

Learning-Based Efficient Approximation of Data-enabled Predictive Control

Yihan Zhou¹, Yiwen Lu¹, Zishuo Li¹, Jiaqi Yan², Yilin Mo¹

Abstract—Data-Enabled Predictive Control (DeePC) bypasses the need for system identification by directly leveraging raw data to formulate optimal control policies. However, the size of the optimization problem in DeePC grows linearly with respect to the data size, which prohibits its application due to high computational costs. In this paper, we propose an efficient approximation of DeePC, whose size is invariant with respect to the amount of data collected, via differentiable convex programming. Specifically, the optimization problem in DeePC is decomposed into two parts: a control objective and a scoring function that evaluates the likelihood of a guessed I/O sequence, the latter of which is approximated with a size-invariant learned optimization problem. The proposed method is validated through numerical simulations on a quadruple tank system, illustrating that the learned controller can reduce the computational time of DeePC by 5x while maintaining its control performance.

I. INTRODUCTION

Optimal trajectory tracking has long been a fundamental problem in control systems. Classical model-based control approaches, such as Model Predictive Control (MPC), typically rely on a *state-space model* of the system to predict future trajectories and determine the optimal control decision. However, defining the state space model can be a laborious part of the control design [1, 2], and even after it is defined, identifying its parameters from noisy observations is known to be difficult [3, 4].

In recent years, direct data-driven control [5–7] has emerged as a promising solution to the aforementioned challenges, with Data-Enabled Predictive Control (DeePC) [8] being a representative algorithm encompassing this concept. A key idea of direct data-driven control is to construct a *non-parametric model* of the system using a data matrix formed by Input/Output (I/O) trajectories of the system collected *a priori*. According to Willem’s fundamental lemma [9], any feasible I/O trajectory of a noise-free linear system must be spanned by these offline trajectories, as long as the persistency of excitation condition [10] is satisfied, which justifies a non-parametric representation. Based on this non-parametric model, DeePC tackles the tracking problem by solving an optimization problem in a receding-horizon manner similarly to MPC, but with the non-parametric model replacing the state-space model for prediction.

DeePC has received significant attention from both theoretical [11, 12] and empirical [13, 14] perspectives. However,

a notable challenge associated with the practical deployment of DeePC lies in its computation demand [6, 15, 16]. Since the I/O data collected from real-world systems are typically contaminated with noise, a large amount of data is required to generate an accurate non-parametric model of the system. However, as the amount of data used for building the non-parametric model increases, width of the data matrix grows accordingly, resulting in an increase in the dimension of a decision variable in DeePC. Therefore, the computational burden of DeePC scales with the amount of data collected, which can be prohibitive for real-time applications such as robotics and power electronics [17].

Previous research efforts have been devoted to mitigating the computational challenge of DeePC by compressing the data matrix. On the one hand, for lossless compression, a representation of predicted trajectories based on the null-space of the data matrix is proposed [13]. On the other hand, approximation methods have been developed to perform lossy compression of the data matrix, including those based on Singular Value Decomposition (SVD) [18], Proper Orthogonal Decomposition (POD) [19], and LQ decomposition [20, 21]. Overall, these methods amount to diminishing the number of effective trajectory segments in the nonparametric model, i.e., reducing the width of the data matrix.

In this paper, we propose an alternative view of DeePC: we show that the DeePC problem is equivalent to minimizing the sum of a control objective $\ell(\tau)$ and an anomaly metric $S(\tau)$, both as functions of the predicted I/O sequence τ . The former encompasses the tracking error, the penalty on control effort, and the constraints on I/O signals, which is independent from the data collected. The latter maps an I/O sequence to a scalar representing how unlikely τ is a trajectory generated by the system, whose evaluation is based on the data. We dub $S(\tau)$ as a *scoring function* or *scoring model*, since it can serve as an oracle whose queries return a score indicating the fitness of a guessed I/O sequence to the system dynamics. Although this reformulation is equivalent to the original DeePC, it motivates us to accelerate computation of the controller via a learning-based approximation $\hat{S}(\tau)$ of the scoring function whose computational complexity is fixed even with more data collected. Specifically, we propose to parameterize the approximate scoring function as a differentiable convex program, and train its parameters via supervised learning against the true scoring function S . With our proposed learning objective based on proximal operators, the minimizer of $\ell + \hat{S}$ is close to the minimizer of $\ell + S$, indicating that the approximate scoring function can be used to solve the overall control problem effectively.

An advantage of our approach is its computational merit in

¹Yihan Zhou, Yiwen Lu, Zishuo Li, and Yilin Mo are with the Department of Automation, Tsinghua University, Beijing, 100084, China. ({zhouyh23, luyw20, lizs19}@mails.tsinghua.edu.cn, ylmo@tsinghua.edu.cn)

²Jiaqi Yan is with the Automatic Control Laboratory, ETH Zurich, Switzerland. (jiayan@ethz.ch)

decoupling the scale of the control problem from the amount of data used. That is, once the form of the approximate scoring function is determined, one can train its parameters *offline* with a large dataset, without affecting the number of variables or constraints in the control problem to be solved *online*. This stands in contrast with the original DeePC, which, as mentioned, has a set of variables whose dimension scales with the amount of data. Hence, the learned approximate scoring function $\hat{S}$ can be viewed as a condensed representation of the system dynamics, which, combined with the control objective, enables computationally efficient control. The above intuitions are supported by numerical simulations, which show that our proposed controller can achieve a similar control performance to DeePC in less computational time.

Our contributions can be summarized as follows: i) We present a reformulation of DeePC based on the notion of scoring model, contributing to a new perspective on the data-enabled control problem. ii) We propose an approximate DeePC control law derived from a learning-based approximation of the scoring function. iii) We demonstrate through numerical simulations that the proposed method can significantly reduce the computational time required for DeePC, with only a minor compromise in control performance.

The rest of the paper is organized as follows. Section II provides a brief overview of the DeePC algorithm and its computational restriction. Section III presents our learning-based approximation of DeePC. Section IV presents the simulation results. Section V concludes the paper.

Notations: The set of all integers is denoted by $\mathbb{Z}$, with subsets of positive integers and non-negative integers represented as $\mathbb{Z}_{>0}$ and $\mathbb{Z}_{\geq 0}$, respectively. In the context of matrices and vectors, $X^\top$ denotes the transpose of matrix X , and $X^\dagger$ represents the Moore-Penrose pseudo-inverse of matrix X . The operator $\text{diag}(x_1, \dots, x_n)$ creates a diagonal matrix with elements $x_1, \dots, x_n$ on its diagonal. The indicator function $\mathbb{I}_{\mathcal{X}}(x)$ equals 0 if $x \in \mathcal{X}$ and $+\infty$ otherwise. The gradient, subdifferential, and proximal operator of a function f are denoted by ∇f , ∂f , and Prox_f , respectively. The operator col is used to denote the column stack of vectors. I denotes the identity matrix of appropriate dimension or the identity operator according to the context. The p -norm of a vector or matrix or operator is denoted by $\|\cdot\|_p$. The quadratic form $\|x\|_Q^2 = x^\top Q x$ signifies the weighted norm squared of vector x . The symbol $\circ$ denotes the composition of two operators, $\otimes$ represents the Kronecker product, and $\langle \cdot, \cdot \rangle$ denotes the inner product. The symbol $\mathcal{N}(\mu, \Sigma)$ denotes a Gaussian distribution with mean μ and covariance Σ .

II. PRELIMINARIES

A. Problem Formulation

In this paper, we consider a discrete-time Linear Time-Invariant (LTI) system with state-space representation

$$x(t+1) = Ax(t) + Bu(t) + w(t), \quad (1a)$$

$$y(t) = Cx(t) + v(t), \quad (1b)$$

where $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times m}$, $C \in \mathbb{R}^{p \times n}$ are state-space matrices, $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}^m$, $y(t) \in \mathbb{R}^p$ are state, input, and output vectors at time $t \in \mathbb{Z}_{\geq 0}$, respectively, and $w(t) \stackrel{\text{i.i.d.}}{\sim} \mathcal{N}(0, \Sigma_w)$ and $v(t) \stackrel{\text{i.i.d.}}{\sim} \mathcal{N}(0, \Sigma_v)$ are process and measurement noises. We assume that (A, B) is controllable and (A, C) is observable.

We consider an optimal tracking problem of the system (1), whose objective is to minimize a quadratic function

$$\limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} (\|y(t) - r\|_Q^2 + \|u(t)\|_R^2), \quad (2)$$

subject to the constraints

$$y(t) \in \mathcal{Y}, u(t) \in \mathcal{U}, \forall t \in \mathbb{Z}_{\geq 0}, \quad (3)$$

where r is a reference signal to be tracked, $\mathcal{Y} \subseteq \mathbb{R}^p$ and $\mathcal{U} \subseteq \mathbb{R}^m$ are convex sets representing the output and input constraints respectively, and Q, R are symmetric positive definite weighting matrices.

Remark 1. The reference signal can be a time-varying signal $r(t)$ for all of the methods in this paper, but we assume that r is a constant for notational simplicity.

We consider a setting where the state-space representation (1) is unavailable to the system operator, i.e., none of the state-space matrices A, B, C , the state vectors $x(t)$, or the state dimension n is known. Instead, only the input and output vectors $u(t)$ and $y(t)$ are observed.

B. Data-Enabled Predictive Control (DeePC)

DeePC [22] tackles the aforementioned optimal tracking problem in a receding horizon manner, where the optimization problem described below is solved at every time step t . We use the convention that the index in the parentheses, e.g., $y(t)$, denotes the actual value of the variable at time t , while the index in subscript, e.g., y_k , denotes the predicted value of the variable at the k -th step in the optimization horizon.

Problem 1 (DeePC Problem at Time Step t).

$$\begin{aligned} \min_{u_{0:N-1}, y_{0:N-1}, g, \sigma_y} \quad & \sum_{k=0}^{N-1} \left(\|y_k - r\|_Q^2 + \|u_k\|_R^2 \right) + \\ & \lambda_{g1} \|g\|_1 + \lambda_{g2} \|g\|_2^2 + \lambda_{y1} \|\sigma_y\|_1 + \lambda_{y2} \|\sigma_y\|_2^2 \\ \text{s.t.} \quad & \underbrace{\begin{matrix} mT_{\text{ini}}\{ \\ mN\{ \\ pT_{\text{ini}}\{ \\ pN\{ \end{matrix} \begin{bmatrix} U_p \\ U_f \\ Y_p \\ Y_f \end{bmatrix}}_M g = \begin{bmatrix} u_{\text{ini}} \\ u_{0:N-1} \\ y_{\text{ini}} \\ y_{0:N-1} \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \sigma_y \\ 0 \end{bmatrix}, \\ & u_k \in \mathcal{U}, \forall k \in \{0, \dots, N-1\}, \\ & y_k \in \mathcal{Y}, \forall k \in \{0, \dots, N-1\}, \end{aligned} \quad (4)$$

where the new notations are defined and explained as follows:

- $N \in \mathbb{Z}_{>0}$ is the prediction horizon, $M \in \mathbb{Z}_{>0}$ is the number of trajectory segments collected offline, and

$T_{\text{ini}} \geq l$ is the length of the initialization sequence, where l is the lag [23, Section 7.2] of the system.

- The matrix $\mathcal{H} \triangleq [U_p^\top \ U_f^\top \ Y_p^\top \ Y_f^\top]^\top$ is the data matrix, each column of which is formed by concatenating the Input/Output (I/O) signals collected offline at $T_{\text{ini}} + N$ consecutive time steps. We call such a column a *trajectory segment*. The data matrix has M columns, and is partitioned into row block matrices U_p, U_f, Y_p, Y_f , with number of rows $mT_{\text{ini}}, mN, pT_{\text{ini}}, pN$ respectively, where the subscripts p and f stand for past and future, respectively. It is assumed that the matrix $[U_p^\top \ U_f^\top]^\top$ is of full row rank.
- $u_{\text{ini}} \in \mathbb{R}^{mT_{\text{ini}}}$ and $y_{\text{ini}} \in \mathbb{R}^{pT_{\text{ini}}}$ are column vectors formed by concatenating the I/O signals of the past T_{ini} time steps, which are inputs to the DeePC problem.
- $u_{0:N-1} \in \mathbb{R}^{mN}$, $y_{0:N-1} \in \mathbb{R}^{pN}$ stand for the predicted I/O sequence over the horizon N , which are optimization variables of the DeePC problem. After solving the DeePC problem, u_0^* from the optimal solution is applied to the system, i.e., $u(t) = u_0^*$.
- $g \in \mathbb{R}^M$ and $\sigma_y \in \mathbb{R}^{pN}$ are auxiliary optimization variables of the DeePC problem, which represent the coefficients for spanning the predicted I/O trajectory sequence from the data matrix and the slack variable for the output constraints, respectively. The scalar parameters $\lambda_{g1}, \lambda_{g2}, \lambda_{y1}, \lambda_{y2} \in \mathbb{R}_{\geq 0}$ are predefined coefficients. When brought together, $\lambda_{g1}\|g\|_1 + \lambda_{g2}\|g\|_2^2 + \lambda_{y1}\|\sigma_y\|_1 + \lambda_{y2}\|\sigma_y\|_2^2$ is viewed as a regularizer essential for the robustness of the controller in the presence of noises [6, 24]. In different versions of DeePC, the regularizer form has been chosen to be l_1 -norm [8, 17], squared l_2 -norm [18, 19], or the hybrid of both [6]. Therefore, we adopt a general formulation that can cover the above cases.

Since the collected I/O data are contaminated with noise, a large amount of data is required for an accurate non-parametric representation of the system dynamics. However, the dimension M of the auxiliary optimization variable g grows linearly with respect to the amount of data, which poses a computational challenge to the online optimization of DeePC.

III. COMPUTATIONALLY EFFICIENT DEEPC VIA LEARNING-BASED APPROXIMATION

We propose a computationally efficient approximate DeePC by introducing the notion of the scoring function. We first show that the DeePC objective can be viewed as minimizing the sum of the control cost and the score, and the evaluation of the latter is costly due to the scale of data matrix. Therefore, we propose to approximate the scoring function with a reduced-order one via differentiable convex programming. The parameters of the reduced scoring function are learned offline. Finally, an approximate control law is formulated as minimizing the sum of the control cost and the learned approximate scoring function. The overall framework is illustrated in Fig 1.

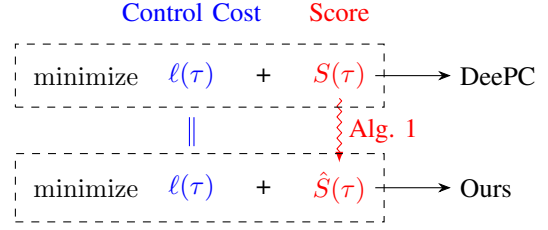


Fig. 1. Illustration of the overall framework. The optimization objective of DeePC can be split into two parts: the control cost $\ell(\tau)$ and the score function $S(\tau)$. Subsequently, $S(\tau)$ is approximated by $\hat{S}(\tau)$, which is learned offline. The learning-based efficient approximation of DeePC is formulated as minimizing the sum of the control cost $\ell(\tau)$ and the learned approximate score $\hat{S}(\tau)$.

A. DeePC through the Lens of Scoring Function

Let $L = T_{\text{ini}} + N$ be the length of the I/O trajectory sequence in the data matrix. we denote

$$\tau \triangleq \text{col}(\mathbf{u}, \mathbf{y})$$

as the predicted I/O trajectory sequence, where $\mathbf{u} \in \mathbb{R}^{mL}$ and $\mathbf{y} \in \mathbb{R}^{pL}$ are the predicted input and output sequences, respectively.

We propose to reformulate Problem 1 and separate the objective into two parts:

$$\min_{\tau} \ell(\tau) + S(\tau), \quad (6)$$

where the objective functions are defined as

$$\ell(\tau) = \|\mathbf{y} - \mathbf{r}\|_{\mathbf{Q}}^2 + \|\mathbf{u}\|_{\mathbf{R}}^2 + \mathbb{I}_{\mathcal{U}^L \times \mathcal{Y}^L}(\tau) + \mathbb{I}_{\{\tau_{\text{ini}}\}}(E_{\text{ini}}\tau), \quad (7)$$

$$S(\tau) = \min_{g, \sigma_y} \lambda_{g1}\|g\|_1 + \lambda_{g2}\|g\|_2^2 + \lambda_{y1}\|\sigma_y\|_1 + \lambda_{y2}\|\sigma_y\|_2^2, \\ \text{s.t. } \tilde{\mathcal{H}} \begin{bmatrix} g \\ \sigma_y \end{bmatrix} = \tau, \quad (8)$$

and $\mathbf{r} \triangleq r \otimes \mathbf{1}_L$, $\mathbf{Q} \triangleq Q \otimes I_L$, $\mathbf{R} \triangleq R \otimes I_L$, $\tau_{\text{ini}} \triangleq \text{col}(u_{\text{ini}}, y_{\text{ini}})$, $E_{\text{ini}} \triangleq \begin{bmatrix} I & 0 & 0 & 0 \\ 0 & 0 & I & 0 \end{bmatrix}$ is a selection matrix that extracts the initial part of I/O sequence from τ , $\tilde{\mathcal{H}} \triangleq \begin{bmatrix} U_p^\top & Y_p^\top & U_f^\top & Y_f^\top \\ 0 & -I & 0 & 0 \end{bmatrix}^\top$. In the notation of $\mathbb{I}_{\mathcal{U}^L \times \mathcal{Y}^L}(\tau)$, “ $\times$ ” denotes Cartesian power and $\mathcal{U}^L$ denotes L -ary Cartesian power of a set $\mathcal{U}$.

Here, $\ell(\cdot)$ and $S(\cdot)$ fulfill two orthogonal roles:

- $\ell(\cdot)$ represents the cost function of the control problem, which is distinct from and unaffected by the system dynamics.
- $S(\cdot)$ is referred to as the scoring function. For any I/O trajectory τ , it evaluates the fitness of τ to the trajectory segments collected in the data matrix. It depends only on the system and is not influenced by any specific costs or constraints.

Note that g, σ_y are auxiliary variables required for the evaluation of the scoring function, not part of the control policy, and the dimension of g grows with the number of trajectory segments in the data matrix. Therefore, we desire to accelerate the evaluation of $S(\tau)$ using a learning-based approach which approximates the scoring function

with fewer auxiliary variables. We denote the approximate scoring function as $\hat{S}(\tau)$, then the overall control problem becomes:

$$\min_{\tau} \ell(\tau) + \hat{S}(\tau). \quad (9)$$

B. Learning Objective

Since the approximation goal is to make control input, i.e., the optimal solution, of the learned problem $\min \ell(\tau) + \hat{S}(\tau)$ close to that of $\min \ell(\tau) + S(\tau)$, we set the learning target to be minimizing the error between the *proximal operators* of the approximate and true scoring functions $\hat{S}$ and S .

The proximal operator [25] of $\hat{S}$ is defined as:

$$\text{Prox}_{\hat{S}}(\tau) = \arg \min_{\hat{\tau}} \hat{S}(\hat{\tau}) + \frac{1}{2} \|\hat{\tau} - \tau\|^2. \quad (10)$$

Consequently, the optimality condition of (9), i.e., $0 \in (\partial \ell + \partial \hat{S})(\hat{\tau}^*)$, can be equivalently expressed with the proximal operators of ℓ and $\hat{S}$ using operator splitting methods. Therefore, the optimal solution $\hat{\tau}^*$ to (9) can be determined once $\text{Prox}_{\hat{S}}$ is given.

C. Approximator Form: Differentiable Convex Program

Given the learning target Prox_S , the subsequent step is to select a parameterized family of $\hat{S}$ that satisfies the following conditions: i) $\hat{S}(\tau)$ is capable of approximately representing the true scoring function $S(\tau)$; ii) $\hat{S}$ is a convex function of τ , and $\text{Prox}_{\hat{S}}$ can be easily parameterized; iii) the control problem $\min_{\tau} \ell(\tau) + \hat{S}(\tau)$ should be computationally tractable. Considering the above requirements, we adopt a differentiable convex programming approach to represent $\hat{S}$. Specially, we adopt the following form:

$$\hat{S}(\tau) = \min_{z \in \mathbb{R}^{n_z}} \|\text{diag}(d_1)z\|_1 + \|\text{diag}(d_2)z\|_2^2 \quad (11a)$$

$$\text{s.t. } Gz + W\tau = 0. \quad (11b)$$

Here, $\hat{S}$ is a mapping from an I/O sequence τ to a scalar standing for the optimal value of a convex optimization problem with n_z variables and m_z constraints, whose decision variable is z and whose parameters include τ as well as learnable parameters $d_1 \in \mathbb{R}^{n_z}, d_2 \in \mathbb{R}^{n_z}, G \in \mathbb{R}^{m_z \times n_z}, W \in \mathbb{R}^{m_z \times L(m+p)}$. To train the differentiable convex program means to update the parameters d_1, d_2, G, W via gradient-based methods, such that the learning target ($\text{Prox}_{\hat{S}} \approx \text{Prox}_S$ in our case) is achieved. The size of the convex program, i.e., n_z and m_z , can be arbitrarily chosen according to a trade-off between the approximation accuracy and the computational cost, but once fixed, it is invariant against the amount of data used for training.

We first show that the chosen form of $\hat{S}$ has adequate representational power to approximate the true scoring function S :

Theorem 1. *With sufficiently large n_z and m_z , the true $S(\tau)$ is representable by (11).*

Proof. Let $d_1 = [\underbrace{\lambda_{g1}, \dots, \lambda_{g1}}_{M \text{ times}}, \underbrace{\lambda_{\sigma_{y1}}, \dots, \lambda_{\sigma_{y1}}}_{pT_{\text{ini}} \text{ times}}]^\top$, $d_2 =$

$[\underbrace{\sqrt{\lambda_{g2}}, \dots, \sqrt{\lambda_{g2}}}_{M \text{ times}}, \underbrace{\sqrt{\lambda_{\sigma_{y2}}}, \dots, \sqrt{\lambda_{\sigma_{y2}}}}_{pT_{\text{ini}} \text{ times}}]^\top$, $G = \tilde{H}$, $W = -I$, then $\hat{S}(\tau)$ defined in (11) is equivalent to $S(\tau)$. $\square$

We next show how to evaluate and learn the proximal operator of the function $\hat{S}$ given in (11). By definition (10) of proximal operator, we have:

$$(z^*, \hat{\tau}^*) = \arg \min_{z, \hat{\tau}} \|\text{diag}(d_1)z\|_1 + \|\text{diag}(d_2)z\|_2^2 + \frac{1}{2} \|\hat{\tau} - \tau\|^2, \quad \text{s.t. } Gz + W\hat{\tau} = 0. \quad (12)$$

To evaluate (12) and find the gradient of τ^* with respect to the learnable parameters d_1, d_2, G, W , we adopt an algorithm unrolling [26, 27] approach, which has been proved efficient for controller learning [28]. In particular, we unroll the Douglas-Rachford Splitting (DRS) [29], which, when applied to the problem (12), yields the following iterations:

$$\begin{bmatrix} z^{k+\frac{1}{2}} \\ \hat{\tau}^{k+\frac{1}{2}} \end{bmatrix} = \begin{bmatrix} \text{sh}_{d_1, d_2}(\xi^k) \\ \frac{\tau + \eta^k}{2} \end{bmatrix}, \quad (13a)$$

$$\begin{bmatrix} z^{k+1} \\ \hat{\tau}^{k+1} \end{bmatrix} = \left(I - \tilde{G}^\dagger \tilde{G} \right) \begin{bmatrix} 2z^{k+\frac{1}{2}} - \xi^k \\ 2\hat{\tau}^{k+\frac{1}{2}} - \eta^k \end{bmatrix}, \quad (13b)$$

$$\begin{bmatrix} \xi^{k+1} \\ \eta^{k+1} \end{bmatrix} = \begin{bmatrix} \xi^k + z^{k+1} - z^{k+\frac{1}{2}} \\ \eta^k + \hat{\tau}^{k+1} - \hat{\tau}^{k+\frac{1}{2}} \end{bmatrix}, \quad (13c)$$

where ξ, η are auxiliary variables, $\tilde{G} = [G \ W]$, and the soft-thresholding operator $\text{sh}_{d_1, d_2}(x)$ is defined as:

$$\text{sh}_{d_1, d_2}(x)_i = \begin{cases} \frac{x_i - |d_{1i}|}{1 + 2(d_{2i})^2} & \text{if } x_i - |d_{1i}| > 0, \\ \frac{x_i + |d_{1i}|}{1 + 2(d_{2i})^2} & \text{if } x_i + |d_{1i}| < 0, \\ 0 & \text{otherwise.} \end{cases} \quad (14)$$

Here, d_{1i} and d_{2i} are the i -th elements of d_1 and d_2 , respectively.

The unrolled iterations resemble a deep neural network that interleaves affine transformations with soft-thresholding nonlinearities, whose weights are functions of the learnable parameters d_1, d_2, G, W . This resemblance, visualized in Fig 2, facilitates the training of the learnable parameters using established deep learning frameworks.

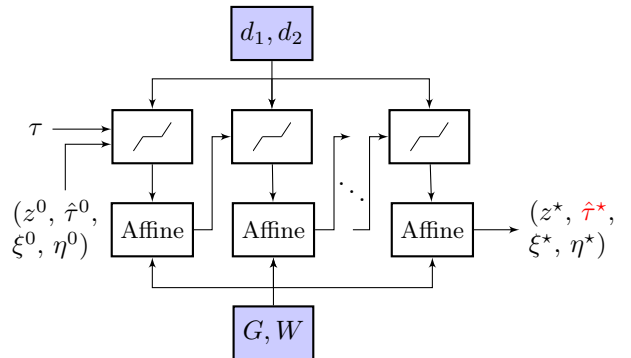


Fig. 2. Illustration of evaluating $\text{Prox}_{\hat{S}}$ with unrolled DRS iterations. The input is an I/O sequence τ and the initial values of z, τ, ξ, η , and the output is $\text{Prox}_{\hat{S}}(\hat{\tau})$, denoted as $\hat{\tau}^*$. The learnable parameters are d_1, d_2, G, W . sh is the soft-thresholding operator (14), and the Affine block is the composition of all affine transformations in each iteration (13).

D. Learning Algorithm

Given a data matrix $\mathcal{H}$ consisting of I/O trajectory sequences, we first collect a dataset $\mathcal{D}$ of I/O trajectory sequences by linear combination of the trajectories in $\mathcal{H}$ and adding random noise to the trajectories to increase the diversity of the dataset. Then, we compute the ground truth $\text{Prox}_S(\tau)$ for each trajectory τ in $\mathcal{D}$. Finally, we learn the parameters $\theta = (d_1, d_2, G, W)$ of the approximate scoring function $\hat{S}$ by minimizing the mean squared error between $\text{Prox}_{\hat{S}}$ and the ground truth Prox_S across all entries in the dataset $\mathcal{D}$.

The learning procedure of the scoring function is summarized in Algorithm 1.

Algorithm 1: Framework for Learning the Scoring Function

Input: Data matrix $\mathcal{H}$ consisting of I/O trajectory sequences, regularization parameters $\lambda_{g1}, \lambda_{g2}, \lambda_{y1}, \lambda_{y2}$

Output: Approximate scoring function parameters $\theta = (d_1, d_2, G, W)$

- 1 Collect a dataset $\mathcal{D}$ of I/O trajectory sequences by applying different noises to the trajectories in $\mathcal{H}$.
- 2 Compute $\text{Prox}_S(\tau)$ for each trajectory τ in $\mathcal{D}$.
- 3 Initialize the parameters $\theta = (d_1, d_2, G, W)$.
- 4 **for** $epoch = 1, 2, \dots$ **do**
- 5 Solve Problem 12 to obtain $\text{Prox}_{\hat{S}}(\tau)$ for each trajectory τ in $\mathcal{D}$.
- 6 Compute the mean squared error denoted by $\ell_{mse}(\theta; \mathcal{D}) = \sum_{\tau \in \mathcal{D}} \|\text{Prox}_S(\tau) - \text{Prox}_{\hat{S}}(\tau)\|_2^2$.
- 7 Update θ according to the loss $\ell_{mse}(\theta; \mathcal{D})$.
- 8 **end**

Note that learning the approximate scoring function is an offline process, and the learned parameters θ are used to substitute the original scoring function $S(\tau)$ to solve the overall control problem (9) online.

IV. SIMULATION

In this section, we present the simulation results to demonstrate the effectiveness of our approach. All computations are performed on a computer with an AMD Ryzen 9 5900X Processor clocked at 3.7GHz. All controllers are implemented using the MOSEK optimizer. Our code is available at <https://github.com/zhou-yh19/redpc>.

Consider a quadruple-tank system [30] whose simulation employs the same linearized system equations and system parameters as the LTI system used in [18]:

$$A = \begin{bmatrix} 0.921 & 0 & 0.041 & 0 \\ 0 & 0.918 & 0 & 0.033 \\ 0 & 0 & 0.924 & 0 \\ 0 & 0 & 0 & 0.937 \end{bmatrix},$$

$$B = \begin{bmatrix} 0.017 & 0.001 \\ 0.001 & 0.023 \\ 0 & 0.061 \\ 0.072 & 0 \end{bmatrix}, C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}.$$

The system is subject to process noise $w(t) \sim \mathcal{N}(0, 0.01I_4)$ and measurement noise $v(t) \sim \mathcal{N}(0, 0.1I_2)$.

We follow the same setting as in [18], such as the control cost matrices $Q = 35 \cdot I_2$ and $R = 10^{-4}I_2$, the control input and output constraints $\mathcal{U} = \mathcal{Y} = [-2, 2]$, the setpoint $r(t) = [0.65, 0.77]^T$, the prediction horizon $N = 20$, and the initial sequence length $T_{\text{ini}} = 10$.

Data Collection: The dataset is collected by applying random input sequences to the system. The input sequences are generated by sampling from a uniform distribution over $\mathcal{U}$. The corresponding output sequences are obtained by simulating the system with the input sequences. In total, 1500 I/O data points are collected to construct the data matrix $\mathcal{H}$ in the form of a Hankel matrix.

We choose the regularization parameters of DeePC as $\lambda_{g1} = 1$, $\lambda_{g2} = 100$, $\lambda_{y1} = 100$, $\lambda_{y2} = 100000$, and incorporates all 1500 I/O data points into the data matrix. And the approximate scoring function is learned with 55 equations, 110 variables and 200 DRS iterations. The training procedure is completed within 260s on a single NVIDIA GeForce RTX 4090 GPU. Fig 3 illustrates the first two states of the system under the control of both the DeePC controller and our proposed approach. The results demonstrate that both controllers track the setpoint effectively.

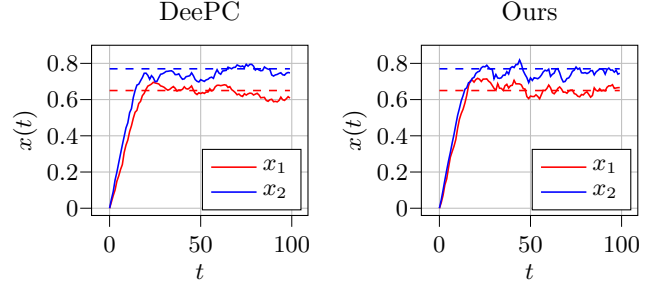


Fig. 3. Tracking performance of DeePC and our approach.

In Table I, we benchmark the control performance of our method with DeePC and MPC, detailing the sample average accumulative cost $\frac{1}{K} \sum_k \sum_t (\|y(t) - r(t)\|_Q^2 + \|u(t)\|_R^2)$, the average and worst computation time per step. Our approach achieves control performance comparable to MPC and DeePC, while significantly enhancing computational efficiency.

TABLE I
PERFORMANCE COMPARISON OF DEEPC, MPC, AND OURS

Method \ Metrics	MPC	DeePC	Ours
Average Accumulative Cost	305.00	290.68	296.25
Average Computation Time (ms)	96.35	153.71	30.41
Worst Computation Time (ms)	136.45	280.44	80.7

Furthermore, we vary the number of I/O sequences in the data matrix of DeePC and the size of the learned approximate scoring function in our method to examine their effects on control performance. The results are shown in Fig. 4. For DeePC, an increase in the number of I/O sequences in the

data matrix results in better control performance but also longer computation time. For our approach, the computation time is mainly determined by the size of the learned approximate scoring function, and there is also trade-off between control performance and computation time. Nevertheless, it can be observed that under the same computational time constraints, our method achieves superior control performance compared to DeePC, and when aiming for comparable control quality, our approach significantly reduces the computation time. This demonstrates the effectiveness of our approach in enhancing the computational efficiency of DeePC while maintaining control performance.

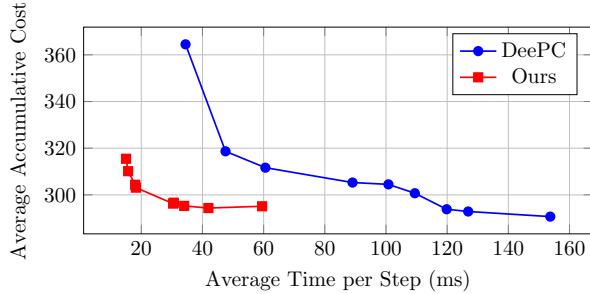


Fig. 4. Comparison of the cost-time curves between DeePC and our approach. The average computation time for DeePC is influenced by the number of I/O sequences in the data matrix, whereas our approach's computation time primarily depends on the size of the learned approximate scoring function.

V. CONCLUSIONS

In this paper, we propose a computationally efficient approximate DeePC by introducing the notion of the scoring function and replacing it with a learned approximate one via differentiable convex programming. The parameters of the reduced scoring function are learned offline, and the control law is formulated as minimizing the sum of the control cost and the learned approximate scoring function. Simulation results demonstrate our approach can achieve a comparable tracking performance to DeePC while significantly reducing the computation time. In the future, we will apply our approach to the real-world control problems and investigate the generalization ability of the learned approximate scoring function.

REFERENCES

- [1] H. Hjalmarsson, "From experiment design to closed-loop control," *Automatica*, vol. 41, no. 3, pp. 393–438, 2005.
- [2] B. A. Ogunnaike, "A contemporary industrial perspective on process control theory and practice," *Annual Reviews in Control*, vol. 20, pp. 1–8, 1996.
- [3] A. Tsiamis and G. J. Pappas, "Linear systems can be hard to learn," in *2021 60th IEEE Conference on Decision and Control (CDC)*. IEEE, 2021, pp. 2903–2910.
- [4] J. Li, S. Sun, and Y. Mo, "Fundamental limit on siso system identification," in *2022 IEEE 61st Conference on Decision and Control (CDC)*. IEEE, 2022, pp. 856–861.
- [5] C. De Persis and P. Tesi, "Formulas for data-driven control: Stabilization, optimality, and robustness," *IEEE Transactions on Automatic Control*, vol. 65, no. 3, pp. 909–924, 2019.
- [6] F. Dörfler, J. Coulson, and I. Markovsky, "Bridging direct and indirect data-driven control formulations via regularizations and relaxations," *IEEE Transactions on Automatic Control*, vol. 68, no. 2, pp. 883–897, 2022.

- [7] F. Zhao, F. Dörfler, A. Chiuso, and K. You, "Data-enabled policy optimization for direct adaptive learning of the lqr," *arXiv preprint arXiv:2401.14871*, 2024.
- [8] J. Coulson, J. Lygeros, and F. Dörfler, "Data-enabled predictive control: In the shallows of the deepc," in *2019 18th European Control Conference (ECC)*. IEEE, 2019, pp. 307–312.
- [9] J. C. Willems and J. W. Polderman, *Introduction to mathematical systems theory: a behavioral approach*. Springer Science & Business Media, 1997, vol. 26.
- [10] J. C. Willems, P. Rapisarda, I. Markovsky, and B. L. De Moor, "A note on persistency of excitation," *Systems & Control Letters*, vol. 54, no. 4, pp. 325–329, 2005.
- [11] J. Berberich, J. Köhler, M. A. Müller, and F. Allgöwer, "Data-driven model predictive control with stability and robustness guarantees," *IEEE Transactions on Automatic Control*, vol. 66, no. 4, pp. 1702–1717, 2020.
- [12] S. Baros, C.-Y. Chang, G. E. Colon-Reyes, and A. Bernstein, "Online data-enabled predictive control," *Automatica*, vol. 138, p. 109926, 2022.
- [13] P. G. Carlet, A. Favato, S. Bolognani, and F. Dörfler, "Data-driven continuous-set predictive current control for synchronous motor drives," *IEEE Transactions on Power Electronics*, vol. 37, no. 6, pp. 6637–6646, 2022.
- [14] E. Elokda, J. Coulson, P. N. Beuchat, J. Lygeros, and F. Dörfler, "Data-enabled predictive control for quadcopters," *International Journal of Robust and Nonlinear Control*, vol. 31, no. 18, pp. 8916–8936, 2021.
- [15] F. Fiedler and S. Lucia, "On the relationship between data-enabled predictive control and subspace predictive control," in *2021 European Control Conference (ECC)*. IEEE, 2021, pp. 222–229.
- [16] L. Huang, J. Coulson, J. Lygeros, and F. Dörfler, "Data-enabled predictive control for grid-connected power converters," in *2019 IEEE 58th Conference on Decision and Control (CDC)*. IEEE, 2019, pp. 8130–8135.
- [17] F. Wegner, "Data-enabled predictive control of robotic systems," Master's thesis, ETH Zurich, 2021.
- [18] K. Zhang, Y. Zheng, C. Shang, and Z. Li, "Dimension reduction for efficient data-enabled predictive control," *IEEE Control Systems Letters*, 2023.
- [19] P. G. Carlet, A. Favato, R. Torchio, F. Toso, S. Bolognani, and F. Dörfler, "Real-time feasibility of data-driven predictive control for synchronous motor drives," *IEEE Transactions on Power Electronics*, vol. 38, no. 2, pp. 1672–1682, 2022.
- [20] V. Breschi, A. Chiuso, and S. Formentin, "Data-driven predictive control in a stochastic setting: A unified framework," *Automatica*, vol. 152, p. 110961, 2023.
- [21] M. Sader, Y. Wang, D. Huang, C. Shang, and B. Huang, "Causality-informed data-driven predictive control," *arXiv preprint arXiv:2311.09545*, 2023.
- [22] J. Coulson, J. Lygeros, and F. Dörfler, "Distributionally robust chance constrained data-enabled predictive control," *IEEE Transactions on Automatic Control*, vol. 67, no. 7, pp. 3289–3304, 2021.
- [23] I. Markovsky, J. C. Willems, S. Van Huffel, and B. De Moor, *Exact and approximate modeling of linear systems: A behavioral approach*. SIAM, 2006.
- [24] I. Markovsky, L. Huang, and F. Dörfler, "Data-driven control based on the behavioral approach: From theory to applications in power systems," *IEEE Control Systems Magazine*, vol. 43, no. 5, pp. 28–68, 2023.
- [25] N. Parikh, S. Boyd *et al.*, "Proximal algorithms," *Foundations and trends® in Optimization*, vol. 1, no. 3, pp. 127–239, 2014.
- [26] K. Gregor and Y. LeCun, "Learning fast approximations of sparse coding," in *Proceedings of the 27th international conference on machine learning*, 2010, pp. 399–406.
- [27] V. Monga, Y. Li, and Y. C. Eldar, "Algorithm unrolling: Interpretable, efficient deep learning for signal and image processing," *IEEE Signal Processing Magazine*, vol. 38, no. 2, pp. 18–44, 2021.
- [28] Y. Lu, Z. Li, Y. Zhou, N. Li, and Y. Mo, "Mpc-inspired reinforcement learning for verifiable model-free control," *arXiv preprint arXiv:2312.05332*, 2023.
- [29] J. Eckstein and D. P. Bertsekas, "On the douglas–rachford splitting method and the proximal point algorithm for maximal monotone operators," *Mathematical programming*, vol. 55, pp. 293–318, 1992.
- [30] K. H. Johansson, "The quadruple-tank process: A multivariable laboratory process with an adjustable zero," *IEEE Transactions on control systems technology*, vol. 8, no. 3, pp. 456–465, 2000.