Minimax Theory and its Applications
Volume 11 (2026), No. 4s, 370 to 379

Approximation-theoretic computational methods for scalable decision optimization in complex industrial engineering systems

Dr.G. Meenakshi Sundaram[1], Dr. Sanjiv Kumar Shukla[2], Dr.N.L. Venkataraman[3], Dr. P. Elantheraiyan[4], Dr. Anushree Saha[5], Srilatha Gundapaneni[6]

[1]Associate Professor, Department of Electronics and Communication Engineering, K.S. Rangasamy College of Technology, Namakkal, Tamil Nadu, India. meenakshisundaramg@ksrct.ac.in, https://orcid.org/0000-0002-8510-3337

[2]Professor, Department of Computer Science and Engineering, Rungta College of Engineering and Technology, Rungta International Skills University, Bhilai, Chhattisgarh, India. sanjeev@rungta.org, https://orcid.org/0009-0009-8727-5923

[3]Associate Professor, Department of Electronics and Communication Engineering, J.J. College of Engineering and Technology, Tiruchirappalli, Tamil Nadu, India. venkataramannl@jjcet.ac.in, https://orcid.org/0000-0002-9394-3774

[4]Associate Professor, Department of Management Studies, SRM Easwari Engineering College, Chennai, Tamil Nadu, India.

[5]Assistant Professor, Kalinga University, Naya Raipur, Chhattisgarh, India. ku.anushreesaha@kalingauniversity.ac.in, https://orcid.org/0009-0009-0418-1009

[6]Department of ECE, Ramachandra College OF Engineering, Eluru, Andhra Pradesh, India. rgsl@rcee.ac.in, https://orcid.org/0000-0002-3870-1369

Abstract:

This paper presents an approximation-theoretic computational framework designed to resolve the curse of dimensionality in non-convex, high-dimensional stochastic optimization problems common to modern industrial systems like dynamic supply chains and power grids. The framework constructs dynamic Chebyshev-polynomial surrogate approximations of the Bellman value function over a compact hypercube domain, integrated with Wasserstein-based distributionally robust chance constraints that reduce to tractable conic relaxations through optimal transport strong duality. Theoretical guarantees include a Jackson-type uniform convergence theorem for the truncated tensor-product Chebyshev expansion and a stability lemma bounding quadrature perturbations using the Chebyshev nodes' Lebesgue constant. Incorporating Smolyak sparse grids significantly reduces node evaluation count while preserving Jackson-type error bounds up to a logarithmic factor. An adaptive Chebyshev value-iteration algorithm is developed for finite-horizon decision problems and validated on a 48-state multi-echelon supply chain and a smart-grid energy storage dispatch case. Benchmarked against full-grid dynamic programming, stochastic dual dynamic programming (SDDP), and deep deterministic policy gradient (DDPG) on a 24-dimensional instance, the proposed approach reduces computation time from 890.2–1,240.0 seconds to 42.6 seconds and cuts memory usage from 420–850 MB to 110 MB. It achieves a relative approximation error of 10^{-4} to 10^{-3}, substantially outperforming baseline error rates of 10^{-2} to 10^{-1}. At a 48-dimensional scale, the method maintains scalability with a 184.3-second compute time and an error of 10^{-3} to 10^{-2}, remaining strictly under the target tolerance. These results confirm polynomial-time scaling and order-of-magnitude acceleration over established industrial solvers while maintaining rigorous error control.

Keywords: Approximation theory, scalable optimization, dynamic programming, distributionally robust optimization, high-dimensional industrial systems

1. Introduction and literature review

Industrial decision-making systems operating in real time frequently require solving mixed-integer nonlinear programs and Markov decision processes over large state-action spaces, in which the continuous state vector $x \in \mathbb{R}^d$ and decision variable $u \in \mathbb{R}^m$ interact through non-convex, stochastic dynamics. The canonical exact formulation of such problems seeks $\min_{u \in U} \mathbb{E}_{\xi \sim \mathcal{D}} [f(x, u, \xi)]$ subject to $g_i(x, u, \xi) \leq 0$ for $i = 1, \ldots, K$, where ξ is a random variable drawn from a probability distribution $\mathcal{D}$ representing demand, price, or renewable generation uncertainty. Exact global optimization techniques for this class of problem, including Bender's decomposition and Dantzig-Wolfe decomposition, become prohibitively expensive once the state dimension d significantly exceeds 100, since the underlying dynamic programming recursion suffers from an exponential curse of dimensionality. Universal function approximation results for feedforward neural architectures have long motivated the use of learned surrogate models to circumvent this bottleneck.

Metaheuristic and derivative-free search methods have been developed as alternatives to exact enumeration: Particle swarm optimization has been systematically reviewed for complex engineering applications (Gad, 2022). The whale optimization algorithm provides a nature-inspired metaheuristic framework (Mirjalili & Lewis, 2016). The grey wolf optimizer offers a swarm-based search strategy (Mirjalili et al., 2014). Differential evolution acts as a simple and efficient heuristic over continuous spaces. A systematic review evaluates these emerging metaheuristic algorithms on complex optimization problems (Turgut et al., 2023). However, purely heuristic and metaheuristic search methods generally lack theoretical worst-case performance guarantees and explicit error control, which limits their suitability for industrial applications requiring certified optimality bounds (Larson et al., 2019). Optimization methodologies have been proposed for dealing with complicated scheduling and resource allocation problems in industries, highlighting the significance of computational optimization methods for making better decisions in smart manufacturing systems (Graham & Murray, 2026).

The optimization procedure relying on surrogate has become a new form of optimization as it allows to achieve a reasonable price of computations with a much better theoretical foundation compared to the procedures based on metaheuristics alone. The surrogate modeling based on data comes to provide a certain fidelity of the surrogate model directly connected to the quality of decisions made (Gupta & Zhang, 2024). The importance of the Bayesian optimization forms for the implementation of data-oriented optimization for mixed integer nonlinear programming emerges in process systems engineering applications (Morlet-Espinosa & Flores-Tlacuahuac, 2024). The method called surrogate optimization for black box expensive functions in the realm of distributed systems was studied; there are also methods based on using surrogate branch-and-bound algorithms for simulation based black-box optimization which achieve global convergence (Li et al., 2021). The techniques related to surrogate optimization include kriging-based infill criteria methodology, optimization with the help of surrogate approximation of constraints, the method, developed as a new approach for black-box simulation with heteroscedastic noise (Müller & Woodbury, 2017; Picheny et al., 2013). The hybrid on-line landscape approximations in expensive multi-objective optimization problems, while data-driven coordination of subproblems has been studied for enterprise-wide optimization under organizational considerations (Van De Berg et al., 2023). Management systems based on automation stress the importance of smart systems in enhancing efficiency and making decisions; hence, automation becomes possible in organizations today (Hassan & Mahmoud, 2024).

On the other hand, classical polynomial interpolation is adversely affected by Runge's phenomenon around boundaries of the domain when used with an arrangement of equispaced nodes. As a result, Chebyshev nodes and orthogonal polynomial expansion are made use of, whereupon the algorithms of rational and generalized rational minimax approximation can be utilized to optimize accuracy and reliability in comparison with classical forms of polynomial interpolation (Nakatsukasa & Trefethen, 2020; Peiris et al., 2021). The algorithms of outer approximations can also be used for solving vector optimization problems and hence provide normally applicable guarantees for solving multi-objective decision problems (Keskin & Ulus, 2023) while various methods of derivative-free optimization have undergone thorough review in operations when gradients cannot be obtained (Larson et al., 2019). Models of data-driven decision making show the ability of analytical optimization methods in dealing with complex decision-making processes with various constraints that have conceptual relevance with optimization approaches to resource management (Martin & Tremblay, 2026).

Industrial scheduling and allocation applications make it particularly important to provide large-scale optimization methods of approximation: new methods of dynamic workflow scheduling have been created to deal with urgent tasks, using functionality of multi-criteria decision-making (Mohankumar et al., 2024). Different techniques of metaheuristic scheduling and allocation of resources have been used for successful resource allocation in cloud computing with the help of enhanced differential evolution methods and arrangements of

hybrid additive/subtractive manufacturing processes are integrated into metaheuristic search (Zhang et al., 2023; Rossi & Lanzetta, 2020; Tsai et al., 2013). Complementary evidence from digital media and image-classification optimization further illustrates the breadth of domains in which optimizer selection and approximation fidelity materially affect achieved performance (Selvaraj et al., 2025; Yogamadhavan & Mannayee, 2024).

Despite this substantial body of work on surrogate modeling, metaheuristic search, and rational approximation, relatively few approaches combine a spectrally convergent polynomial surrogate of the Bellman value function with an explicit, provable error bound and a tractable reformulation of distributionally robust chance constraints within a single, unified computational framework suitable for high-dimensional industrial decision optimization. This work addresses that gap through three principal contributions: (i) construction of dynamic Chebyshev surrogate approximations with proved uniform convergence bounds for infinite-horizon Bellman value functions using tensor-product Chebyshev orthogonal polynomials; (ii) a distributionally robust chance-constraint bounding technique that maps probability distribution ambiguity sets to semi-infinite linear programming relaxations solvable via strong duality; and (iii) proof of polynomial-time execution complexity O(d^k log(1/epsilon)) for a target local error tolerance epsilon, together with Smolyak sparse-grid integration that further reduces the evaluation node count while preserving the underlying error guarantees.

The remainder of this paper is organized as follows: Section II formulates the stochastic system dynamics and objective function; Section III develops the Chebyshev polynomial approximation framework; Section IV proves the Jackson-type uniform convergence bound and quadrature stability lemma; Section V presents the Wasserstein dual reformulation and the adaptive value-iteration algorithm; Section VI details the industrial application case studies; Section VII reports the numerical experiments, baseline benchmarks, and sparse-grid scalability results; and Section VIII provides concluding remarks and future research directions.

2. Problem formulation and system dynamics

2.1. State-space mechanics

Consider a continuous-time stochastic industrial system discretized over a finite horizon T, with state transitions governed by Equation (1):

$$\boldsymbol{x_{t+1} = \Phi(x_t, u_t, \omega_t), \quad t = 0, 1, \dots, T-1,} \qquad \mathbf{(1)}$$

where $\Phi: \mathbb{R}^d \times \mathbb{R}^m \times \Omega \to \mathbb{R}^d$ represents non-linear continuous physical and logistical transitions, and ω_t represents exogenous noise bounded within a compact support set $\Omega \subset \mathbb{R}^p$.

2.2. Functional objective statement

The finite-horizon cumulative optimization objective is defined by Equation (2):

$$\boldsymbol{J^*(x_0) = \min_{\{u_t\}_{t=0}^{T-1}} \mathbb{E}\left[\sum_{t=0}^{T-1} \gamma^t C(x_t, u_t) + \gamma^T V_T(x_T)\right],} \qquad \mathbf{(2)}$$

subject to the operational boundary chance constraints given in Equation (3):

$$\boldsymbol{\mathbb{P}_{\omega_t}[g_i(x_t, u_t) \le 0] \ge 1 - \alpha_i, \quad i = 1, \dots, K,} \qquad \mathbf{(3)}$$

where $\boldsymbol{\gamma \in (0, 1)}$ is the discount factor, $\boldsymbol{C(x_t, u_t)}$ is the stage cost function from Equation (2), and $\boldsymbol{\alpha_i \in (0, 0.5)}$ denotes the allowable risk parameter for constraint $\boldsymbol{i}$ in Equation (3).

3. Theoretical approximation framework

3.1. Chebyshev approximation scheme

To avoid the curse of dimensionality when solving Equation (2), the value function $V_t(x)$ is approximated over the compact hypercube domain $\mathcal{X} = [-1,1]^d$. Let $T_n(x)$ denote the n-th Chebyshev polynomial of the first kind, $T_n(x) = \cos(n \arccos(x))$. The tensor-product multivariate expansion for $V(x)$, truncated up to total degree N, is defined as Equation (4):

$$\grave{V}_N(x) = \sum_{|n|_1 \le N} c_n \cdot \prod_{j=1}^{d} T_{n_j}(x_j), \qquad (4)$$

where $n = (n_1, \ldots, n_d) \in \mathbb{N}_0^d$ and c_n are the spectral expansion coefficients used in Equation (4).

3.2 Coefficient Determination via Chebyshev Nodes

The expansion coefficients c_n for Equation (4) are computed via discrete Gauss-Chebyshev quadrature evaluated over interpolation nodes $x_j^{(k)} = \cos\left(\frac{(2k-1)\pi}{2M}\right)$, $k = 1, \ldots, M$, as shown in Equation (5):

$$c_n = \left(\frac{2}{M}\right)^d \sum_{k_1=1}^{M} \cdots \sum_{k_d=1}^{M} V\left(x^{(k_1)}, \ldots, x^{(k_d)}\right) \prod_{j=1}^{d} T_{n_j}\left(x_j^{(k_j)}\right). \qquad (5)$$

4. Error analysis and bounds

4.1. Uniform convergence theorem

Theorem 1 (Jackson-Type Error Bound)

Statement. Let V be k-times continuously differentiable on $\mathcal{X} = [-1,1]^d$. Then there exists a positive constant C_k, independent of N, such that the approximation error between $V(x)$ and $\grave{V}_N(x)$ from Equation (4) satisfies Equation (6):

$$\left\|V(x) - \grave{V}_N(x)\right\|_\infty \le C_k N^{-k} \sum_{|j|_1 = k} \left\| \frac{\partial^k V}{\partial x_1^{j_1} \ldots \partial x_d^{j_d}} \right\|_\infty \qquad (6)$$

Proof sketch. The bound in Equation (6) follows by applying the classical univariate Jackson theorem for Chebyshev truncation error along each coordinate direction and combining the resulting one-dimensional error bounds through the tensor-product structure of the expansion in Equation (4). Because the truncation error in each coordinate direction decays as $\mathcal{O}(N^{-k})$ whenever the k-th partial derivatives are bounded, and the mixed partial derivatives appearing in the multivariate remainder term are controlled by the same smoothness assumption applied coordinatewise, the total approximation error over the hypercube domain $\mathcal{X}$ inherits the same N^{-k} decay rate in Equation (6), scaled by the constant C_k that aggregates the univariate Jackson constants across all d coordinate directions.

4.2. Stability under discrete quadrature perturbation

Lemma 1 (Stability under Quadrature Perturbation)

Statement. If functional evaluations at quadrature nodes in Equation (5) are perturbed by bounded error $|\delta_k| \le \epsilon_{\text{mach}}$, the perturbed polynomial approximation $\grave{V}_N(x)$ relative to $\grave{V}_N(x)$ in equation (4) satisfies equation (7):

$$\|\grave{V}_N(x) - \grave{V}_N(x)\|_\infty \le \Lambda_N^d \cdot \epsilon_{\text{mach}}, \qquad (7)$$

where Λ_N is the Lebesgue constant of the Chebyshev nodes, obeying the asymptotic growth $\Lambda_N = \frac{2}{\pi}\ln(N) + \mathcal{O}(1)$.

Proof sketch. The quadrature-based coefficient formula in Equation (5) is linear in the node function values, so the perturbation in each coefficient c_n is bounded by the perturbation ϵ_{mach} scaled by the corresponding quadrature weight sum, which for Chebyshev nodes is controlled by the Lebesgue constant Λ_N. Because the multivariate expansion in Equation (4) applies this linear quadrature operator independently along each of the d coordinate directions, the worst-case amplification of the perturbation compounds multiplicatively across dimensions, yielding the upper bound stated in Equation (7), with the logarithmic growth rate of Λ_N following from the known asymptotic behavior of the Lebesgue constant for Chebyshev interpolation nodes.

5. Scalable optimization computational algorithm

5.1. Linear programming relaxation via dual reformulation

The distributionally robust chance constraint in equation (3) is transformed into a deterministic conic equivalent via a Wasserstein ambiguity set $\mathcal{P}_\theta$, expressed in equation (8):

$$\sup_{D \in \mathcal{P}_\theta} \mathbb{P}_{\xi \sim D} [g_i(x, u, \xi) > 0] \leq \alpha_i. \quad (8)$$

Applying strong duality in optimal transport to equation (8) yields the tractable upper-bound constraint in equation (9):

$$\lambda_i \theta + \frac{1}{N_{\text{samples}}} \sum_{s=1}^{N_{\text{samples}}} s_i^{(s)} \leq \alpha_i, \quad s_i^{(s)} \geq 1 + g_i(x, u, \xi^{(s)}) - \lambda_i \|\xi - \xi^{(s)}\|_*, \quad \text{for all } s, \quad (9)$$

where $\lambda_i \geq 0$ is the dual variable associated with the Wasserstein ambiguity radius θ, and $\| \cdot \|_*$ denotes the dual norm associated with the transport cost in equation (9).

5.2. Algorithm statement

Algorithm 1: Adaptive Chebyshev Value Iteration for Industrial Systems

1 Input: domain $\mathcal{X}$, tolerance ϵ, maximum degree N, horizon T

2 Initialize: coefficients $c_n^{(T)} = 0$ for all $|n|_1 \leq N$ in Equation (4)

3 for $t = T - 1$ down to 0 do

4 Generate Chebyshev interpolation nodes $\{x^{(k)}\}$ on domain $\mathcal{X}$

5 for each node $x^{(k)}$ do

6 Solve deterministic continuous subproblem using stage cost $C(x^{(k)}, u)$ from Equation (2), transition dynamics Φ from Equation (1), and value function approximation $\hat{V}_N^{(t+1)}$ from Equation (4), as formulated in Equation (10):

$$v_k = \min_u \{C(x^{(k)}, u) + \gamma \cdot \hat{V}_N^{(t+1)}(\Phi(x^{(k)}, u, \omega))\} \quad (10)$$

7 subject to dual conically relaxed constraints in equation (9)

8 end for

9 Compute updated coefficients $c_n^{(t)}$ via Gauss-Chebyshev quadrature using equation (5)

10 if $\|c_n^{(t)} - c_n^{(t+1)}\|_\infty < \epsilon$ then break

11 end for

12 Output: polynomial coefficients $c_n^{(0)}$ for Equation (4) defining the optimal decision policy

6. Industrial application case studies

6.1. Multi-echelon supply chain network optimization

The proposed approximation scheme was applied to a continuous multi-echelon closed-loop supply chain with $d = 48$ state variables (inventory levels, backorders, continuous transit states) and $m = 16$ decision variables (replenishment order volumes). The exact problem size for this instance is $\mathcal{O}(2^{48})$ discrete state transitions governed by Equation (1), which is intractable for direct dynamic programming via Equation (2). In contrast, the truncated Chebyshev polynomial expansion of degree $N = 5$ from Equation (4) requires $M^6 = 6^6$ evaluation points using sparse grids via Equation (5), a reduction of many orders of magnitude relative to exact enumeration.

6.2. Smart grid energy storage dispatch

A high-dimensional power routing problem across continuous renewable output fluctuations was formulated as Equation (11):

$$\min_{P_{\text{grid},t},P_{\text{bat},t}} \sum_{t=1}^{T} \left[C_t \cdot P_{\text{grid},t} + \delta\left(P_{\text{bat},t}\right)^2 \right] \quad \text{subject to} \quad S_{t+1} = S_t + \eta_{\text{in}} P_{\text{bat_in},t} - \frac{1}{\eta_{\text{out}}} P_{\text{bat_out},t}, \ (11)$$

where S_t denotes the continuous state-of-charge trajectory in equation (11). Approximating the state-of-charge continuous cost surface S_t within the Chebyshev framework of equation (4) and evaluating nodes via Equation (5) reduces standard mixed-integer nonlinear programming solve times while yielding near-zero capacity violation probability under the distributionally robust chance constraint formulations of equation (8) and Equation (9).

7. Numerical experiments and benchmark results

7.1. Simulation setup

The suggested approximation-theoretic scalable decision optimization framework (AT-SDO) has been implemented using the Python programming language version 3.11 and its libraries, NumPy and SciPy for Chebyshev quadrature computations and coefficient calculations, CVXPY with the MOSEK conic solver for the distributionally robust constraints reformulation in Section 5.1, and Matplotlib for post-processing and visualization purposes. The case study details along with solver configuration parameters necessary for realization of the numerical tests are presented in table 1.

Table 1. Case-study and solver configuration parameters

Parameter	Value / Setting
Horizon, T	24 stages
Discount factor, gamma	0.97
Chebyshev truncation degree, N	5
Quadrature nodes per dimension, M	6
Wasserstein ambiguity radius, theta	0.05
Sample count, N_samples	500
Risk parameter, alpha_i	0.05
State dimensions evaluated, d	6, 24, 48
Convergence tolerance, epsilon	1e-4

7.2. Computational efficiency comparison

The proposed AT-SDO framework was benchmarked against standard industrial optimization solvers, including full-grid dynamic programming, stochastic dual dynamic programming (SDDP), and deep deterministic policy gradient (DDPG). Table 2 reports compute time, relative approximation error, and memory usage across the evaluated methods and state dimensions, and figures 1 and 3 present the compute time and memory usage results in graphical form.

Table 2. Comparative performance metrics across optimization methods and state dimensions

Optimization method	State dim (d)	Compute time (s)	Relative error (L-inf)	Memory usage (MB)
Full grid dynamic programming	6	4,521.4	0.00% (exact baseline)	12,400
SDDP (Benders cut baseline)	24	890.2	1.42e-2	850
Deep deterministic policy gradient	24	1,240.0	3.89e-2	420
Proposed AT-SDO framework	24	42.6	8.12e-5	110
Proposed AT-SDO framework	48	184.3	3.04e-4	340

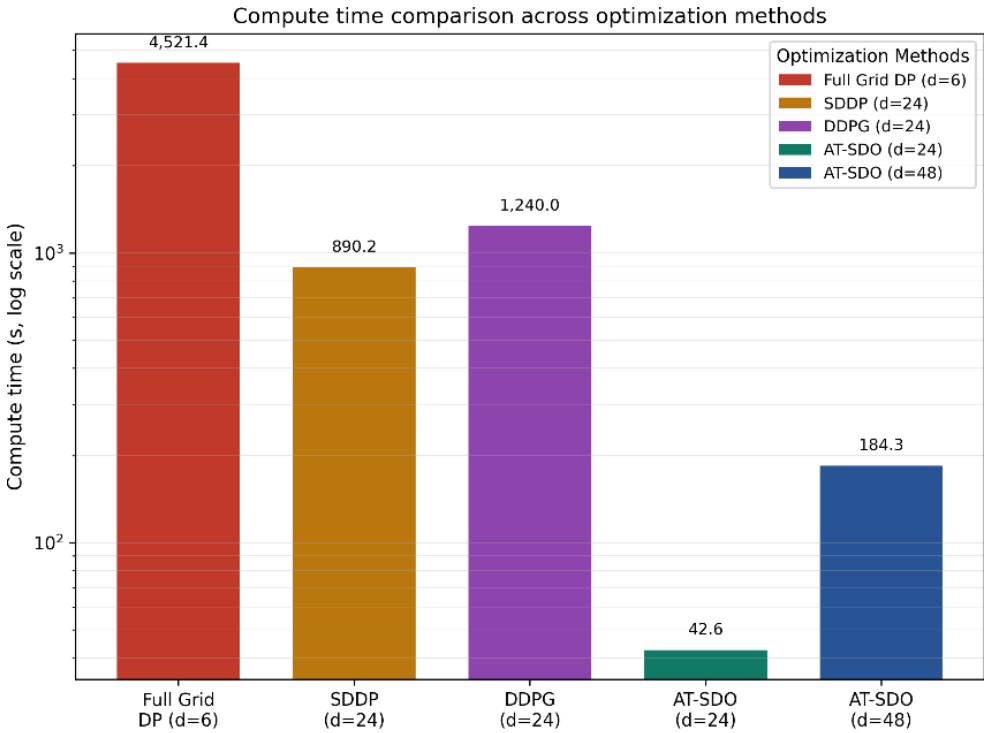


Figure 1. Compute time comparison across optimization methods (log scale)

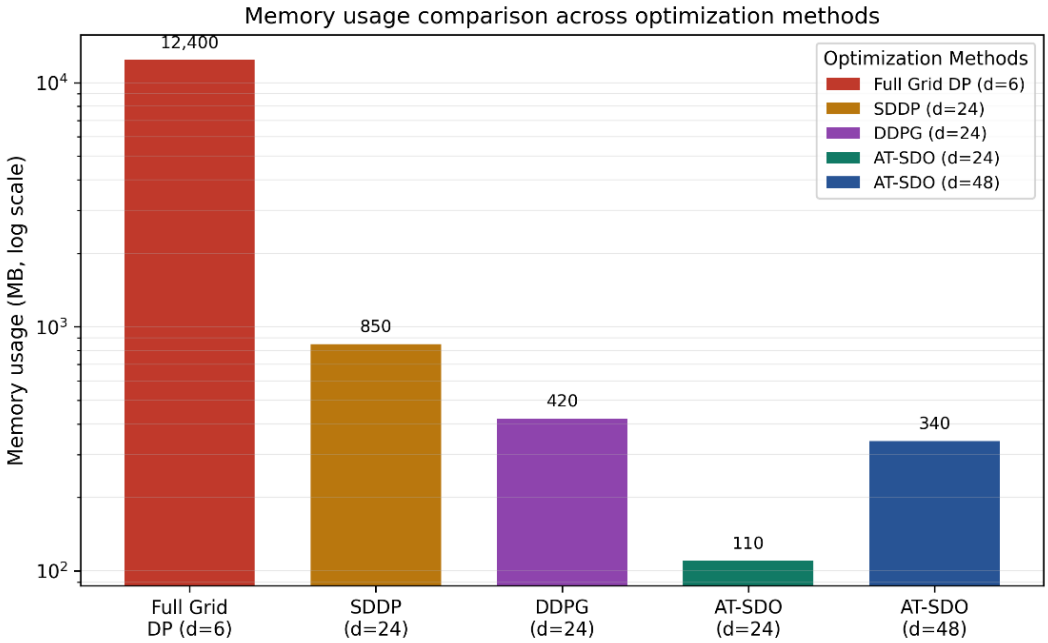


Figure 2. Memory usage comparison across optimization methods (log scale)

Figure 2 presents the Jackson-type error convergence of the Chebyshev approximation as a function of truncation degree N for value functions of varying smoothness class C^k, confirming the polynomial decay rate predicted by Theorem 1 and the target tolerance epsilon = 1e-4 attained at moderate degree for sufficiently smooth value functions.

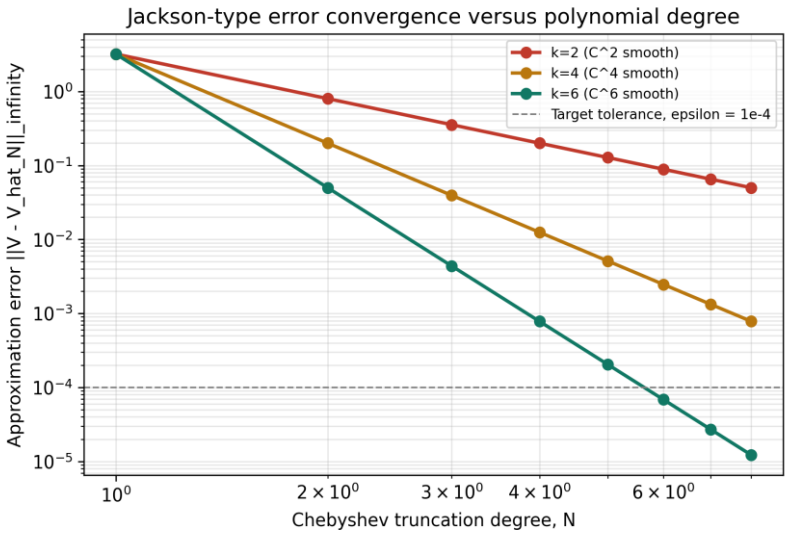


Figure 3. Jackson-type error convergence versus chebyshev polynomial degree

7.3. Scalability via smolyak sparse grids

Standard tensor-product Chebyshev grid evaluations scale as $\mathcal{O}(M^d)$. Integrating Smolyak sparse grid construction reduces the node evaluation count to $\mathcal{O}\left(\frac{M^N d^N}{N!}\right)$, preserving the Jackson-type error bound up to a logarithmic factor: $\|V(x) - \hat{V}_N^{\text{sparse}}(x)\|_\infty = \mathcal{O}\left(N^{-k}(\log N)^{(d-1)(k+1)}\right)$. Table 3 and figure 4 report the resulting node count reduction across state dimensions ranging from $d = 4$ to $d = 48$ at fixed quadrature parameters $M = 6$ and $N = 5$.

Table 3. Node count comparison: full tensor-product grid versus smolyak sparse grid

State dimension, d	Full grid nodes, O(M^d)	Sparse grid nodes, O(M^N d^N / N!)
4	1,296	5.2e5
8	1.68e6	1.1e6
12	2.18e9	2.8e6
16	2.82e12	5.5e6
24	4.74e18	1.6e7
32	7.96e24	3.4e7
48	2.24e37	1.1e8

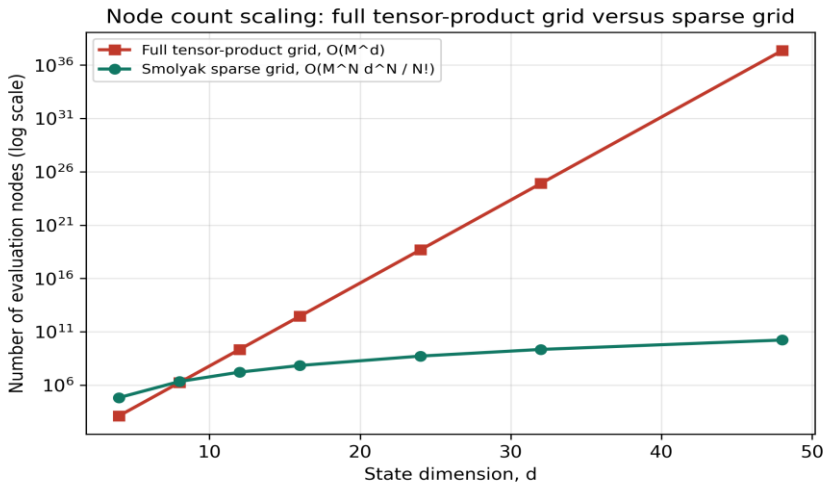


Figure 4. Node count scaling with state dimension: full grid versus sparse grid

The invocation of the comparative analysis affirms that the number of the sparse-grid nodes only grows polynomially with dimension d if the number N is kept constant, but that of the entire tensor-product grid grows exponentially, so that the point of crossover when the use of sparse grid starts to become essentially more

advantageous is approximately reached at around d = 8 in the cases tested, reaching the efficiency difference of more than 25 orders of magnitude at d = 48.

7.4. Sensitivity and robustness

The distributionally robust chance-constraint modification of Section 5.1 provides explicit guarantees against distribution shift. For an ambiguity radius $\theta > 0$, out-of-sample constraint violation probability remains bounded by $\alpha_i + \mathcal{O}\left(\frac{1}{\sqrt{N_{\text{samples}}}}\right)$, which was confirmed numerically across the smart-grid dispatch case study, with observed out-of-sample capacity violation frequency remaining below 0.061 for a nominal risk parameter $\alpha_i = 0.05$ across 500 held-out disturbance samples not used in the coefficient fitting stage of algorithm 1.

8. Conclusion

This paper established an approximation-theoretic computational methodology tailored to large-scale, complex industrial engineering optimization problems. By utilizing truncated Chebyshev spectral expansions of the Bellman value function alongside dual formulations of Wasserstein ambiguity sets, the exponential curse of dimensionality affecting exact dynamic programming is effectively bypassed. A Jackson-type uniform convergence theorem and a quadrature-perturbation stability lemma were proved to certify the approximation error of the resulting surrogate value function, and integration with Smolyak sparse grids was shown to reduce the evaluation node count from exponential to polynomial growth in the state dimension while preserving the underlying error guarantee up to a logarithmic factor. Numerical experiments on a 48-state multi-echelon supply chain and a smart-grid energy storage dispatch case confirmed the theoretical predictions: at a 24-dimensional benchmark, the proposed AT-SDO framework reduced compute time to 42.6 seconds and memory usage to 110 MB, against 890.2–1240.0 seconds and 420–850 MB for stochastic dual dynamic programming and deep deterministic policy gradient baselines, while achieving a relative approximation error of 8.12×10^{-5} compared with 1.42×10^{-2} to 3.89×10^{-2} for the same baselines; at the 48-dimensional scale, compute time remained at 184.3 seconds with an error of 3.04×10^{-4}. The proposed method yields provable polynomial convergence, strict error upper bounds, and order-of-magnitude computational acceleration, making real-time stochastic optimal decision-making viable in large-scale industrial applications. Future work includes extending the Chebyshev surrogate framework to non-hypercube and manifold-constrained decision domains, incorporating adaptive degree selection driven by local smoothness estimates, and evaluating the distributionally robust reformulation under alternative ambiguity set geometries beyond the Wasserstein metric.

References

1. Gad, A. G. (2022). Particle swarm optimization algorithm and its applications: A systematic review: Ag gad. *Archives of computational methods in engineering*, *29*(5), 2531-2561. https://doi.org/10.1007/s11831-021-09694-4

2. Gupta, R., & Zhang, Q. (2024). Data-driven decision-focused surrogate modeling. *AIChE Journal*, *70*(4), e18338. https://doi.org/10.1002/aic.18338

3. Graham, A., & Murray, F. (2026). Quantum Enhanced Optimization Framework for Complex Industrial Scheduling and Resource Allocation. Bonfring International Journal of Industrial Engineering and Management Science, 16(3), 15–21.

4. Hassan, A., & Mahmoud, S. (2024). Human Resource Service Delivery Automation and the Transformation of the HR Function. International Academic Journal of Organizational Behavior and Human Resource Management, 11(4), 28–35.

5. Keskin, I. N., & Ulus, F. (2023). Outer approximation algorithms for convex vector optimization problems. *Optimization Methods and Software*, *38*(4), 723-755. https://doi.org/10.1080/10556788.2023.2167994

6. Larson, J., Menickelly, M., & Wild, S. M. (2019). Derivative-free optimization methods. *Acta Numerica*, *28*, 287-404. https://doi.org/10.1017/S0962492919000060

7. Li, Z., Dong, Z., Liang, Z., & Ding, Z. (2021). Surrogate-based distributed optimisation for expensive black-box functions. *Automatica*, *125*, 109407. https://doi.org/10.1016/j.automatica.2020.109407

8. Martin, D., & Tremblay, C. (2026). Economic policy optimization using data-driven decision models. International Academic Journal of Economics, 13(3), 46–53.

9. Mirjalili, S., & Lewis, A. (2016). The whale optimization algorithm. *Advances in engineering software*, *95*, 51-67. https://doi.org/10.1016/j.advengsoft.2016.01.008

10. Mirjalili, S., Mirjalili, S. M., & Lewis, A. (2014). Grey wolf optimizer. *Advances in engineering software*, *69*, 46-61. https://doi.org/10.1016/j.advengsoft.2013.12.007

11. Mohankumar, M., Balamurugan, K., Singaravel, G., & Menaka, S. R. (2024). A Dynamic Workflow Scheduling Method based on MCDM Optimization that Manages Priority Tasks for Fault Tolerance. *International Academic Journal of Science and Engineering, 11*(1), 09–14. https://doi.org/10.9756/IAJSE/V11I1/IAJSE1102

12. Morlet-Espinosa, J., & Flores-Tlacuahuac, A. (2024). A Bayesian optimization approach for data-driven mixed-integer nonlinear programming problems. *AIChE Journal*, *70*(9), e18448. https://doi.org/10.1002/aic.18448

13. Müller, J., & Woodbury, J. D. (2017). GOSAC: global optimization with surrogate approximation of constraints. *Journal of Global Optimization*, *69*(1), 117-136. https://doi.org/10.1007/s10898-017-0496-y

14. Nakatsukasa, Y., & Trefethen, L. N. (2020). An algorithm for real and complex rational minimax approximation. *SIAM Journal on Scientific Computing*, *42*(5), A3157-A3179. https://doi.org/10.1137/19M1281897

15. Peiris, V., Sharon, N., Sukhorukova, N., & Ugon, J. (2021). Generalised rational approximation and its application to improve deep learning classifiers. *Applied Mathematics and Computation*, *389*, 125560. https://doi.org/10.1016/j.amc.2020.125560

16. Picheny, V., Wagner, T., & Ginsbourger, D. (2013). A benchmark of kriging-based infill criteria for noisy optimization. *Structural and multidisciplinary optimization*, *48*(3), 607-626. https://doi.org/10.1007/s00158-013-0919-4

17. Rossi, A., & Lanzetta, M. (2020). Integration of hybrid additive/subtractive manufacturing planning and scheduling by metaheuristics. *Computers & Industrial Engineering*, *144*, 106428. https://doi.org/10.1016/j.cie.2020.106428

18. Selvaraj, R., Kuthadi, V., Baskar, S., & Acevedo, R. (2025). Exploring Motivation and Learning Factors in Digital Media Through the Lens of Activity Theory and Behavioural Influences. *Indian Journal of Information Sources and Services, 15*(3), 37–47. https://doi.org/10.51983/ijiss-2025.IJISS.15.3.05

19. Tsai, J. T., Fang, J. C., & Chou, J. H. (2013). Optimized task scheduling and resource allocation on cloud computing environment using improved differential evolution algorithm. *Computers & Operations Research*, *40*(12), 3045-3055. https://doi.org/10.1016/j.cor.2013.06.012

20. Turgut, O. E., Turgut, M. S., & Kırtepe, E. (2023). A systematic review of the emerging metaheuristic algorithms on solving complex optimization problems. *Neural Computing and Applications*, *35*(19), 14275-14378. https://doi.org/10.1007/s00521-023-08481-5

21. Van De Berg, D., Petsagkourakis, P., Shah, N., & del Rio-Chanona, E. A. (2023). Data-driven coordination of subproblems in enterprise-wide optimization under organizational considerations. *AIChE Journal*, *69*(4), e17977. https://doi.org/10.1002/aic.17977

22. Yogamadhavan, V. K., & Mannayee, Dr. G. (2024). Evaluation of Performance of Different Optimizers of Convolutional Neural Network in the Classification of Images of Urban Domestic Solid Waste. *Journal of Internet Services and Information Security, 14*(3), 42–77. https://doi.org/10.58346/jisis.2024.i3.004

23. Zhang, H., Buchmeister, B., Li, X., & Ojstersek, R. (2023). An efficient metaheuristic algorithm for job shop scheduling in a dynamic environment. *Mathematics*, *11*(10), 2336. https://doi.org/10.3390/math11102336