly, for a left-invertible matrix A, it holds that $\bar{\sigma}(AB) = \bar{\sigma}((A^*A)^{1/2}B)$. Here, $\bar{\sigma}(\cdot)$ denotes the maximum singular value of a matrix. Applying these identities to (3.1) yields the following expression for the maximum singular value:

$$\bar{\sigma}(\tilde{M}_2(-P_1 + P_2)M_1) = \bar{\sigma}((\tilde{M}_2^*\tilde{M}_2)^{1/2}(P_2 - P_1)(M_1 M_1^*)^{1/2}). \quad (3.2)$$

To express $(M_1 M_1^*)^{-1}$ and $(\tilde{M}_2^*\tilde{M}_2)^{-1}$ in terms of P_1 and P_2, the derivation begins with the normalization condition for the right coprime factorization of P_1:

$$N_1^* N_1 + M_1^* M_1 = I. \quad (3.3)$$

Pre-multiplying by $(M_1^*)^{-1}$ and post-multiplying by M_1^{-1} on both sides of this equation yields the following identity:

$$(M_1^*)^{-1}N_1^*N_1(M_1)^{-1} + I = (M_1^*)^{-1}(M_1)^{-1}. \tag{3.4}$$

By using $P_1 = N_1M_1^{-1}$ and the matrix identity $(M_1^*)^{-1}(M_1)^{-1} = (M_1M_1^*)^{-1}$, (3.4) can be reformulated as:

$$P_1^*P_1 + I = (M_1M_1^*)^{-1}. \tag{3.5}$$

Similarly, for the P_2, the derivation proceeds using the normalization condition for the left coprime factorization:

$$\tilde{N}_2\tilde{N}_2^* + \tilde{M}_2\tilde{M}_2^* = I. \tag{3.6}$$

Pre-multiplying by $\tilde{M}_2^{-1}$ and post-multiplying by $(\tilde{M}_2^*)^{-1}$ on both sides of this equation yields:

$$\tilde{M}_2^{-1}\tilde{N}_2\tilde{N}_2^*(\tilde{M}_2^*)^{-1} + I = \tilde{M}_2^{-1}(\tilde{M}_2^*)^{-1}. \tag{3.7}$$

By substituting $P_2 = \tilde{M}_2^{-1}\tilde{N}_2$ and applying the matrix identity $\tilde{M}_2^{-1}(\tilde{M}_2^*)^{-1} = (\tilde{M}_2^*\tilde{M}_2)^{-1}$, this equation can be rewritten as:

$$P_2P_2^* + I = (\tilde{M}_2^*\tilde{M}_2)^{-1}. \tag{3.8}$$

Substituting (3.5) and (3.8) into (3.2) yields the frequency-domain representation

$$\bar{\sigma}(\tilde{G}_2G_1)(j\omega) = \bar{\sigma}((I + P_2P_2^*)^{-1/2}(P_2 - P_1)(I + P_1^*P_1)^{-1/2})(j\omega), \tag{3.9}$$

which serves as the core expression underlying the modified ν-gap metric. The relation in (3.9) provides the theoretical basis for the proposed modified ν-gap. It establishes the mathematical equivalence between the abstract graph metric kernel $\tilde{G}_2G_1$ and a normalized, weighted frequency-domain error representation. This equivalence rigorously justifies the definition of the modified ν-gap in (3.10) as a cumulative performance metric rooted in coprime factorization and ν-gap theory, rather than as a heuristic construction. This structure suggests a form of frequency-dependent weighting, emphasizing discrepancies in frequency regions where the system exhibits greater sensitivity. The modified ν-gap metric provides a frequency-resolved assessment of the discrepancy between two systems. In contrast to the conventional approach that focuses solely on the worst-case behavior, this formulation aggregates differences across multiple frequencies. It is defined as

$$\nu_{mod}(P_1, P_2) = \sum_{i=1}^{N} \bar{\sigma}(\tilde{G}_2G_1)(j\omega_i), \tag{3.10}$$

where $\omega_i \in [\omega_l, \omega_h]$ represents discrete frequency points sampled between the lower and upper frequency bounds ω_l and ω_h, respectively, and N is the total number of frequency samples. The metric in (3.10) is well defined only when the WNC in (2.1) is satisfied. As N increases, the modified ν-gap metric provides an increasingly detailed evaluation of system behavior across the frequency spectrum. Moreover, the flexibility of this approach allows its tailoring to various applications by adjusting the frequency bounds or applying frequency-dependent weights. This makes the modified ν-gap metric particularly effective in applications in which specific frequency bands are of primary interest or where transient and time-domain behavior must be accurately captured.

To better understand the motivation behind and implications of the proposed enhancement, revisiting the characteristics and limitations of the conventional ν-gap metric would be useful. The

conventional ν-gap metric, defined in (2.2), measures the discrepancy between two systems by evaluating the worst-case deviation in their closed-loop responses. Mathematically, it corresponds to the H_∞-norm of the transfer function $\tilde{G}_2 G_1$, which selects the maximum singular value achieved across all frequencies. This approach offers a conservative and robust measure, particularly valuable in theoretical analysis and controller synthesis. However, its supremum-based formulation inherently limits its ability to capture nuanced differences that may occur at frequencies outside the peak response, particularly in applications where performance is sensitive to frequency-localized dynamics or time-domain fidelity. The modified ν-gap metric proposed herein addresses these limitations by adopting a frequency-resolved perspective. Instead of isolating a single worst-case frequency, the metric aggregates the singular value responses over a prescribed range of frequency samples. This cumulative formulation retains the same internal structure as the original metric; however, it shifts the focus from extremal analysis to spectral integration. Accordingly, fine-grained deviations that may be critical for applications such as model reduction, system identification, or transient response preservation can be captured. This formulation naturally supports improved frequency resolution, as it accounts for discrepancies across the full spectrum instead of only the peak. Furthermore, owing to summing over multiple frequency points, the modified metric becomes less susceptible to localized anomalies or outliers, leading to a more robust and representative assessment of model similarity. The increased sensitivity to subtle differences also makes it suitable for systems in which small perturbations can accumulate and degrade performance. From a practical standpoint, the modified ν-gap aligns well with the reality of empirical modeling. In many real-world scenarios, frequency response data are available only at discrete points, particularly in data-driven or experimental settings. Although the conventional H_∞-based ν-gap can still be computed using standard tools, its continuous-domain formulation may not entirely reflect the discrete nature of available data. By contrast, the proposed approach is inherently compatible with discretely sampled frequency responses, making it more naturally applicable in such environments. Furthermore, its structure allows for the introduction of frequency-dependent weighting, enabling designers to emphasize specific frequency bands or prioritize certain time-domain characteristics according to application requirements. In addition to its interpretability and adaptability, the modified formulation is computationally efficient. Instead of searching for a global maximum over a continuous domain, it evaluates singular values at a fixed set of frequencies, reducing computational burden while maintaining accuracy. The modified ν-gap metric represents a conceptual shift: from a worst-case, peak-focused philosophy to a frequency-aware and integrated framework more suitable for modern control and model-reduction tasks. In summary, the modified ν-gap offers a principled and practical alternative to the standard formulation. By integrating information across the frequency spectrum and enabling targeted customization through weighting, it provides a more flexible and insightful metric for comparing dynamical systems in both theoretical and applied contexts.

Example 2. *To validate the practical utility of the modified ν-gap metric in (3.10), once again consider the reduced-order models $R_1(s)$ and $R_2(s)$ from Example 1. Despite their conventional ν-gap values being nearly identical—$\delta_\nu(P(s), R_1(s)) = 0.21361$ and $\delta_\nu(P(s), R_2(s)) = 0.21367$—their corresponding time-domain performances differ significantly as shown in Figure 1. This discrepancy arises because the conventional ν-gap metric evaluates only the worst-case (maximum) singular value across all frequencies and thus fails to reflect how discrepancies are distributed over the frequency spectrum. By contrast, the modified ν-gap metric aggregates the frequency-wise singular value*

differences, providing a more detailed assessment of system similarity. As illustrated in Figure 2, the frequency-wise ν-gap profile of $R_1(s)$ exhibits broad discrepancies across a wide frequency band, whereas $R_2(s)$ shows a localized peak but remains close to zero over most of the spectrum. Because the modified ν-gap metric sums these values over all frequencies, $R_2(s)$ yields a substantially smaller metric value ($\nu_{mod}(P, R_2) = 53.0$) than $R_1(s)$ ($\nu_{mod}(P, R_1) = 1259.4$) as given in Table 2, indicating a closer overall match to the original system. This observation is further supported by time-domain response accuracy, evaluated using the following standard performance indices:

$$ISE = \int_0^\infty (y_p(t) - y_r(t))^2 \, dt, \qquad IAE = \int_0^\infty \left| y_p(t) - y_r(t) \right| \, dt, \tag{3.11}$$

where $y_p(t)$ and $y_r(t)$ denote the unit step responses of the original and reduced-order models $P(s)$ and $R(s)$, respectively. Table 2 reveals that $R_2(s)$ achieves substantially lower error values (integral square error (ISE) = 9.3, integral absolute error (IAE) = 95.6) compared to $R_1(s)$ (ISE = 214.0, IAE = 594.2), despite both models having nearly equal conventional ν-gap values. This confirms that the modified ν-gap metric reflects the response similarity between the original and reduced-order systems more faithfully, and it is thus more suitable for use in model reduction contexts where time-domain fidelity is critical.

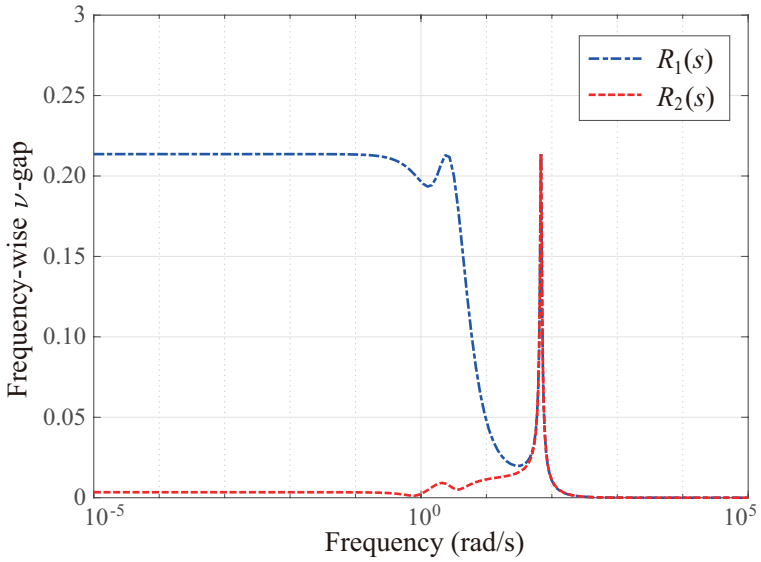


Figure 2. Frequency-wise ν-gap profiles of reduced-order models $R_1(s)$ and $R_2(s)$, computed using Eq (3.9).

Table 2. Quantitative comparison of reduced-order models based on ν-gap metrics and time-domain error indices.

Model	$\delta_\nu(P(s), R_i(s))$	$\nu_{mod}(P(s), R_i(s))$	ISE	IAE
$R_1(s)$	0.21361	1259.4	214.0	594.2
$R_2(s)$	0.21367	**53.0**	**9.3**	**95.6**

3.2. PSO-driven model reduction with frequency-sensitive accuracy using the modified ν-gap metric

This subsection presents a parameter identification framework for reduced-order models, in which the model coefficients are explicitly optimized to minimize the modified ν-gap metric introduced earlier. The objective is to construct a reduced-order transfer function that closely matches the original system, with particular emphasis on preserving frequency-sensitive and time-domain behavior. To this end, the model reduction problem is formulated as a constrained optimization task, where the modified ν-gap serves as the objective function and WNC is enforced as a critical feasibility constraint to ensure the metric's well-posedness.

To concretize this framework, the MOR problem is formulated to identify the coefficients of a reduced-order transfer function that closely approximates the original high-order model. Given an original mth-order model $P(s)$, the structure of the reduced-order model is fixed as

$$R(s, \boldsymbol{x}) = \frac{x_n s^b + x_{n-1} s^{b-1} + \cdots + x_{n-b+1} s + x_{n-b}}{s^a + x_{n-b-1} s^{a-1} + x_{n-b-2} s^{a-2} + \cdots + x_2 s + x_1}, \tag{3.12}$$

where $a < m$ and $n = a + b + 1$. The integers a and b are determined by the designer, where the order a may be selected by examining the Hankel singular values (HSVs) of the original model, which quantify the relative energy contributions of system states. The HSV spectrum decay profile provides a heuristic guideline for selecting an appropriate reduced order. The objective of the reduction problem is to minimize the discrepancy between $P(s)$ and $R(s)$, as measured by the modified ν-gap metric introduced earlier. To achieve this objective, the optimization problem is formulated with respect to the decision vector $\boldsymbol{x} \in \mathbb{R}^n$, which contains the unknown coefficients of the reduced-order model $R(s)$, arranged as $\boldsymbol{x} := (x_1,\ x_2 \cdots, x_{n-b-2},\ x_{n-b-1}, x_{n-b}, \cdots, x_{n-1},\ x_n)^T \in \mathbb{R}^n$, where the entries x_1 through x_{n-b-1} and x_{n-b} through x_n correspond to the denominator coefficients (excluding the leading coefficient, which is normalized to 1) and the numerator coefficients, respectively. This partitioning ensures that the total number of unknown coefficients is $n = a + b + 1$, consistent with the reduced-order model structure defined in (3.12).

The resulting optimization problem poses notable computational challenges due to its nonconvex nature [19] and the requirement that the WNC be satisfied in order for the modified ν-gap metric to be well-defined. Traditional gradient-based optimization techniques are often ill-suited for such problems, because they tend to become trapped in local minima and are not inherently equipped to handle structural constraints such as WNC. To address this limitation, a PSO approach is employed, which is particularly effective for navigating nonconvex search spaces. In this framework, each particle in the swarm represents a candidate solution $\boldsymbol{x}$, which is a possible set of coefficients for the reduced-order model $R(s, \boldsymbol{x})$. The optimization objective is to minimize the following cost function:

$$\min_{\boldsymbol{x}} \mathcal{F}(\boldsymbol{x}) := \begin{cases} \arctan\left(\nu_{mod}(P(s), R(s, \boldsymbol{x}))\right) - \dfrac{\pi}{2}, & \text{if WNC is satisfied,} \\ 1, & \text{otherwise.} \end{cases} \tag{3.13}$$

This objective function $\mathcal{F}(\boldsymbol{x})$ is deliberately constructed to guide the swarm away from infeasible regions that violate the WNC. If a particle's position results in a reduced model $R(s, \boldsymbol{x})$ that fails the WNC, the function is set to one, representing a flat (nondescent) direction and discouraging further exploration of that region. By contrast, feasible points yield strictly negative objective values (because $\nu_{mod} > 0$ and $\arctan(\cdot) - \pi/2 < 0$), thereby promoting convergence toward valid and high-fidelity

reduced models. To verify whether a candidate $\boldsymbol{x}$ satisfies the WNC, the matrix product $G_2^{\sim}(s)G_1(s)$ is evaluated, where $G_2^{\sim}(s) := G_2^*(-s^*)$ denotes the para-Hermitian conjugate of the transfer matrix $G_2(s)$. This formulation ensures that the WNC is checked using the correct topological configuration. The determinant $\det(G_2^{\sim}(s)G_1(s))$ is analyzed by counting its right-half-plane (RHP) poles and zeros. The WNC is considered satisfied if the difference between the net number of RHP poles and RHP zeros is zero, which guarantees that the Nyquist plot of the determinant does not encircle the origin. This topological condition is critical for ensuring that the modified ν-gap metric is well defined for the candidate reduced-order model $R(s, \boldsymbol{x})$.

The optimization problem defined in (3.13) is addressed using a customized PSO framework known as cyclic-neighborhood topology PSO with quantum infusion (CNT-PSO-QI) [21]. This approach is particularly effective for solving nonconvex and constrained problems such as the present model-reduction task, where feasibility with respect to the WNC must be maintained while optimizing a nonsmooth and nonlinear objective. The swarm comprises n_p particles denoted by $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_{n_p}$, which iteratively explore the parameter space to locate the global minimizer $\boldsymbol{x}^*$ of the cost function $\mathcal{F}(\boldsymbol{x})$ defined in (3.13). Each particle maintains its own position and velocity in the n-dimensional design space, corresponding to the coefficient vector $\boldsymbol{x}$ defining the reduced-order model $R(s, \boldsymbol{x})$. Specifically, the i-th particle at generation k is represented by

$$\boldsymbol{x}_i^k := (x_{i,1}^k, x_{i,2}^k, \cdots, x_{i,n}^k)^T \in \mathbb{R}^n, \quad \boldsymbol{v}_i^k := (v_{i,1}^k, v_{i,2}^k, \cdots, v_{i,n}^k)^T \in \mathbb{R}^n, \tag{3.14}$$

denoting its current solution candidate and associated search velocity, respectively. At each iteration, particles update their states by aggregating three influence sources: (i) the momentum from the previous velocity (inertia), (ii) the attraction toward the particle's own best historical solution (personal best), and (iii) guidance from the best solutions found by its neighborhood and a quantum-inspired offspring mechanism (social best). The detailed update mechanism of each particle's state is as follows:

$$\boldsymbol{v}_i^{k+1} = c_0\boldsymbol{v}_i^k + c_1 r_{1,i}^k(\boldsymbol{x}_{Pbest,i}^k - \boldsymbol{x}_i^k) + c_2 r_{2,i}^k(\boldsymbol{x}_{Sbest,i}^k - \boldsymbol{x}_i^k), \quad \boldsymbol{x}_i^{k+1} = \boldsymbol{x}_i^k + \boldsymbol{v}_i^{k+1}, \tag{3.15}$$

where c_0, c_1, and c_2 are the inertia, cognitive, and social coefficients, respectively, and $r_{1,i}^k$ and $r_{2,i}^k$ are uniformly distributed random scalars in $[0, 1]$ independently drawn at each generation. The personal best position $\boldsymbol{x}_{Pbest,i}^k$ tracks the best solution encountered by the i-th particle up to generation k, based on the objective function value $\mathcal{F}(\boldsymbol{x})$ defined in (3.13) as

$$\boldsymbol{x}_{Pbest,i}^k := \arg \min_{\boldsymbol{x} \in \{\boldsymbol{x}_i^c | c=1,\cdots,k\}} \mathcal{F}(\boldsymbol{x}). \tag{3.16}$$

The social best position $\boldsymbol{x}_{Sbest,i}^k$ is derived from a comparison between the cyclic neighborhood best and a quantum-infused offspring, ensuring both local exploitation and exploratory diversity:

$$\boldsymbol{x}_{Sbest,i}^k := \arg \min_{\boldsymbol{x} \in \{\boldsymbol{x}_{social,i}^k,\ \boldsymbol{x}_{offspring,i}^k\}} \mathcal{F}(\boldsymbol{x}), \tag{3.17}$$

where the best neighbor $\boldsymbol{x}_{social,i}^k$ is defined as

$$\boldsymbol{x}_{social,i}^k := \arg \min_{\boldsymbol{x} \in \{\boldsymbol{x}_j^k | j=i-\frac{n_s}{2},\cdots,i+\frac{n_s}{2}\}} \mathcal{F}(\boldsymbol{x}), \tag{3.18}$$

with $\boldsymbol{x}_j^k := \boldsymbol{x}^k_{(j-1 \bmod n_p)+1}$ for $j < 1$ or $n_p+1 \leq j$, and n_s is the predefined neighborhood size assumed to be even. The quantum-infused offspring vector $\boldsymbol{x}^k_{offspring,i}$ introduces additional diversity to the search and is computed as:

$$\boldsymbol{x}^k_{offspring,i} := \alpha \boldsymbol{x}^k_{social,i} + (1-\alpha)\boldsymbol{x}^k_{Pbest,i} \pm \beta \left|\boldsymbol{x}^k_{Mbest,i} - \boldsymbol{x}^k_r\right| \ln\left(\frac{1}{u}\right), \tag{3.19}$$

where α and u are random scalars uniformly distributed over $[0, 1]$ and $(0, 1]$, respectively, while β is a scaling constant and the $\pm$ sign is chosen at random with equal probability. The term $\boldsymbol{x}^k_r$ is a randomly selected vector from the social neighborhood, and the mean best position is defined as follows:

$$\boldsymbol{x}^k_{Mbest,i} := \frac{1}{n_s+1} \sum_{j=i-n_s/2}^{i+n_s/2} \boldsymbol{x}^k_{Pbest,j}, \tag{3.20}$$

where $\boldsymbol{x}^k_{Pbest,j} := \boldsymbol{x}^k_{Pbest,(j-1 \bmod n_p)+1}$ for $j < 1$ or $n_p + 1 \leq j$. This iterative update procedure continues over successive generations until a termination criterion is satisfied, such as a maximum number of iterations or stagnation in solution improvement. Through this evolutionary process, the swarm gradually converges toward the global minimizer $\boldsymbol{x}^*$, yielding a reduced-order model that minimizes the modified ν-gap metric while satisfying the WNC. This combination of structured neighborhood topology and probabilistic offspring generation distinguishes the implemented method from conventional PSO variants and has demonstrated superior performance in previous studies on challenging non-convex problems [21]. This hybrid mechanism balances local exploitation and global exploration by leveraging both deterministic memory and probabilistic diversity while remaining robust to local minima. The quantum infusion component further ensures that the swarm avoids premature convergence by injecting controlled randomness based on the distribution of elite solutions. Overall, the CNT-PSO-QI algorithm provides a robust and scalable approach to model reduction, optimized for minimizing the modified ν-gap under strict feasibility constraints, and delivering high-fidelity system approximation across diverse practical scenarios. Its flexibility and effectiveness highlight its value as a general-purpose tool for frequency-sensitive system simplification. For reproducibility, the parameter settings used for the CNT-PSO-QI algorithm in this study are summarized in Table 3.

Table 3. Definitions and parameter values used in the CNT-PSO-QI algorithm.

Symbol	Description	Value
c_0	Inertia coefficient	0.7298
c_1	Cognitive coefficient	1.4962
c_2	Social coefficient	1.4962
β	Quantum scaling factor	1.5
n_p	Population size	200
n_s	Cyclic neighborhood size	20
$\mathbb{D}$	Search domain of the decision variables	$[-10^3, 10^3]$
k_{max}	Maximum number of iterations	1000

To validate the practical effectiveness of the proposed framework, this section presents two

comparative examples, where the performance of the modified ν-gap–based model reduction approach is evaluated in comparison with those of existing methods reported in the literature.

Example 3. *The effectiveness and robustness of the proposed CNT-PSO-QI algorithm—explicitly designed to minimize the ν-gap metric δ_ν while ensuring feasibility through strict enforcement of the WNC—are demonstrated through a comparison with the BMI-based optimization method proposed by Taamallah [20]. Specifically, consider the original fifth-order system defined as*

$$P(s) = \frac{s+2}{(0.0025s+1)^2(0.1s+1)(0.05s^2-1)}, \tag{3.21}$$

which is approximated by the following second-order reduced-order model structure:

$$R(s,\boldsymbol{x}) = \frac{x_1 s + x_2}{s^2 + x_3 s + x_4}. \tag{3.22}$$

In this formulation, the decision vector is defined as $\boldsymbol{x} := (x_1, x_2, x_3, x_4)^T \in \mathbb{R}^4$, where x_1 and x_2 represent the numerator coefficients and x_3 and x_4 correspond to the denominator coefficients of the reduced-order model $R(s,\boldsymbol{x})$. In the study by Taamallah [20], the objective was to minimize the conventional ν-gap metric $\delta_\nu(P(s), R(s))$ defined in (2.2), without explicit enforcement of the WNC. Although the BMI-based formulation enables structured optimization using SDP tools such as YALMIP, it presents two main limitations: the WNC is not guaranteed during optimization, and the performance of the method exhibits high sensitivity to the choice of initial conditions.

To evaluate the advantages of the proposed method under identical problem settings, a numerical comparison was conducted between the BMI-based method and the CNT-PSO-QI approach. For the CNT-PSO-QI implementation, the swarm size and maximum number of iterations were set to $n_p = 200$ and $k_{max} = 1000$, respectively. The algorithm parameters were selected as $c_0 = 0.7298$, $c_1 = c_2 = 1.4962$, and quantum scaling factor $\beta = 1.5$. The cyclic neighborhood size was $n_s = 20$, and the search space was defined as $\mathbb{D} := \{\boldsymbol{x} \in \mathbb{R}^4 : -10^3 \le x_i \le 10^3\}$. Following common practice in metaheuristic optimization, the bounds used in this benchmark are explicitly reported and were deliberately chosen to be sufficiently conservative to allow adequate exploration of the search space. In the reported results, the obtained optimal solution was found to lie well inside the prescribed bounds, indicating that the optimization outcome is not artificially constrained by the selected initialization limits. To ensure fairness, both algorithms were executed 30 times using randomly generated initial conditions, including the initialization recommended by Taamallah [20]. Figure 3 depicts the convergence profiles of the objective function values over 1000 iterations. Evidently, the proposed CNT-PSO-QI algorithm achieves more stable and superior convergence behavior, whereas the BMI-based method exhibits greater variability due to its initialization dependence. Table 4 summarizes the statistical outcomes of the objective function values from the 30 independent trials. The BMI-based method yielded a wide performance range (best: 0.1106, worst: 0.7445, mean: 0.4753, standard deviation: 0.1424), indicating inconsistent convergence. By contrast, the CNT-PSO-QI algorithm consistently achieved the optimal value of 0.0965 with zero variance, demonstrating both high accuracy and robustness.

The corresponding optimal reduced-order models obtained from the two methods are as follows:

$$R^*_{BMI}(s) = \frac{-7.66s + 363.8}{s^2 + 28.23s - 134}, \tag{3.23}$$

$$R^*_{PSO}(s) = \frac{-4.81s + 345.5}{s^2 + 24.15s - 132.8}, \tag{3.24}$$

with associated conventional ν-gap values $\delta_\nu(P(s), R^*_{BMI}(s)) = 0.1106$ and $\delta_\nu(P(s), R^*_{PSO}(s)) = 0.0965$. Notably, neither the initial set proposed by Taamallah [20] nor other random initializations led to the best solution, which highlights the inherent sensitivity of the BMI-based approach to initial conditions and its limited capability for global convergence. By contrast, the CNT-PSO-QI algorithm exhibits strong global convergence behavior, consistently attaining the optimal solution regardless of initial conditions. Moreover, by explicitly enforcing the WNC at every evaluation step, it guarantees that all candidate models remain within the feasible domain, thereby ensuring the validity of the ν-gap metric throughout the optimization process. This integrated mechanism affirms the algorithm's robustness, constraint satisfaction, and practical effectiveness in solving non-convex model reduction problems.

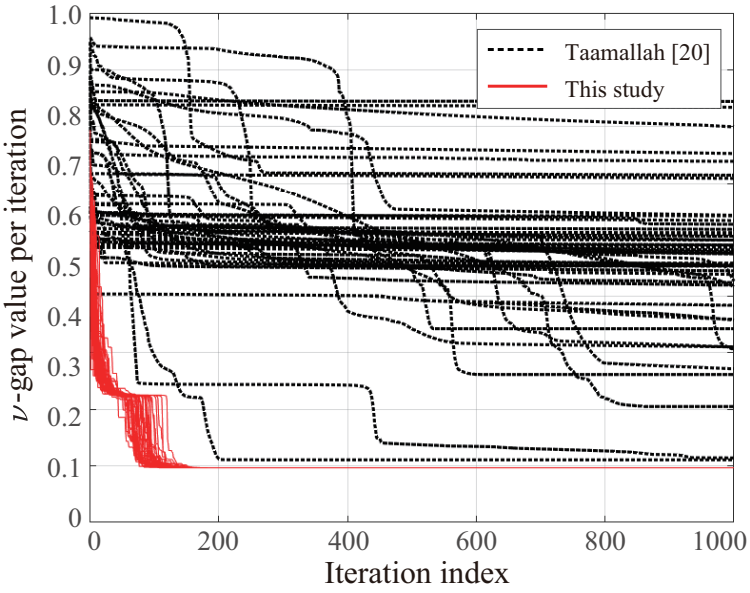


Figure 3. Convergence trajectories of the ν-gap metric δ_ν over iterations.

Table 4. Performance comparison for the ν-gap-based model reduction task.

	$\delta_\nu(P(s), R(s))$ statistics over 30 trials			
	Best	Worst	Mean	Standard Deviation
Taamallah [20]	0.1106	0.7445	0.4753	0.1424
Sootla [19]	0.2535	1.000	0.4158	0.1720
This paper	**0.0965**	**0.0965**	**0.0965**	**0.0000**

It is worth noting that mainstream model reduction techniques, such as balanced truncation or Krylov subspace methods, are not included in the comparative analysis presented in this section. These approaches are primarily designed to minimize open-loop approximation errors measured by H_2 or H_∞ norms and do not explicitly enforce the WNC, which is a fundamental requirement for the

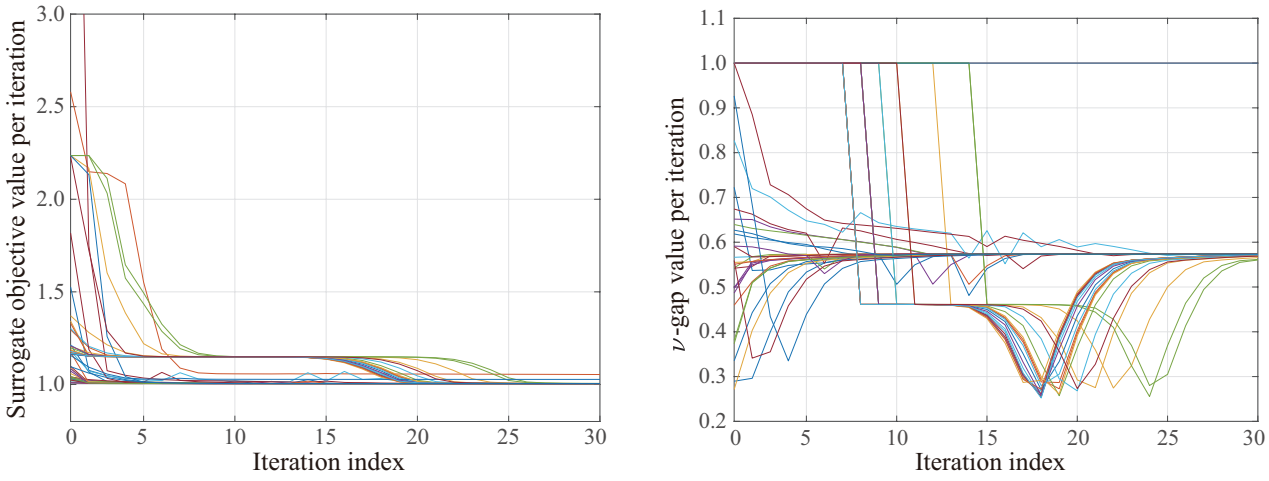
ν-gap metric to be well defined. When the WNC is violated, the ν-gap metric is taken as 1 by definition, yielding an uninformative result for ν-gap-based benchmarking regardless of time-domain response similarity. For this reason, the comparisons in this section are intentionally restricted to existing methods that explicitly target the ν-gap metric and address the WNC, thereby ensuring a mathematically consistent and meaningful evaluation under identical problem settings.

To further validate the proposed method under different optimization paradigms, the next example presents a comparison with the frequency-domain SDP-based approach introduced by Sootla [19].

Example 4. *This example evaluates the performance of the proposed CNT-PSO-QI algorithm against the model reduction approach presented by Sootla [19], which formulates the problem as a frequency-domain SDP task. In that work, the ν-gap between the original model $P(s)$ and the reduced-order model $R(s)$ is indirectly minimized by minimizing an upper bound γ, as expressed by*

$$\delta_\nu = \sqrt{1 - \frac{1}{\gamma_{opt}^2}}, \quad \gamma_{opt} \le \gamma, \tag{3.25}$$

where γ serves as a surrogate objective function. However, because γ is only an upper bound on the true optimal γ_{opt}, its reduction during optimization does not necessarily imply effective reduction of the actual ν-gap. To illustrate this limitation, the algorithm in Sootla [19] is applied to the same model reduction problem introduced in Example 3 using YALMIP. Frequency samples are uniformly drawn over $[10^{-5}, 10^{5}]$ (rad/s) with a total of 500 points. Initial optimization values are randomly generated, including the Hankel-based initialization recommended by Sootla [19]. Figure 4 presents the optimization traces for both the surrogate objective value γ and the corresponding computed ν-gap values δ_ν. The plots reveal that although γ decreases consistently, the actual ν-gap often fluctuates and does not improve accordingly. This highlights a key shortcoming of using γ as an indirect metric: the true system similarity as measured by the ν-gap is not necessarily improved.



(a) Optimization trajectory of the surrogate objective γ

(b) Computed ν-gap values δ_ν over iterations

Figure 4. Convergence behavior of the surrogate objective γ and the computed ν-gap δ_ν during optimization.

Another limitation of the method proposed by Sootla [19] concerns how feasibility with respect to the WNC is imposed. Instead of directly verifying the condition defined in (2.1), the method enforces a sufficient condition based on right coprime factorizations of the original and reduced models. Specifically, let $P(s) = N_1 M_1^{-1}$ *and* $R(s) = \tilde{N}_2 \tilde{M}_2^{-1}$*, with their adjoint coprime factors denoted by* $\tilde{M}_1$*,* $\tilde{N}_1$ *(i.e.,* $\tilde{G}_1 = [-\tilde{M}_1 \ \tilde{N}_1]$*). Then, the following inequality is imposed:*

$$\mathrm{Re}\left(\frac{\tilde{M}_2\tilde{M}_1 + \tilde{N}_2\tilde{N}_1}{M_c\tilde{M}_1 + N_c\tilde{N}_1}\right) > 0, \tag{3.26}$$

where M_c *and* N_c *are the right coprime factors of a central transfer function iteratively updated during the optimization process to satisfy:*

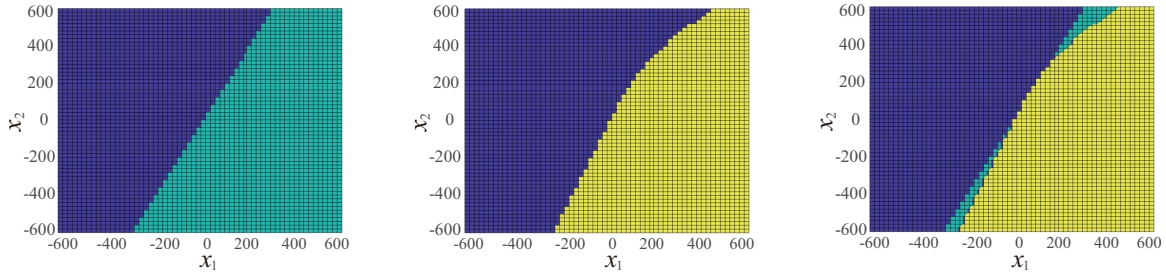
$$\mathrm{wno}(M_c\tilde{M}_1 + N_c\tilde{N}_1) = 0, \tag{3.27}$$

$$M_c\tilde{M}_1 + N_c\tilde{N}_1 \neq 0. \tag{3.28}$$

Although the condition in (3.26) ensures that the reduced model satisfies the WNC, it is only a sufficient condition, not a necessary one. Therefore, the feasible set of reduced-order models is overly conservative: it excludes certain valid solutions that would satisfy the true WNC but not this stricter condition. This limitation can result in suboptimal performance, particularly in nonconvex settings where broad exploration of the solution space is critical for achieving a minimal ν*-gap. To illustrate this conservativeness, consider the second-order structure*

$$R(s) = \frac{-4.81s + x_1}{s^2 + 24.15s + x_2}, \tag{3.29}$$

where (x_1, x_2) *are coefficients to be optimized. The WNC is tested across a grid of candidate coefficients, and the feasible regions under different conditions are shown in Figure 5. As illustrated in Figure 5(a), the WNC can be directly verified using the winding number test from (2.1). Figure 5(b) shows the feasible region under the sufficient condition (3.26), which proves to be a subset of the full WNC-satisfying region. Figure 5(c) overlays the two, confirming that the criterion used in Sootla [19] excludes many valid candidates. In summary, the method by Sootla [19] is limited in two important aspects. First, its optimization is based on a surrogate objective that does not directly minimize the* ν*-gap metric. Second, the WNC is enforced through a sufficient but not necessary condition, which may overly constrain the feasible set and lead to suboptimal solutions. By contrast, the proposed CNT-PSO-QI method directly minimizes the* ν*-gap metric and enforces the WNC exactly, enabling more reliable and accurate model reduction. To further corroborate this point, the statistical performances from* 30 *independent runs are compared in Table 4. The method of Sootla [19] yields* ν*-gap values with significant variation (best:* 0.2535*, worst:* 1.0000*, mean:* 0.4158*, standard deviation:* 0.1720*), indicating inconsistent convergence and a tendency to fall short of globally optimal solutions. By contrast, the proposed CNT-PSO-QI method reliably achieves the optimal value of* 0.0965 *with zero variance, demonstrating superior accuracy, robustness, and consistency across trials.*



(a) Feasible region (green) under the WNC condition in (2.1)

(b) Feasible region (yellow) under the sufficient condition in (3.26)

(c) Overlay of (a) and (b): green for (2.1), yellow for (3.26)

Figure 5. Comparison of WNC-feasible coefficient regions: direct test vs. sufficient condition.

The proposed CNT-PSO-QI method has exhibited excellent performance in minimizing the modified ν-gap, providing high-fidelity reduced models that align well with the time-domain behavior. However, the conventional ν-gap offers distinct advantages in robustness analysis and closed-loop stability assessment. To completely leverage the complementary strengths of both metrics, the next section introduces a multi-objective optimization framework that integrates the two perspectives for balanced and robust model reduction.

Remark 1. *In the numerical examples of Section 3.2, the CNT-PSO-QI algorithm is employed as a single-objective solver to identify reduced-order models satisfying the WNC. Because feasibility is governed by the WNC, the admissible region may become highly non-convex and fragmented, where feasible candidates occupy only narrow subsets of the decision space. Under such conditions, population-based optimization methods generally require sufficiently broad exploration to reliably detect and preserve feasible particles. For this reason, a relatively large swarm was adopted to enhance numerical robustness under the WNC constraint, rather than to heuristic tuning for individual examples. The particle update coefficients were fixed to the standard constriction-factor-based configuration* $\{c_0, c_1, c_2\} = \{0.7298, 1.4962, 1.4962\}$*, following classical PSO stability guidelines that balance exploration and convergence while preventing particle divergence. In addition, the quantum-infusion parameter* β *was set to a moderate default value (*$\beta = 1.5$*) to balance global exploration and convergence stability, consistent with the original CNT-PSO-QI framework. The cyclic neighborhood size was also fixed to preserve diversity without ad hoc tuning. All algorithmic parameters were kept unchanged across the examples in Section 3.2 to ensure reproducibility. The statistical results summarized in Table 4 further indicate highly consistent outcomes over* 30 *independent runs, supporting the numerical stability and robustness of the adopted parameter configuration for WNC-constrained model reduction.*

4. Multi-objective model reduction using complementary ν-gap metrics

Although the proposed CNT-PSO-QI algorithm has demonstrated outstanding performance in minimizing the modified ν-gap metric ν_{mod}, it is important to recognize that the conventional ν-gap metric δ_ν remains highly relevant, particularly for assessing closed-loop robustness and stability

margins. By contrast, ν_{mod} offers improved resolution in capturing frequency-localized discrepancies and often aligns more closely with time-domain accuracy. Nevertheless, these two metrics are not always tightly coupled: minimizing ν_{mod} does not necessarily guarantee minimization of δ_ν, and a trade-off between the two can emerge in practical scenarios. Furthermore, the computation of ν_{mod} depends on frequency sampling and may not include the full analytic structure captured by the conventional ν-gap, particularly in cases involving limited frequency grids or weighting filters. Thus, the two metrics can be considered complementary, each reflecting distinct yet critical aspects of model quality.

Motivated by this complementarity and the potential trade-off, in this section, we introduce a multi-objective model reduction framework that simultaneously minimizes δ_ν and ν_{mod}. By integrating their respective advantages through a tailored MOPSO scheme, the proposed approach aims to construct reduced-order models that achieve a balanced compromise between robustness and fidelity, making it well-suited for control-oriented applications demanding both theoretical rigor and practical effectiveness.

4.1. Formulation and algorithmic structure

To simultaneously optimize the conventional and modified ν-gap metrics, a multi-objective optimization problem is formulated. The proposed method adopts a MOPSO framework, which evolves a population of candidate solutions to generate and iteratively refine a Pareto-optimal set of reduced-order models. Each particle in the swarm encodes a candidate solution represented by the parameter vector $\boldsymbol{x} \in \mathbb{R}^n$, defining the coefficients of the reduced-order model. The objectives are defined to penalize candidate models that violate the WNC. To achieve this, the fitness functions are transformed using an arctangent-based scheme to ensure that infeasible models are effectively excluded from the search process. The two objective functions are defined as follows:

$$\mathcal{F}_1(\boldsymbol{x}) := \begin{cases} \arctan\left(\delta_\nu(P(s), R(s, \boldsymbol{x}))\right) - \dfrac{\pi}{2}, & \text{if WNC is satisfied,} \\ 1, & \text{otherwise,} \end{cases} \tag{4.1}$$

$$\mathcal{F}_2(\boldsymbol{x}) := \begin{cases} \arctan\left(\nu_{mod}(P(s), R(s, \boldsymbol{x}))\right) - \dfrac{\pi}{2}, & \text{if WNC is satisfied,} \\ 1, & \text{otherwise.} \end{cases} \tag{4.2}$$

Both objectives yield values in $[-\pi/2, 0)$ when WNC holds and return 1 otherwise, thereby imposing strict separation between feasible and infeasible solutions. This enables the swarm to effectively explore only the WNC-satisfying region. Furthermore, because the conventional ν-gap (δ_ν) and the modified ν-gap (ν_{mod}) have different physical interpretations and numerical magnitudes, both metrics are transformed using the bounded nonlinear mapping $\mathcal{F}(\cdot) = \arctan(\cdot) - \frac{\pi}{2}$. This transformation acts as a normalization step that compresses their disparate scales into a unified, dimensionless interval $[-\pi/2, 0)$, preventing scale domination and enabling balanced multi-objective optimization under Pareto dominance.

The detailed implementation steps of the proposed MOPSO algorithm are as follows:

Step 0 (Initialization). Determine the denominator order a for the reduced-order model in (3.12) based on the HSVs of the original system. Set the numerator order b (e.g., $b = a - 1$), and define

the search dimension as $n = a + b + 1$. Select algorithm parameters: iteration limit k_{max}, population size n_p, repository size n_r, grid resolution n_d, mutation rate q_r, and quantum scaling parameter β. Define the search space bounds $\boldsymbol{x}_{min} \in \mathbb{R}^n$ and $\boldsymbol{x}_{max} \in \mathbb{R}^n$, specifying the optimization domain as $\mathbb{D} := \{\boldsymbol{x} \in \mathbb{R}^n : x_{min,j} \leq x_j \leq x_{max,j},\ j = 1, 2, \ldots, n\}$, and initialize the iteration index $k = 1$. The Pareto repository *REP* to store Pareto-optimal solutions is initialized to be empty.

Step 1 (Particle Initialization). Randomly initialize particle positions and set their initial velocities to zero:

$$\boldsymbol{x}_i^1 = \boldsymbol{x}_{min} + \boldsymbol{r}_{1,i}^1(\boldsymbol{x}_{max} - \boldsymbol{x}_{min}), \tag{4.3}$$

$$\boldsymbol{v}_i^1 = \boldsymbol{0}_n, \quad i = 1, 2, \ldots, n_p, \tag{4.4}$$

where $\boldsymbol{r}_{1,i}^1 \in \mathbb{R}^{n\times n}$ is a diagonal matrix of uniformly distributed random numbers in $[0, 1]$, and $\boldsymbol{0}_n$ denotes an n-dimensional zero vector.

Step 2 (Fitness Evaluation). Evaluate both fitness functions $\mathcal{F}_1(\boldsymbol{x}_i^k)$ and $\mathcal{F}_2(\boldsymbol{x}_i^k)$ for all particles at iteration k.

Step 3 (Repository and Personal Best Update). In multi-objective optimization, a solution $\boldsymbol{X}$ is said to dominate another solution $\boldsymbol{Y}$ if it is no worse than $\boldsymbol{Y}$ in all objectives, and strictly better in at least one. This concept of Pareto dominance is used to identify non-dominated solutions that collectively define the Pareto-optimal front. For each particle $\boldsymbol{x}_i^k$, based on its current objective values $\mathcal{F}_1(\boldsymbol{x}_i^k)$ and $\mathcal{F}_2(\boldsymbol{x}_i^k)$ obtained in Step 2, if $\boldsymbol{x}_i^k$ is not dominated by any other particle in the current swarm or in the Pareto-optimal repository *REP*, then it is added to the repository. If $\boldsymbol{x}_i^k$ dominates any existing entries in the repository, those dominated entries are removed.

The personal best position $\boldsymbol{x}_{p,i}^k$ is then updated by comparing the current particle position $\boldsymbol{x}_i^k$ against the previous personal best $\boldsymbol{x}_{p,i}^{k-1}$ using their respective objective values. The update rule is defined as follows:

$$\boldsymbol{x}_{p,i}^k = \begin{cases} \boldsymbol{x}_i^k, & \text{if } \mathcal{F}_1(\boldsymbol{x}_i^k) < \mathcal{F}_1(\boldsymbol{x}_{p,i}^{k-1}) \text{ and } \mathcal{F}_2(\boldsymbol{x}_i^k) < \mathcal{F}_2(\boldsymbol{x}_{p,i}^{k-1}), \\ \boldsymbol{x}_{p,i}^{k-1}, & \text{if } \mathcal{F}_1(\boldsymbol{x}_i^k) > \mathcal{F}_1(\boldsymbol{x}_{p,i}^{k-1}) \text{ and } \mathcal{F}_2(\boldsymbol{x}_i^k) > \mathcal{F}_2(\boldsymbol{x}_{p,i}^{k-1}), \\ \boldsymbol{x}_i^k \text{ or } \boldsymbol{x}_{p,i}^{k-1}, & \text{otherwise, chosen randomly with probability 0.5,} \end{cases} \tag{4.5}$$

with initial setting $\boldsymbol{x}_{p,i}^1 := \boldsymbol{x}_i^1$. It is important to note that the random selection in (4.5) is applied only as a tie-breaking mechanism when the current solution and the previous personal best are mutually non-dominated. In such cases, neither solution is strictly superior in the Pareto sense, and selecting either does not degrade solution quality. This randomized update helps prevent bias toward a specific region of the objective space and mitigates stagnation by preserving diversity along the Pareto front. Furthermore, the global convergence behavior of the swarm remains stable, as the overall search direction is primarily governed by the leaders selected from the external repository, as described in the subsequent steps.

Step 4 (Repository-based Leader Selection). To promote exploration and maintain diversity across the Pareto front, the repository *REP* is adaptively partitioned into a set of hypercubes. This

partitioning is governed by the grid resolution parameter n_d, which specifies the number of divisions along each objective axis in the objective space. A hypercube is defined as the smallest axis-aligned cell bounded by this grid, and each Pareto-optimal solution is assigned to a specific hypercube based on its objective values $(\mathcal{F}_1, \mathcal{F}_2)$. Next, a density estimation process is performed: the number of particles within each hypercube is counted, and a fitness value is assigned to each hypercube inversely proportional to this count. In particular, sparsely populated hypercubes receive higher fitness, encouraging the algorithm to explore less crowded regions of the Pareto front. Using these fitness values, a roulette-wheel selection mechanism is applied to probabilistically select a hypercube. Each hypercube is assigned a segment on a conceptual wheel whose length is proportional to its fitness value. A random number is drawn uniformly in $[0, 1]$, and the corresponding hypercube is selected as per the location where the wheel stops. This process ensures that less-crowded hypercubes have a higher chance of being selected, promoting diversity across the front. Finally, a global best solution $\boldsymbol{x}_{rep,i}^k$ is randomly selected from within the chosen hypercube to guide the i-th particle's update in the next step.

Step 5 (Velocity and Position Update). Each particle updates its state using:

$$\boldsymbol{v}_i^{k+1} = c_0\boldsymbol{v}_i^k + c_1 r_{2,i}^k(\boldsymbol{x}_{p,i}^k - \boldsymbol{x}_i^k) + c_2 r_{3,i}^k(\boldsymbol{x}_{rep,i}^k - \boldsymbol{x}_i^k), \tag{4.6}$$

$$\boldsymbol{x}_i^{k+1} = \boldsymbol{x}_i^k + \boldsymbol{v}_i^{k+1}, \tag{4.7}$$

where c_0, c_1, and c_2 are PSO coefficients, and $r_{2,i}^k$, $r_{3,i}^k \in [0, 1]$ are independent scalar random variables sampled uniformly at each iteration for each particle.

Step 6 (Quantum-inspired Mutation). With a predefined fraction q_r, select a subset of particles to undergo quantum-inspired mutation. The new position for each selected particle is generated as follows:

$$\boldsymbol{x}_i^{k+1} = r_{4,i}^k\boldsymbol{x}_{p,i}^k + (1 - r_{4,i}^k)\boldsymbol{x}_{rep,i}^k \pm \beta|\boldsymbol{x}_m^k - \boldsymbol{x}_i^{k+1}| \ln\frac{1}{r_{5,i}^k}, \quad \boldsymbol{x}_m^k = \frac{1}{n_p}\sum_{j=1}^{n_p}\boldsymbol{x}_{p,j}^k, \tag{4.8}$$

where $r_{4,i}^k$ and $r_{5,i}^k$ are scalar random variables sampled uniformly in $[0, 1]$ and $(0, 1]$, respectively. The operator $\pm$ is applied with equal probability (i.e., 50%), and β is a tuning parameter that controls the mutation magnitude. Here, $\boldsymbol{x}_b^k := r_{4,i}^k\boldsymbol{x}_{p,i}^k + (1 - r_{4,i}^k)\boldsymbol{x}_{rep,i}^k$ denotes the base position generated from a linear combination of the personal and global bests, which is computed first and then perturbed using the quantum-inspired offset.

Step 7 (Boundary Enforcement). Maintain feasibility by clamping updated particle positions within the prescribed search boundaries:

$$x_{i,j}^{k+1} = \begin{cases} x_{i,j}^{k+1}, & \text{if } x_{min,j} \le x_{i,j}^{k+1} \le x_{max,j}, \\ x_{min,j}, & \text{if } x_{i,j}^{k+1} < x_{min,j}, \\ x_{max,j}, & \text{if } x_{i,j}^{k+1} > x_{max,j}. \end{cases} \tag{4.9}$$

Step 8 (Termination and Output). If the termination criterion (e.g., $k = k_{max}$) is satisfied, return the Pareto-optimal solutions stored in REP; otherwise, increment k and repeat from Step 2. Although

various termination strategies exist for metaheuristic algorithms, a predefined maximum iteration criterion is adopted here as a standard and reproducible stopping rule. The maximum iteration count is selected conservatively to ensure sufficient convergence, in line with common practice, rather than being adaptively adjusted within the optimization process.

This structured multi-objective optimization framework leverages the complementary characteristics of the conventional and modified ν-gap metrics. By incorporating both objectives into a unified PSO-driven search process, the algorithm facilitates the generation of reduced-order models that effectively balance frequency-domain robustness and time-domain fidelity. Consequently, this approach provides a practical and theoretically grounded solution for control-oriented model reduction tasks that require both stability assurance and accurate dynamic representation.

4.2. Benchmark validation and trade-off analysis using Pareto-based model reduction

To validate the effectiveness of the proposed multi-objective optimization framework, three benchmark systems are considered from the standard collection of test models presented in [27]. It is worth noting that the benchmark systems considered in this study—the multi-story building model, the International Space Station (ISS) model, and the clamped beam—are well-established finite-element-based representations of real mechanical structures. These models rely on the standard small-deformation assumption around an equilibrium configuration, under which the nonlinear physical dynamics can be accurately approximated by high-order linear LTI models. This assumption provides the physical justification for using these systems as high-fidelity linear benchmarks in model reduction studies. Each case involves approximating a high-order continuous-time linear model by a reduced-order version using the MOPSO algorithm. The resulting models are evaluated in terms of both frequency-domain metrics, namely, the conventional ν-gap δ_ν and the modified metric ν_{mod}, as well as time-domain step response errors measured using ISE and IAE. The modified ν-gap metric, ν_{mod}, was computed using 5000 uniformly spaced frequency samples over the interval $[10^{-5}, 10^5]$ rad/s. For time-domain performance evaluation, the unit-feedback step response was simulated using a sampling interval of 0.01 seconds. The MOPSO design parameters were configured as follows: population size $n_p = 3000$, repository size $n_r = 200$, maximum iteration count $k_{max} = 2000$, PSO coefficients $\{c_0, c_1, c_2\} = \{0.7298, 1.4962, 1.4962\}$, and quantum mutation scaling factor $\beta = 1.5$. The search space for each model coefficient was uniformly bounded as $\boldsymbol{x}_{min} = [-10^{10}, \ldots, -10^{10}]^T$ and $\boldsymbol{x}_{max} = [10^{10}, \ldots, 10^{10}]^T$, applied elementwise to all decision variables. The parameter settings of the MOPSO algorithm used for each benchmark problem are summarized in Table 5.

Table 5. MOPSO parameters and their settings for different benchmark problems.

Symbol	Description	Value		
		Building	ISS	Clamped beam
c_0	Inertia coefficient	0.7298		
c_1	Cognitive coefficient	1.4962		
c_2	Social coefficient	1.4962		
β	Quantum scaling factor	1.5		
n_r	Repository size	200		
$x_{min,j}, x_{max,j}$	Lower and upper bounds of the j-th component of the decision variable	$-10^{10}, 10^{10}$		
n_p	Population size	3000	20000	20000
k_{max}	Maximum number of iterations	2000	10000	10000

The first example concerns a 48th-order continuous-time building model characterized by lightly damped structural dynamics. The system's Hankel singular values are listed in descending order: for $i \in [1, 2, \cdots, 48]$,

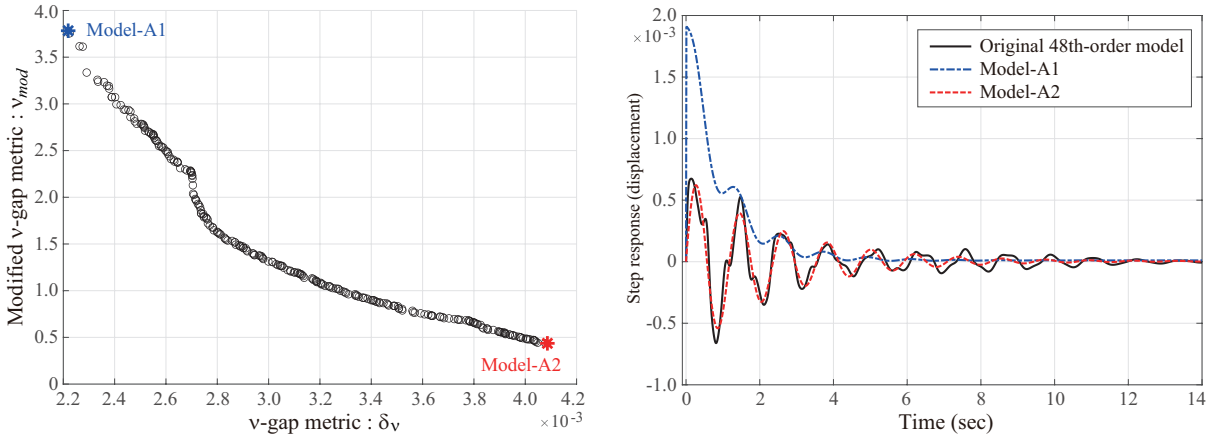
$$\sigma_i = (0.25,\ 0.24,\ 0.19,\ 0.19,\ 0.07,\ 0.07,\ 0.06, \dots) \times 10^{-2}. \tag{4.10}$$

These values indicate that a fourth-order approximation is appropriate, and this order is adopted as the target model order. The proposed algorithm then performs parameter optimization under this fixed order, ensuring both fidelity and consistency with the dominant dynamics indicated by the HSV profile. The Pareto front generated by the MOPSO procedure is shown in Figure 6(a), illustrating the inherent trade-off between δ_ν and ν_{mod}. From the Pareto set, two representative reduced-order models are extracted for further analysis:

$$\text{Model-A1: } R_{A1}(s) = \frac{6.235s^3 + 13.38s^2 + 179s + 1.027}{s^4 + 3256s^3 + 7412s^2 + 98060s + 95950}, \tag{4.11}$$

$$\text{Model-A2: } R_{A2}(s) = \frac{0.02247s^3 + 12.75s^2 + 5.663s + 0.2755}{s^4 + 3245s^3 + 5227s^2 + 93710s + 72410}. \tag{4.12}$$

As summarized in Table 6, although Model-A1 has a smaller δ_ν, Model-A2 significantly outperforms it in ν_{mod} and time-domain accuracy. Model-A2 achieves a significantly lower ν_{mod} and outperforms Model-A1 in terms of both ISE and IAE, indicating better time-domain fidelity. This conclusion is visually supported by the step responses in Figure 6(b), where Model-A2 replicates the original system dynamics more accurately. This example illustrates the importance of jointly optimizing δ_ν and ν_{mod}. A selection based solely on the conventional ν-gap could erroneously favor Model-A1, overlooking its inferior transient performance. By contrast, the proposed MOPSO framework enables a comprehensive evaluation of trade-offs and guides the selection of reduced-order models that balance robustness with dynamic fidelity, depending on the specific needs of the application.



(a) Pareto front of reduced-order models for the building system

(b) Unit-feedback step responses of the original and reduced-order models

Figure 6. Reduced-order models for the building system obtained using the proposed MOPSO algorithm.

Table 6. Performance comparison of representative reduced-order models for the building system.

Models	Objective function values		Response error	
	δ_ν	ν_{mod}	ISE	IAE
$A1$	**0.0022**	3.7832	1.52×10^{-4}	0.26
$A2$	0.0041	**0.4363**	$\mathbf{5.86\times10^{-6}}$	**0.08**

The second benchmark case focuses on a high-order model of the international space station (ISS), specifically the 270th-order continuous-time system introduced in [27]. Here, model reduction is performed for the transfer function mapping the first input to the first output. On the basis of the HSV decay, the reduced order of (3.12) was set to 6, which captures the dominant dynamics while discarding negligible modes. The MOPSO algorithm was applied using the same parameter settings as before, except for an increased population size $n_p = 20{,}000$ and a longer iteration limit $k_{max} = 10{,}000$ to ensure convergence in the larger parameter space. The resulting Pareto front is depicted in Figure 7(a), showing a clear spread in both δ_ν and ν_{mod} values. Two reduced-order models, denoted as Model-B1 and Model-B2, are selected from the extremes of the front for comparative evaluation:

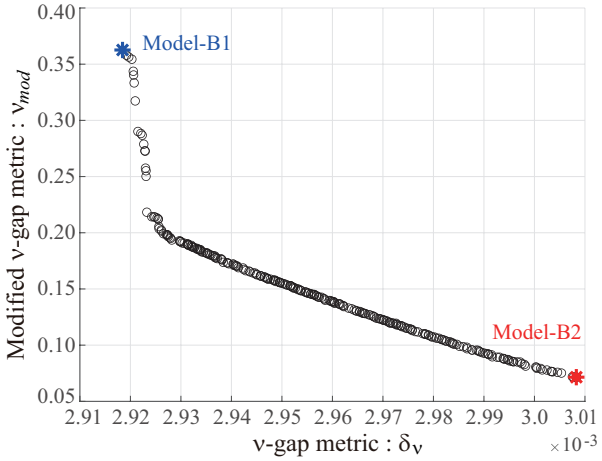
$$\text{Model-B1: } R_{B1}(s) = \frac{0.006s^5 + 0.0985s^4 + 2.266s^3 + 0.095s^2 + 5.646s - 0.158}{s^6 + 0.568s^5 + 1458s^4 + 42.53s^3 + 6633s^2 + 62.56s + 3457}, \tag{4.13}$$

$$\text{Model-B2: } R_{B2}(s) = \frac{0.0056s^5 + 0.0018s^4 + 2.287s^3 + 0.0363s^2 + 5.722s - 2.87\times10^{-6}}{s^6 + 0.41s^5 + 1448s^4 + 41.79s^3 + 6596s^2 + 62.61s + 3440}. \tag{4.14}$$

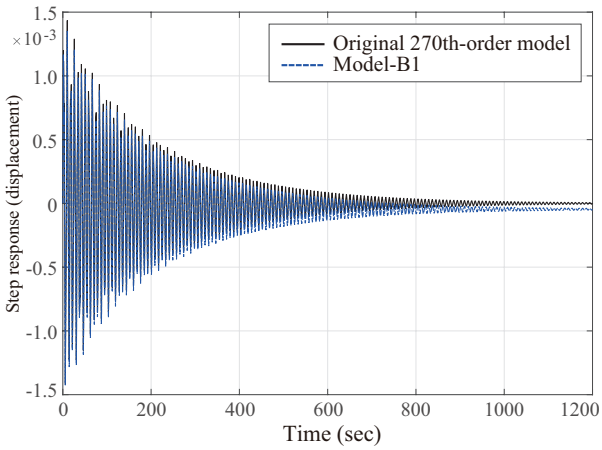
The performance comparison in Table 7 reveals that the values of δ_ν are nearly identical across the two models. However, ν_{mod} differs by a factor of almost five, and this distinction is evident in the time-domain metrics: Model-B2 exhibits dramatically lower ISE and IAE, indicating considerably better time-domain fidelity. This observation is further supported by the unit-feedback step responses presented in Figure 7(b),(c), which illustrate the time-domain performances of Model-B1 and Model-B2, respectively. The response of Model-B1 exhibits higher overshoot and slower attenuation, indicating inferior tracking behavior. By contrast, Model-B2 replicates the dynamics of the original high-order system more accurately, thereby validating its superior fidelity in both frequency and time domains. This example reinforces the central thesis of the proposed framework: while conventional ν-gap values may indicate comparable robustness, they can obscure significant differences in frequency-localized error and time-domain performance. These results further support the necessity of adopting a dual-objective framework.

Table 7. Performance comparison of representative reduced-order models for the ISS system.

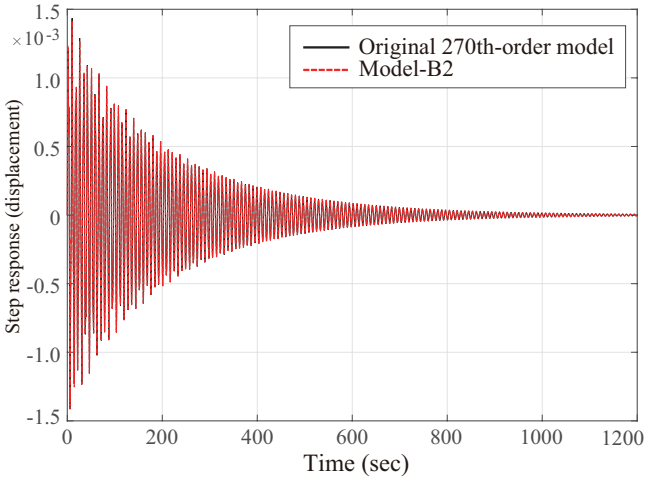
Models	Objective function values		Response error	
	δ_ν	ν_{mod}	ISE	IAE
$B1$	**0.0029**	0.3626	2.66×10^{-4}	5.5183
$B2$	0.0030	**0.0715**	$\mathbf{1.29\times10^{-6}}$	**0.1010**



(a) Pareto front of reduced-order models for the ISS system



(b) Step response comparison between the original system and Model-B1



(c) Step response comparison between the original system and Model-B2

Figure 7. Reduced-order models for the ISS system obtained using the proposed MOPSO algorithm.

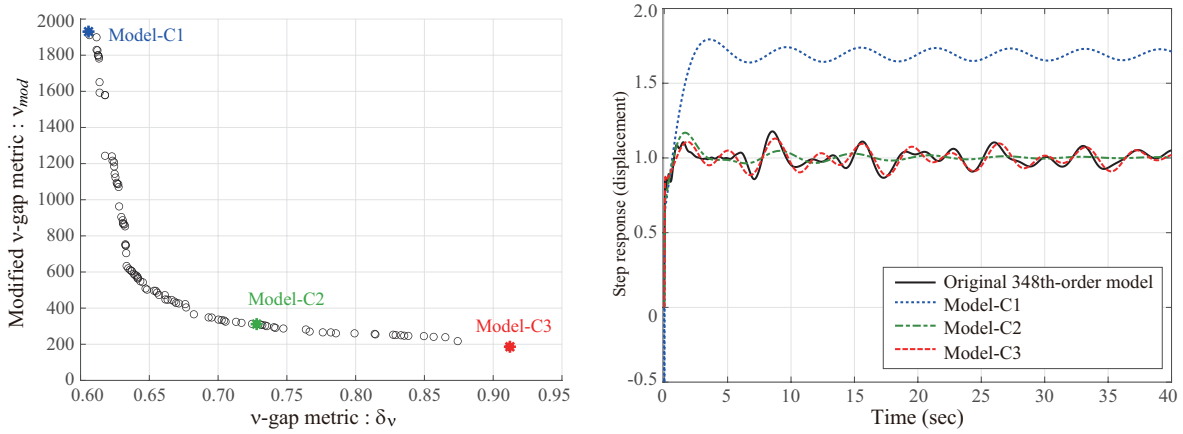
The final case addresses the reduction of a 348th-order clamped beam model introduced in [27]. Following the trend of Hankel singular values, the target order is set to six. The optimization parameters are configured as $n_p = 20{,}000$ and $k_{max} = 10{,}000$. As indicated by the Pareto front in Figure 8(a), the proposed MOPSO algorithm reveals a wide variation in the trade-off behavior between δ_ν and ν_{mod}. Three representative reduced-order models are selected from the Pareto set for further analysis:

$$\text{Model-C1: } R_{C1}(s) = \frac{-1255s^5 + 13000s^4 + 34390s^3 + 65130s^2 + 43020s + 58300}{s^6 + 1599s^5 + 11560s^4 + 5591s^3 - 6030s^2 + 2031s - 23830}, \tag{4.15}$$

$$\text{Model-C2: } R_{C2}(s) = \frac{-11.7s^5 + 3138s^4 + 5188s^3 + 14750s^2 + 7382s + 12210}{s^6 + 123.7s^5 + 1554s^4 - 86.38s^3 + 2144s^2 - 826.9s - 30.85}, \tag{4.16}$$

$$\text{Model-C3: } R_{C3}(s) = \frac{5.317s^5 + 1911s^4 + 239.5s^3 + 9040s^2 + 382.6s + 8067}{s^6 + 65.94s^5 + 355.5s^4 + 202.4s^3 + 680.5s^2 + 124.3s + 24.83}. \tag{4.17}$$

As summarized in Table 8, Model-C1 yields the smallest value of the conventional ν-gap; however, it suffers from excessive response error, indicating poor dynamic fidelity. By contrast, Model-C3 shows the smallest modified ν-gap and the lowest ISE and IAE values, showcasing its superior performance in terms of time-domain accuracy. Model-C2 strikes a balance between robustness and fidelity, making it a practical choice under moderate performance constraints. The unit-feedback step responses in Figure 8(b) clearly illustrate the difference in transient behavior among the models. Compared to Model-C1, both Model-C2 and Model-C3 achieve significantly improved tracking of the original system response. This final example underscores the utility of the Pareto front for flexible model selection as per design priorities. For instance, if the ν-gap distance between the original and reduced model is smaller than the robustness margin $b_{P,C}$ of a predesigned controller (e.g., $b_{P,C} = 0.75$), then Model-C2 can be safely adopted without violating the stability guarantee, because $\delta_\nu < b_{P,C}$ ensures closed-loop stability as stated in Proposition 1.



(a) Pareto front of reduced-order models for the clamped beam system

(b) Unit-feedback step responses of the original and reduced-order models

Figure 8. Reduced-order models for the clamped beam system obtained using the proposed MOPSO algorithm.

Table 8. Performance comparison of representative reduced-order models for the clamped beam system.

Models	Objective function values		Response error	
	δ_ν	ν_{mod}	ISE	IAE
$C1$	**0.6058**	1930.6	2843.9	4083.6
$C2$	0.7280	311.6	14.0	235.1
$C3$	0.9121	**186.3**	**4.9**	**140.1**

In summary, the benchmark examples confirm that the proposed MOPSO framework not only delivers accurate reduced-order models but also provides clear insights into trade-offs between robustness and fidelity, facilitating informed model selection tailored to application-specific performance and stability requirements.

Remark 2. *In the numerical examples of this section, a MOPSO framework is employed to construct reduced-order models that jointly minimize the conventional ν-gap and the modified ν_{mod} metrics under the WNC. Compared to the single-objective setting in Section 3.2, the admissible solution space is further constrained by the simultaneous satisfaction of multiple performance objectives and feasibility requirements, leading to a highly fragmented and irregular Pareto-feasible region. To ensure reliable exploration and stable approximation of the Pareto front, the algorithmic parameters were selected based on consistent design principles rather than being tuned for individual benchmark cases. In particular, a relatively large swarm size and a fixed repository structure were adopted to preserve a sufficient number of nondominated solutions and to mitigate the premature loss of diversity induced by the WNC constraint. While parameters such as swarm size were adjusted to accommodate differences in search dimensionality across the numerical examples, these adjustments were not intended as case-specific tuning. The particle update coefficients were fixed to the constriction-factor-based configuration* $\{c_0, c_1, c_2\} = \{0.7298, 1.4962, 1.4962\}$*, following classical PSO stability guidelines that balance exploration and convergence. Similarly, the quantum-infusion parameter was set to a robust default value of* $\beta = 1.5$*, consistent with the original CNT-PSO-QI framework, to maintain a balance between global exploration and convergence stability to maintain a balanced influence of global exploration and convergence stability in the multi-objective setting. Overall, the adopted parameter configuration reflects a conservative design choice aimed at clearly demonstrating the effectiveness, numerical stability, and robustness of the proposed multi-objective ν-gap-based model reduction framework under stringent feasibility constraints. The increased computational cost arising from this conservative setup was accepted as a deliberate trade-off for the improved reliability and interpretability of the obtained Pareto-optimal solutions.*

Remark 3. *In the multi-objective numerical examples of Section 4.2, particle positions are initialized randomly within the prescribed search bounds. Although such randomness may potentially influence convergence behavior in population-based optimization, the proposed MOPSO framework exhibited stable and consistent solution patterns across all benchmark cases considered in this study. From a numerical perspective, the underlying CNT-PSO-QI search mechanism, which constitutes the core engine of the proposed MOPSO framework, demonstrated highly consistent convergence behavior under random initialization, as verified in Section 3.2 (Table 4). This numerical robustness of the base optimizer provides a stable foundation for the multi-objective extension. From a structural perspective, the multi-objective framework employs an external repository that continuously preserves high-quality nondominated solutions identified during the search process. As the optimization proceeds, swarm evolution becomes increasingly guided by these archived leaders rather than by the initial random particle distribution. This repository-guided mechanism effectively mitigates sensitivity to initialization randomness and promotes consistent convergence toward well-distributed Pareto sets. While a rigorous theoretical convergence proof for stochastic multi-objective PSO variants with repository-based leader selection remains an open problem, the smooth and consistent Pareto-front patterns presented in Section 4.2 (Figures 6–8) provide strong empirical evidence that the proposed MOPSO framework maintains stable convergence behavior in practice, despite random sampling initialization.*

5. Conclusions

This study proposed a comprehensive framework for MOR based on the ν-gap metric and its frequency-sensitive enhancement. The primary objective was to generate reduced-order models that preserve both robustness characteristics and the dynamic fidelity of the original high-order system. Although the conventional ν-gap metric offers a powerful tool to assess closed-loop stability under model uncertainty, it frequently fails to capture discrepancies in time-domain performance. To address this inadequacy, a modified ν-gap metric was developed to better reflect frequency-wise deviations in the system response. To solve the resulting nonconvex optimization problem, a PSO algorithm incorporating quantum-inspired perturbation (CNT-PSO) was introduced. This method improves search diversity and convergence reliability while systematically enforcing the WNC to ensure feasibility in closed-loop control settings. Comparative results against existing methods validated the algorithm's capability to generate reduced models with significantly improved optimization consistency and objective function values. Building on the single-objective formulation, a multi-objective optimization algorithm (MOPSO-QI) was proposed to simultaneously minimize the conventional ν-gap and the modified metric. This dual-objective approach enables a balanced trade-off between robustness and fidelity. The resulting Pareto front offers a spectrum of candidate models, empowering users to select an appropriate reduced-order model according to the specific stability margin or response accuracy required by the application. In particular, the framework enables controller-aware model selection by verifying that the ν-gap distance does not exceed the robustness margin $b_{P,C}$ of a given predesigned controller. The effectiveness of the proposed framework was demonstrated through benchmark simulations involving high-order systems such as a building model, the ISS, and a clamped beam. The results confirmed that the proposed methods consistently outperformed existing ν-gap-based techniques in both accuracy and robustness. By explicitly revealing the trade-off structure and offering feasible models with tunable properties, the proposed framework provides a flexible and practically valuable solution to the model reduction problem in control-oriented applications.

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

This research was supported by a National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. 2021R1A2C1008686), and a Chung-Ang University Graduate Research Scholarship in 2024.

Conflict of interest

The authors declare there are no conflicts of interest.

References

1. R. Kumar, D. Ezhilarasi, A state-of-the-art survey of model order reduction techniques for large-scale coupled dynamical systems, *Int. J. Dyn. Control*, **11** (2023), 900–916. https://doi.org/10.1007/s40435-022-00985-7
2. S. K. Suman, A. Kumar, Investigation and implementation of model order reduction technique for large scale dynamical systems, *Arch. Comput. Methods Eng.*, **29** (2022), 3087–3108. https://doi.org/10.1007/s11831-021-09690-8
3. K. Lu, K. Zhang, H. Zhang, X. Gu, Y. Jin, S. Zhao, et al., A review of model order reduction methods for large-scale structure systems, *Shock Vib.*, **2021** (2021), 6631180. https://doi.org/10.1155/2021/6631180
4. O. M. Agudelo, C. Vermeersch, B. De Moor, Globally optimal H_2-norm model reduction: A numerical linear algebra approach, *IFAC-PapersOnLine*, **54** (2021), 564–571. https://doi.org/10.1016/j.ifacol.2021.06.117
5. S. Lagauw, O. M. Agudelo, B. De Moor, Globally optimal SISO H_2-norm model reduction using Walsh's theorem, *IEEE Control Syst. Lett.*, **7** (2023), 1670–1675. https://doi.org/10.1109/LCSYS.2023.3276299
6. L. Yu, J. Xiong, H_2 model reduction for negative imaginary systems, *Int. J. Control*, **93** (2020), 588–598. https://doi.org/10.1080/00207179.2018.1482501
7. Y. Sakai, T. Wada, Y. Fujisaki, Mixed H_2/H_∞ balanced truncations for discrete time linear systems, in *2019 12th Asian Control Conference (ASCC)*, (2019), 301–306.
8. M. A. Hamadi, K. Jbilou, A. Ratnani, A data-driven Krylov model order reduction for large-scale dynamical systems, *J. Sci. Comput.*, **95** (2023), 2. https://doi.org/10.1007/s10915-023-02122-8
9. A. Bunse-Gerstner, D. Kubalińska, G. Vossen, D. Wilczek, H_2-norm optimal model reduction for large scale discrete dynamical MIMO systems, *J. Comput. Appl. Math.*, **233** (2010), 1202–1216. https://doi.org/10.1016/j.cam.2008.12.029
10. M. Vidyasagar, The graph metric for unstable plants and robustness estimates for feedback stability, *IEEE Trans. Autom. Control*, **29** (1984), 403–418. https://doi.org/10.1109/TAC.1984.1103547
11. A. El-Sakkary, The gap metric: Robustness of stabilization of feedback systems, *IEEE Trans. Autom. Control*, **30** (1985), 240–247. https://doi.org/10.1109/TAC.1985.1103926
12. T. T. Georgiou, On the computation of the gap metric, *Syst. Control Lett.*, **11** (1988), 253–257. https://doi.org/10.1016/0167-6911(88)90067-9
13. J. H. Steele, G. Vinnicombe, Closed-loop time-domain model validation in the nu-gap metric, in *Proceedings of the 40th IEEE Conference on Decision and Control*, (2001), 4332–4337. https://doi.org/10.1109/CDC.2001.980882
14. S. Taamallah, Nu-gap metric a sum-of-squares and linear matrix inequality approach, in *2014 18th International Conference on System Theory, Control and Computing (ICSTCC)*, (2014), 405–411. https://doi.org/10.1109/ICSTCC.2014.6982450

15. G. Buskes, M. Cantoni, Reduced order approximation in the ν-gap metric, in *2007 46th IEEE Conference on Decision and Control*, (2007), 4367–4372. https://doi.org/10.1109/CDC.2007.4434557
16. M. Cantoni, On model reduction in the ν-gap metric, in *Proceedings of the 40th IEEE Conference on Decision and Control*, (2001), 3665–3670. https://doi.org/10.1109/CDC.2001.980431
17. H. Kondo, Y. Ochi, I-PD flight controller design based on integral-type optimal servomechanism, *J. Control Meas. Syst. Integr.*, **3** (2010), 448–455. https://doi.org/10.9746/jcmsi.3.448
18. H. Kondo, Y. Ochi, Plant-model reduction for PID controller design, in *49th IEEE Conference on Decision and Control (CDC),* (2010), 6201–6206. https://doi.org/10.1109/CDC.2010.5717005
19. A. Sootla, Nu-gap model reduction in the frequency domain, *IEEE Trans. Autom. Control*, **59** (2014), 228–233. https://doi.org/10.1109/TAC.2013.2290817
20. S. Taamallah, Model order reduction in the $\mathcal{L}_2$-gap metric, in *2016 American Control Conference (ACC)*, (2016), 2453–2459. https://doi.org/10.1109/ACC.2016.7525285
21. Y. Hwang, Y. Ko, Y. Lee, T. Kim, Frequency-domain tuning of robust fixed-structure controllers via quantum-behaved particle swarm optimizer with cyclic neighborhood topology, *Int. J. Control Autom. Syst.*, **16** (2018), 426–436. https://doi.org/10.1007/s12555-016-0766-3
22. S. M. Djouadi, On robustness in the gap metric and coprime factor uncertainty for LTV systems, *Syst. Control Lett.*, **80** (2015), 16–22. https://doi.org/10.1016/j.sysconle.2015.03.006
23. A. Mondal, S. Ganguli, P. Sarkar, Reduced order modeling of delta operator systems by optimal frequency fitting approach, *IFAC J. Syst. Control*, **27** (2024), 100240. https://doi.org/10.1016/j.ifacsc.2024.100240
24. E. H. Houssein, M. K. Saeed, G. Hu, M. M. Al-Sayed, Metaheuristics for solving global and engineering optimization problems: Review, applications, open issues and challenges, *Arch. Comput. Methods Eng.*, **31** (2024), 4485–4519. https://doi.org/10.1007/s11831-024-10168-6
25. D. Freitas, L. G. Lopes, F. Morgado-Dias, Particle swarm optimisation: A historical review up to the current developments, *Entropy*, **22** (2020), 362. https://doi.org/10.3390/e22030362
26. G. Vinnicombe, Frequency domain uncertainty and the graph topology, *IEEE Trans. Autom. Control*, **38** (1993), 1371–1383. https://doi.org/10.1109/9.237648
27. Y. Chahlaoui, P. Van Dooren, A collection of benchmark examples for model reduction of linear time invariant dynamical systems, 2002. Available from: https://eprints.maths.manchester.ac.uk/1040/.

AIMS Press