Highlights

Certifiable approximation of model predictive control laws: a shape-constrained polynomial read-out

Adilkhan Salkimbayev

- Kolmogorov-Arnold network (KAN) error and symbolic-extraction error are separated and reported independently.
- Off-the-shelf symbolic read-out loses most of the KAN's fidelity; a convex refit recovers it.
- A gated parameterization makes zero offset and local stability structural.
- Constraint satisfaction and robustness of the explicit policy are measured.
- Matched-budget comparison against direct polynomial regression, multi-layer perceptron (MLP), DeepONet and linear quadratic regulator (LQR).

Certifiable approximation of model predictive control laws: a shape-constrained polynomial read-out

Adilkhan Salkimbayev[a,*]

[a]School of Information Technology and Engineering, Kazakh-British Technical University, Tole bi 59, Almaty, 050005, Almaty oblysy, Kazakhstan

ARTICLE INFO

Keywords:
approximate model predictive control
explicit control laws
shape-constrained regression
closed-loop certification
Kolmogorov–Arnold networks
non-minimum-phase systems

ABSTRACT

Approximating a model predictive control (MPC) policy offline removes the online optimizer from the loop, but it also removes the guarantees that motivated using MPC in the first place. What can be recovered analytically depends less on how well the approximant fits than on the class it belongs to: a control law that is polynomial in the measured state and linear in its own coefficients admits an analytic closed-loop Jacobian, convex shape constraints and an identifiability analysis, none of which is available for a dense network of comparable accuracy. Symbolic Kolmogorov–Arnold networks (KANs) are one route into that class, and this paper asks what this route presents, using the Johansson quadruple-tank benchmark in both its minimum-phase (MP) and non-minimum-phase (NMP) configurations. Four findings are reported. First, the symbolic read-out is a dominant error source: an off-the-shelf extraction turns a 2.4% imitation error into 7.8% of the actuator range in the MP regime (4.4% into 11.0% in the NMP one), and a convex re-estimation of the coefficients on the KAN-selected monomial support recovers most of the loss. Reporting the network and symbolization errors separately, which prior work does not do, is necessary for any claim about symbolic controllers. Second, an unconstrained symbolic read-out violates the elementary negative-feedback condition $\partial u_j/\partial e_j \geq 0$ (where u_j and e_j are the control law's term and its respective error term) on up to 30% of the operating box; an isolated occurrence of this looks like a sign error inviting manual repair, and we trace it to a rank-deficient distillation design on which 18% of random seeds produce the wrong sign. Imposing the condition as an affine inequality inside a convex least-squares read-out eliminates it for at most 0.2 percentage points of accuracy with no human in the loop. Third, embedding the learned term behind a gate that vanishes quadratically at the setpoint makes zero steady-state offset and local exponential stability structural: substituting random coefficients leaves the closed-loop spectral radius unchanged to 10^{-10}. Fourth, on a matched arithmetic budget the KAN-selected support is statistically indistinguishable from choosing the same number of monomials by orthogonal matching pursuit, and a dense multilayer perceptron imitates the teacher at least as closely as any symbolic form; what the symbolic step provides is constant-time execution and auditability, not accuracy and (on this problem) not structure selection either. The deployed law is a fixed chain of 30–203 multiply–accumulate operations with no data-dependent control flow, and remains stable under ±20% parameter perturbation in 97–100% of Monte-Carlo trials. On this benchmark its closed-loop performance is within 16% of the MPC and comparable to a well-tuned gain-scheduled linear quadratic regulator (LQR), so we also report the sharper conclusion this implies: distillation is worth its complexity only where the optimizer's advantage over a linear controller is itself large. Unlike the MPC, the distilled law carries no state-constraint guarantee, which we quantify rather than assert.

1. Introduction

Model predictive control (MPC) owes its position to a single mechanism: it solves a constrained optimal control problem at every sampling instant, and so handles multivariable coupling, actuator limits and state constraints within one framework (Borrelli et al., 2017; Mayne et al., 2000). That mechanism is also its principal implementation cost. The online quadratic program (QP) must be solved before the next sample, and its cost is not fixed: the number of iterations an operator-splitting or active-set solver takes depends on which constraints are active, and therefore on the state. Two classical routes remove that dependence by moving the optimization offline, and the choice between them is the subject of a large literature.

*Corresponding author
a_salkimbayev@kbtu.kz (A. Salkimbayev)
ORCID(s): 0009-0003-9763-0294 (A. Salkimbayev)

1.1. Approximate explicit MPC, and what has to be certified

The first route is exact: explicit MPC (Bemporad et al., 2002) solves the multi-parametric QP offline and stores the resulting piecewise-affine optimal policy; Alvarado et al. (2011) applied it to the quadruple-tank benchmark used here. Its scaling is well known — the number of polyhedral regions grows quickly with horizon and state dimension — but for the present problem a structural limitation is decisive. Exact explicit MPC presupposes a fixed parametric QP. In the linear time-varying (LTV) formulation used here the plant is re-linearized at every sampling instant, so the QP data (A_k, B_k, d_k) are themselves functions of the state and the parametric solution would have to be recomputed online. Explicit MPC is therefore discussed here as related work rather than benchmarked as a drop-in alternative; a gain-scheduled LQR, which is a legitimate constant-cost controller for this plant, is evaluated in its place.

The second route is approximate: fit a function approximator to the optimal policy offline and evaluate it online. The

idea goes back at least to Parisini and Zoppoli (1995), and the modern literature is largely concerned not with whether the fit is possible but with what can be guaranteed about the closed loop it produces. Chen et al. (2018) approximate explicit MPC with constrained neural networks; Karg and Lucia (2020) characterize the representational efficiency of deep ReLU networks for MPC laws; Lucia et al. (2021) deploy such a controller on power-electronics hardware; and Hertneck et al. (2018) obtain closed-loop guarantees by pairing a robust MPC design with statistical validation of the learned approximant. Operator-learning architectures such as DeepONet (Lu et al., 2021) offer a further parameterization in which the reference enters through a separate branch network.

Read as a control problem rather than a regression problem, approximate MPC raises three questions that a fit statistic does not answer. Does the approximate closed loop retain a stable equilibrium, and at the commanded setpoint? Does it retain the constraint satisfaction that motivated the use of MPC? And when the approximant behaves non-physically, is that detectable before deployment? Hertneck et al. (2018) answer the first statistically, at the cost of a robust MPC design and a validation budget. The alternative, pursued here, is to make the answers structural — to choose a parameterization for which the relevant properties can be established analytically, and to pay for that choice in approximation accuracy rather than in validation effort.

1.2. The read-out class decides what can be certified

What an approximate controller admits by way of analysis depends less on how accurately it fits than on the class it belongs to. All three properties provided below are available for a control law that is a polynomial in the measured state and linear in its own coefficients, and unavailable for a dense composition of fixed activations:

1. the closed-loop map has an analytic Jacobian, so local exponential stability at a setpoint can be certified by an eigenvalue computation rather than argued from simulation;
2. physical shape requirements — that a positive tracking error must not reduce the corresponding actuator command — are affine inequalities in the coefficients, so they can be imposed inside a convex program instead of being checked afterwards;
3. the coefficients are identifiable quantities, so a non-physical read-out can be diagnosed as a property of the experiment design rather than treated as an anomaly.

This is a narrower and more useful notion than interpretability in the sense usually invoked against black-box controllers (Rudin, 2019; Amodei et al., 2016).

It also determines the role that Kolmogorov–Arnold networks (KANs) (Liu et al., 2024) play in this paper, which is instrumental. KANs place learnable univariate functions on the edges of the network rather than fixed activations at the nodes. Each edge can therefore be replaced independently by a symbolic primitive and the composition expanded in closed form, which yields a sparse polynomial without committing to a dictionary in advance, as sparse regression must. Whether that route produces a better polynomial than fitting one directly is an empirical question, and one that the literature has not asked; we ask it, and report the answer.

1.3. Gaps in existing accounts

Three things are missing from published accounts of symbolic distillation for control, and they organise the paper.

1. **The two error sources are conflated.** A symbolic controller distilled from a network carries two distinct approximation errors: the network's failure to reproduce the MPC policy, and the symbolic read-out's failure to reproduce the network. They have different causes and different remedies, and the second can be the larger, because the symbolic step happens after training with nothing in the training objective encouraging the learned activations to be symbolically representable. Recent work attacks this at source by folding symbolic structure into the objective (Bagrow and Bongard, 2025; Faroughi et al., 2026; Howard et al., 2026); we instead measure the gap and close it with a convex post-processing step.
2. **Non-physical read-outs are treated as anomalies rather than as a failure mode with a cause.** Reports of a distilled controller acquiring a positive feedback gain that is then repaired by hand are, on inspection, symptoms of a degenerate experiment design. With a single constant reference the policy regressor is exactly rank-deficient, so individual coefficients are not identifiable and their signs are settled by numerical jitter rather than by data.
3. **The network is not compared against fitting its own output class directly.** If the deployed object is a polynomial, the question that decides whether the network earned its place is whether it produces a better polynomial than sparse regression on the same data at the same arithmetic cost.

1.4. Contributions

1. A **parameterization with structural guarantees** (Section 4.3): an analytic feedforward term, a detuned linear quadratic regulator (LQR) core, and a learned correction multiplied by a gate that vanishes quadratically at the setpoint. Because the gate and its gradient vanish there, the closed-loop linearization at every reachable reference is exactly $A(r) - BK$ and the steady-state input is exactly $u_{\mathrm{eq}}(r)$ for any learned term. Zero steady-state offset and local exponential stability are therefore properties of the architecture rather than of the quality of the fit.
2. A **two-step symbolic read-out** in which the KAN supplies only the monomial support, and the coefficients are re-estimated by a convex quadratic program. The physical requirement that a positive level error must not reduce the corresponding pump command is

imposed as a set of affine inequalities on the coefficients, evaluated on a dense sample of the operating box (Section 4.5). The requirement is thus enforced by construction rather than by manual sign correction after the fact: the procedure has a unique solution and takes no expert input, and it is expressible only because the read-out is linear in its parameters.

3. An **error decomposition** that reports MPC→KAN and KAN→symbolic errors separately, together with the coefficient of determination of every individual edge's symbolic fit, so that badly represented activations can be located (Section 5.1).
4. A **matched-budget comparison** against sparse polynomial regression, a multi-layer perceptron (MLP), a DeepONet and a gain-scheduled LQR, in open loop and in closed loop, from which we report a negative result: across eleven term budgets in two operating regimes the KAN-selected support is statistically indistinguishable from one chosen directly by sparse regression (Section 5.2).
5. A **diagnosis of the sign pathology** that reproduces it from the rank structure of the distillation design, quantifies how far a manual sign flip moves the identifiable part of the policy, and shows that the constrained read-out removes the pathology at a measured, small cost (Section 5.7).
6. A **stability, robustness and constraint-satisfaction assessment** that states plainly what is guaranteed (local exponential stability, structurally, and verified through the analytic closed-loop Jacobian) and what is only numerical evidence (region-of-attraction estimate, sampled Lyapunov decrease, Monte-Carlo robustness). A scalar derivative test is not treated as formal verification anywhere in this account (Section 5.5). The deployed complexity is itself selected by the certificate rather than merely checked against it.

Two properties of the benchmark are settled before any distillation result is reported. First, the reference such as $[10, 10, 2, 2]$ is not an equilibrium of the plant at all, so no controller can hold it (Section 3.1); the reachable set has to be characterized before a tracking target is chosen. Second, the teacher's prediction horizon must exceed the plant's inverse response, which in the non-minimum-phase (NMP) configuration excludes horizons of a few seconds by a factor of twenty (Section 4.1). Either condition, violated, prevents the closed loop from converging whatever is distilled from it.

1.5. Scope of the study

This study is entirely software-in-the-loop. We make no measured claims about execution time on any particular processor. What is claimed about computational cost is structural and is established by counting: the deployed control law is a fixed sequence of multiply–accumulate operations with no data-dependent branching, whereas an iterative quadratic-programming solver performs a number of iterations that depends on the active constraint set and therefore on the state. That distinction survives any choice of hardware. The quadruple-tank process itself has dominant time constants of tens of seconds and does not need fast control; it is used here because it is a well-characterized benchmark on which the approximation question can be studied cleanly, not as evidence that any plant needs microsecond actuation.

The NMP configuration is chosen deliberately. A transmission zero in the open right half-plane makes the plant respond initially in the direction opposite to its steady-state target (Skogestad and Postlethwaite, 2005), so the optimal policy is nonlinear and cannot be recovered by a controller that reasons only about the next few samples. That is what makes it a demanding target for a static approximant — and, as Section 4.1 shows, what makes the teacher's own horizon a question that has to be settled before any distillation result means anything.

The remainder of the paper is organized as follows. Section 2 states the problem and recalls the Kolmogorov–Arnold representation. Section 3 describes the plant, including its equilibrium manifold, which determines which references are reachable at all. Section 4 details the distillation pipeline. Section 5 reports the experiments, and Section 6 discusses what the results support and what they do not.

2. Problem Formulation

2.1. The Kolmogorov–Arnold representation

The Kolmogorov–Arnold representation theorem (Kolmogorov, 1957; Arnold, 1957) states that any continuous multivariate function $f : [0,1]^n \to \mathbb{R}$ admits a representation as a finite composition of continuous univariate functions,

$$f(x) = \sum_{q=1}^{2n+1} \Phi_q\left(\sum_{p=1}^{n} \phi_{q,p}(x_p)\right), \tag{1}$$

where Φ_q and $\phi_{q,p}$ are univariate. This is a statement about exact representation with a specific compositional structure, and it is not in competition with the universal approximation theorem, which concerns approximation by shallow networks of fixed activations; either result can be invoked to justify a KAN's expressive power. What the Kolmogorov–Arnold form contributes here is purely architectural: because the learnable objects $\phi_{q,p}$ are univariate functions on the edges, each can be replaced independently by a symbolic primitive, and the composition can then be expanded in closed form. Following Liu et al. (2024), the $\phi_{q,p}$ are parameterized as B-splines during training.

2.2. Control objective

The plant is a multiple-input multiple-output (MIMO) system: two pump voltages act on four liquid levels, with strong cross-coupling and, in one valve configuration, a right-half-plane transmission zero. Let $x \in \mathbb{R}^4$ denote the tank levels, $u \in \mathbb{R}^2$ the pump voltages and $r \in \mathbb{R}^4$ a

commanded reference. The objective is to regulate the two upper levels h_1, h_2 to their references under input saturation $0 \le u_i \le 12$ V, level limits $0 \le h_i \le 20$ cm, measurement noise and unmodelled actuator degradation.

An LTV-MPC solving this problem online is treated as the teacher. The question addressed is how well a controller of the form

$$u = \mathrm{sat}\Big(\underbrace{u_{\mathrm{eq}}(r)}_{\text{feed-forward}} + \underbrace{Ke}_{\text{LQR core}} + \underbrace{w(e)\,U_{\max}\,f_\theta\big(\varphi(x,r)\big)}_{\text{learned correction}} \Big), \tag{2}$$

in which $e = r - x$ is the tracking error, $\varphi(x, r)$ a fixed normalized feature map of the state and reference, $U_{\max} = 12$ V the actuator range, and f_θ an explicit closed-form expression, can replace it and what has to be verified before it may. The three terms, the feature map φ and the gate w are all specified in Section 4.3; the essential point is that the split is what makes the resulting controller certifiable.

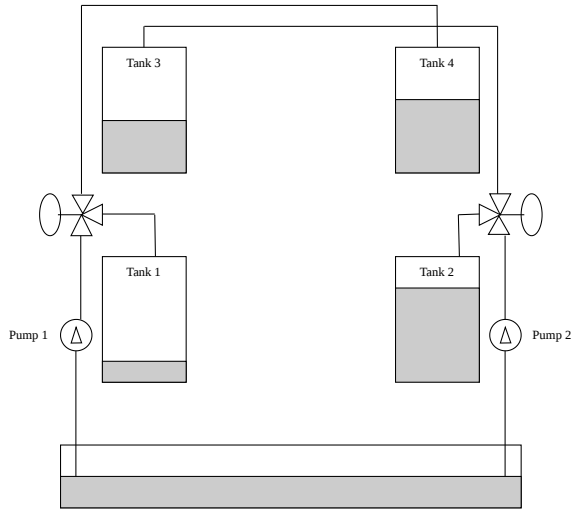

Figure 1: Quadruple-tank process. Four interconnected tanks and two pumps with strong cross-coupling. The valve split $\gamma_1 + \gamma_2$ selects the minimum-phase or non-minimum-phase configuration.

3. Plant Model and Reachable References

The quadruple-tank process (Johansson, 2000) consists of four interconnected tanks driven by two pumps through two three-way valves. Its dynamics follow from Bernoulli's principle and Torricelli's law:

$$\begin{aligned}
\dot{h}_1 &= -\tfrac{a_1}{A_1}\sqrt{2gh_1} + \tfrac{a_3}{A_1}\sqrt{2gh_3} + \tfrac{\gamma_1 k_1}{A_1}u_1,\\
\dot{h}_2 &= -\tfrac{a_2}{A_2}\sqrt{2gh_2} + \tfrac{a_4}{A_2}\sqrt{2gh_4} + \tfrac{\gamma_2 k_2}{A_2}u_2,\\
\dot{h}_3 &= -\tfrac{a_3}{A_3}\sqrt{2gh_3} + \tfrac{(1-\gamma_2)k_2}{A_3}u_2,\\
\dot{h}_4 &= -\tfrac{a_4}{A_4}\sqrt{2gh_4} + \tfrac{(1-\gamma_1)k_1}{A_4}u_1,
\end{aligned} \tag{3}$$

with $g = 981$ cm/s^2. Fig. 1 shows the plant and Table 1 the parameters, taken from Johansson (2000).

Table 1
Plant parameters for the MP and NMP configurations, following Johansson (2000). Initial states are those used for the nominal experiments.

Parameter	Symbol	MP	NMP
Valve ratios	γ_1, γ_2	$0.70, 0.60$	$0.43, 0.34$
Pump gains (cm^3/Vs)	k_1, k_2	$3.33, 3.35$	$3.14, 3.29$
Outlet cross-sections (cm^2)	a_i	$[0.071, 0.057, 0.071, 0.057]^\top$	
Tank cross-sections (cm^2)	A_i	$[28, 32, 28, 32]^\top$	
Initial state (cm)	x_0	$[12.4, 12.7, 1.8, 1.4]^\top$	$[12.6, 13.0, 4.8, 4.9]^\top$
Transmission zeros (1/s)		$-0.0194,\ -0.0674$	$+0.0146,\ -0.0634$

3.1. The equilibrium manifold

With two pumps the equilibrium set is two-dimensional: commanding $(h_1^{\mathrm{ref}}, h_2^{\mathrm{ref}})$ fixes the steady inputs through

$$\begin{bmatrix} \gamma_1 k_1 & (1-\gamma_2)k_2 \\ (1-\gamma_1)k_1 & \gamma_2 k_2 \end{bmatrix} u_{\mathrm{eq}} = \begin{bmatrix} a_1\sqrt{2gh_1^{\mathrm{ref}}} \\ a_2\sqrt{2gh_2^{\mathrm{ref}}} \end{bmatrix}, \tag{4}$$

and the lower levels then follow as

$$\begin{aligned}
h_3^{\mathrm{eq}} &= \big((1-\gamma_2)k_2 u_{\mathrm{eq},2}/a_3\big)^2/2g,\\
h_4^{\mathrm{eq}} &= \big((1-\gamma_1)k_1 u_{\mathrm{eq},1}/a_4\big)^2/2g.
\end{aligned} \tag{5}$$

They cannot be assigned independently. For the NMP configuration at $h_1^{\mathrm{ref}} = h_2^{\mathrm{ref}} = 10$ cm this gives $x_{\mathrm{eq}} = [10,\ 10,\ 4.16,\ 3.44]^\top$ cm with $u_{\mathrm{eq}} = [2.61,\ 2.95]^\top$ V.

Consequently, a commanded state like $[10, 10, 2, 2]^\top$ is not an equilibrium in the NMP regime, and no controller — MPC included — can drive the tracking error to zero for it. Every reference used in this paper is therefore the exact solution of (4), which is what allows the quadratic cost to converge to zero rather than to a constant offset. Equation (4) additionally supplies the feedforward term of (2): two square roots and four multiply–accumulate operations, exact and physically interpretable.

Linearizing about x_{eq} and taking $y = [h_1, h_2]^\top$ gives transmission zeros $\{-0.0194, -0.0674\}$ s^{-1} for the MP configuration and $\{+0.0146, -0.0634\}$ s^{-1} for the NMP configuration, confirming the right-half-plane zero of the latter, with an associated time constant of 69 s.

3.2. Linearization for the teacher

To keep the teacher's optimization convex (Boyd and Vandenberghe, 2004), the nonlinear model is linearized at each sampling instant about the current estimate, giving an affine LTV model

$$x_{k+1} = A_k x_k + B_k u_k + d_k, \qquad y_k = C x_k, \tag{6}$$

with $A_k = I + \Delta t\, \partial f/\partial x|_{\hat{x}_k}$, $B_k = \Delta t\, B_c$ and d_k the linearization offset, where

$$\frac{\partial f}{\partial x} = \begin{pmatrix} -\frac{a_1}{A_1}\sqrt{\frac{g}{2h_1}} & 0 & \frac{a_3}{A_1}\sqrt{\frac{g}{2h_3}} & 0 \\ 0 & -\frac{a_2}{A_2}\sqrt{\frac{g}{2h_2}} & 0 & \frac{a_4}{A_2}\sqrt{\frac{g}{2h_4}} \\ 0 & 0 & -\frac{a_3}{A_3}\sqrt{\frac{g}{2h_3}} & 0 \\ 0 & 0 & 0 & -\frac{a_4}{A_4}\sqrt{\frac{g}{2h_4}} \end{pmatrix}, \tag{7}$$

$$B_c = \begin{pmatrix} \frac{\gamma_1 k_1}{A_1} & 0 \\ 0 & \frac{\gamma_2 k_2}{A_2} \\ 0 & \frac{(1-\gamma_2)k_2}{A_3} \\ \frac{(1-\gamma_1)k_1}{A_4} & 0 \end{pmatrix}. \tag{8}$$

The benchmark assumes full state measurement, $C = I_4$, so that the comparison between controllers is not confounded by observer transients; structural observability of the standard partial-output configuration $y = [h_1, h_2]^\top$ was verified separately and is consistent with Johansson (2000).

4. Method

4.1. The teacher: LTV-MPC

This section provides a comprehensive specification of the LTV-MPC, specified as the creator of the nominal trajectory, that is, a teacher.

The prediction horizon must span the inverse response. The plant's dominant time constants are 50–70 s and the non-minimum-phase transmission zero sits at 0.0146 s^{-1}, i.e. a time constant of 69 s. The consequence is not a small loss of optimality. A horizon shorter than the inverse response makes the controller effectively myopic, and on a non-minimum-phase plant a myopic controller moves in the wrong direction: a 3 s horizon drives the closed loop steadily away from a reachable equilibrium, ending at $[9.33, 11.06, 1.06, 7.93]$ cm against a reference of $[10, 10, 4.16, 3.44]$ cm after 200 s, and continuing to drift. Widening the horizon fixes it: at $\Delta t_p = 1, 2, 4$ s (horizons of 30, 60, 120 s) the final state is $[9.53, 10.76, 2.98, 4.80]$, $[9.77, 10.40, 3.89, 3.71]$ and $[9.93, 10.20, 4.13, 3.45]$ cm respectively.

This paper therefore uses $N = 30$ at $\Delta t_p = 4$ s, a 120 s horizon, while the control loop still runs at $\Delta t = 0.1$ s — a coarse prediction grid with a fine control period. We record this requirement explicitly because it interacts with everything downstream: a teacher that does not converge cannot be distilled into a student that does.

The optimal control problem. At each instant the controller solves

$$\begin{aligned} \min_{x_{0:N},\, u_{0:N-1}} \quad & \sum_{k=0}^{N-1} \left[(x_k - r)^\top Q\, (x_k - r) + u_k^\top R\, u_k \right] \\ & + (x_N - r)^\top Q\, (x_N - r) \\ \text{s.t.} \quad & x_{k+1} = A_k x_k + B_k u_k + d_k, \quad x_0 = \hat{x}, \\ & 0 \le u_k \le 12\ \text{V}, \quad 0 \le x_{k+1} \le 20\ \text{cm}, \end{aligned} \tag{9}$$

with $Q = \mathrm{diag}(40, 40, 5, 5)$, $R = 10^{-3} I_2$, $N = 30$ and $\Delta t_p = 4$ s. The problem is specified in CVXPY (Diamond and Boyd, 2016) and solved with OSQP (Stellato et al., 2020). State estimation uses a canonical extended Kalman filter (EKF) (Kalman, 1960) with $H = I_4$.

4.2. Dataset design

Because the identifiability of a symbolic read-out depends entirely on how the distillation data are excited, the dataset design is stated in full.

For each regime, samples come from two sources. On-policy: 24 closed-loop LTV-MPC trajectories of 60 s at $\Delta t = 0.1$ s, each with an independently drawn reference and initial condition. Off-policy: 6000 independent and identically distributed draws of (x, r) from the operating box, each labeled by a single MPC solve. References are drawn uniformly from the box $[7, 15]^2$ cm and mapped to exact equilibria through (4); initial levels are drawn uniformly from $[1, 18]$ cm (upper tanks) and $[0.5, 9]$ cm (lower tanks).

Two design choices deserve comment.

Why the reference must vary. If the reference is held constant, as in a single regulation trajectory, the error features are exact affine functions of the level features, $e_i = r_i - h_i$ with r_i constant. The 8-dimensional regressor plus bias then has rank 5, not 9: only the differences between each level gain and the corresponding error gain are identifiable, and the four remaining degrees of freedom are fixed by numerical noise the fitting procedure happens to fit. This is the mechanism analyzed in Section 5.7.

Why the lower-tank references are jittered off the manifold. On the equilibrium manifold both h_3^{eq} and h_4^{eq} have the form $\alpha h_1 + \beta h_2 + \gamma\sqrt{h_1 h_2}$, so one fixed linear combination of the four reference levels is constant for every reachable reference and the regressor remains rank-8. The off-policy samples therefore perturb the commanded $h_3^{\mathrm{ref}}, h_4^{\mathrm{ref}}$ by ± 1.5 cm, which restores full rank and, as a by-product, teaches the policy how to respond to inconsistent commands.

Splits are made by trajectory and by block, not by row, so that no test sample is temporally adjacent to a training sample. Table 2 reports the resulting sizes and conditioning, alongside a single-trajectory design at fixed reference for contrast.

Table 2
Design of the policy-distillation datasets. The regressor is the 8-dimensional policy input augmented with a bias column; its numerical rank determines whether the individual coefficients of a linear symbolic read-out are identifiable at all. $c(e_1, e_2)$ is the Pearson correlation between the two error features and $\kappa(\Phi)$ the condition number of the regressor Φ; both measure how far the design is from degeneracy. A single regulation trajectory towards a fixed reference is shown for contrast: it is the cheapest design to collect and the one on which the read-out is not identifiable.

Dataset	Total	Train	Val.	Test	On-policy	Off-policy	Groups	$c(e_1, e_2)$	$\kappa(\Phi)$	rank
Single trajectory, fixed r	500	400	–	100	500	0	1	0.964	2.0×10^{17}	5/9
MP	20400	14100	2900	3400	14400	6000	36	0.050	91	9/9
NMP	20400	14200	2900	3300	14400	6000	36	-0.457	154	9/9

4.3. Policy parameterization and its structural guarantees

Every surrogate is embedded in the same skeleton (2), with the 8-dimensional policy input

$$\varphi(x, r) = \left[\frac{x}{X_n},\ \frac{r-x}{X_n}\right] \in \mathbb{R}^8, \qquad X_n = 20\ \text{cm}, \tag{10}$$

and learns only the residual f_θ that the first two terms leave unexplained. Each term has a distinct role.

$u_{\text{eq}}(r)$ is the exact steady-state input of (4): two square roots and four multiply–accumulate operations. It removes the steady-state offset that a static approximate policy inherits directly from its imitation error. Without such a term the distilled controller settles wherever its residual command error places it, and does not reach the commanded level at all — an offset that no amount of additional training data removes, because it is a property of approximating a policy rather than of approximating it badly.

K is a discrete LQR gain for the plant linearized at the nominal equilibrium, solved once offline and held constant. It is deliberately detuned relative to the teacher's own input weight: with $R = 10^{-3}$ the unconstrained LQR gain exceeds 100 V/cm and saturates on any appreciable error, leaving nothing for the learned term to do, so the core uses $R_{\text{core}} = I$ and produces gains of a few volts per centimeter. The core supplies stability; the learned term supplies the predictive, constraint-aware behavior that distinguishes MPC from LQR.

The gate

$$w(e) = \min\left(\frac{\|e\|^2}{X_n^2 s^2},\ 1\right), \qquad s = 0.25, \tag{11}$$

vanishes quadratically as $e \to 0$. This is the device that turns a heuristic into a guarantee. Because both w and ∇w vanish at $e = 0$, the learned correction and its Jacobian vanish at every reachable setpoint, for any f_θ whatsoever. Two consequences follow immediately and require no assumption about the quality of the fit:

(P1) **Zero steady-state offset.** At a reachable reference the commanded input is exactly $u_{\text{eq}}(r)$, so $x_{\text{eq}}(r)$ is an exact fixed point of the closed loop.

(P2) **Structural local stability.** The closed-loop linearization at that fixed point is exactly $A(r) - BK$, independent of what was learned. Local exponential stability therefore follows from LQR theory rather than from a property of the approximation, and is verified numerically at six setpoints in Section 5.5.

Away from the setpoint $w \to 1$ and the learned term takes over, which is precisely where the MPC's behavior differs from a linear controller. All surrogates — KAN, MLP, DeepONet, polynomial regression — use the identical skeleton, so the comparison between them is a comparison of function approximators and nothing else.

4.4. KAN training

The network is a $[8, 5, 2]$ KAN with cubic B-splines on a grid of 8 intervals, initialized with seed 42, trained with L-BFGS for 40 steps ($\lambda = 10^{-4}$, $\lambda_{\text{entropy}} = 2.0$), pruned at edge threshold 5×10^{-3} and node threshold 10^{-3}, then retrained for 25 further steps. Following Liu et al. (2024), pykan trains with full-batch L-BFGS Wright et al. (1999) throughout. Training a single network takes 2–4 minutes on a consumer CPU (AMD Ryzen 7, 16 GB RAM). Fig. 2 shows the topology.

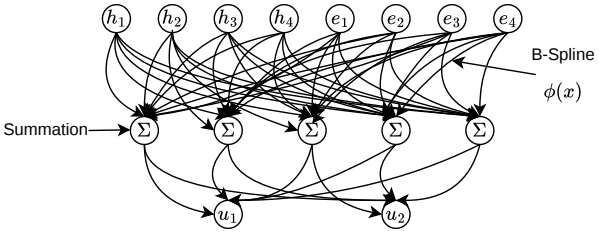

Figure 2: Network topology. Four tank levels h_i and four level errors e_i enter the $[8, 5, 2]$ KAN; the two outputs are the deviations of the pump commands from the analytic feedforward input.

4.5. Symbolic read-out

The read-out is deliberately split into two steps with different roles.

Step 1: structure selection. Each surviving edge is replaced by the best-fitting candidate from the declared library $\{x,\ x^2\}$, chosen by the coefficient of determination of the univariate fit and not weighted towards simplicity, and the resulting composition is expanded symbolically. The library is restricted to the polynomial family so that the deployed

law is a fixed-length multiply–accumulate chain with no transcendental calls, which is what makes its execution time constant. Expanding a two-layer composition of quadratics yields monomials of total degree at most 4; the union of the monomials that appear is the KAN-selected support S, and it is the only thing carried forward from the network.

Step 2: shape-constrained coefficient fit. The coefficients on S are then re-estimated from the training data, rather than inherited from the network. Let $m(\varphi)$ collect the monomials of S, so that the deployed law is $u_j = u_{\mathrm{eq},j}(r) + (Ke)_j + w(e)\,U_{\max}\,c_j^{\top} m(\varphi)$. The read-out solves, for each pump independently,

$$\begin{aligned} \min_{c_j} \quad & \left\| W A\, c_j - y_j \right\|_2^2 + \varepsilon \|c_j\|_2^2 \\ \text{s.t.} \quad & \left.\frac{\partial u_j}{\partial e_j}\right|_{\varphi^{(l)}} \geq 0, \qquad l = 1, \ldots, L, \end{aligned} \tag{12}$$

where A is the monomial design matrix on the training set, W the diagonal matrix of gate values $w(e^{(i)})$, and $\{\varphi^{(l)}\}$ are $L = 1600$ points drawn from the operating box (half resampled training points, half uniform). The constraint states that a positive error on level j must not decrease pump j's command — elementary negative feedback. Writing it out,

$$\frac{\partial u_j}{\partial e_j} = K_{jj} + \frac{U_{\max}}{X_n}\left[\frac{\partial w}{\partial \varphi_{j+4}}\, m(\varphi)^{\top} + w\, \frac{\partial m}{\partial \varphi_{j+4}}^{\top}\right] c_j, \tag{13}$$

which is affine in c_j, because differentiating a polynomial in the monomial basis is linear in its coefficients. Problem (12) is therefore a convex quadratic program with a unique solution, solved here with CLARABEL through CVXPY. No expert judgement, and in particular no manual sign editing, enters at any point.

It is worth noting what this construction relies on. The constraint is expressible because the read-out is linear in its parameters; the same requirement applied to a multilayer perceptron would be a non-convex constraint on a non-convex problem. This is the concrete sense in which a symbolic read-out buys something that a black-box approximator of equal accuracy does not.

We stress the modest status of this constraint. It is a shape restriction on the control law, in the spirit of monotone regression (Gupta et al., 2016); it is not a stability proof, and Section 5.5 treats stability separately. Its role is to exclude a class of physically absurd read-outs by construction rather than to detect them after the fact.

Simplification and budget selection. Deploying all 145 expanded monomials is unnecessary. For a term budget k, orthogonal matching pursuit (OMP) (Pati et al., 1993) selects the k most useful monomials of S and (12) is re-solved on that subset. The deployed budget is not chosen by imitation accuracy alone: because the law is an explicit polynomial, the closed-loop Jacobian of every candidate can be formed and its spectral radius computed, so the selection rule is the smallest budget that is certified locally exponentially stable at all six test setpoints and whose imitation error is within one percentage point of the best certified candidate. The verification is thus inside the design loop, and the law that ships is the law that was certified. All candidate budgets are reported in Section 5.2.

4.6. Gain scheduling and its realizability

Two laws are distilled, one per valve configuration, and the active law is selected by $\gamma_1 + \gamma_2 \gtrless 1$. On the quadruple-tank rig the split ratios are set by motorized three-way valves, so γ_i is an actuator setpoint known to the controller, not an unmeasured parameter; a plant reconfiguration is a commanded move of those valves over a finite time. Section 5 therefore ramps γ from the NMP to the MP configuration over 10 s and lets the scheduler read the commanded valve position. Where the ratios are fixed by hardware, the scheduler degenerates to a compile-time constant and the question does not arise.

5. Experimental Results

Unless stated otherwise, results are software-in-the-loop simulations at $\Delta t = 0.1$ s with the EKF and measurement noise ($\sigma^2 = 2.37$) in the loop.

5.1. Imitation fidelity and where the error comes from

Table 3 reports open-loop imitation of the MPC policy on the held-out split for every surrogate, and Fig. 3 isolates the contribution of each stage of the read-out.

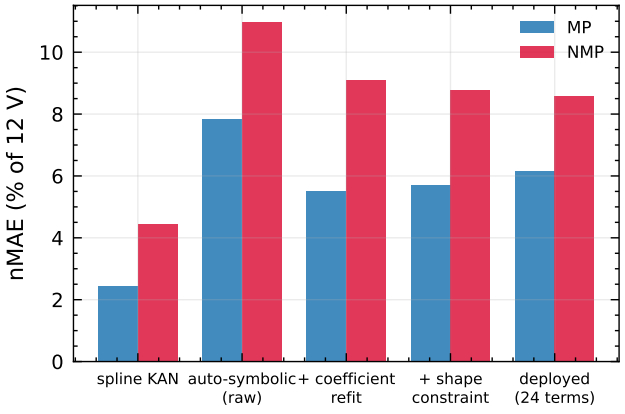

Figure 3: Where the approximation error is introduced. The spline KAN imitates the MPC policy well; the off-the-shelf symbolic extraction loses most of that fidelity; re-estimating the coefficients on the KAN-selected support recovers the bulk of it; the shape constraint costs little and in the NMP regime helps; the sparsified deployed law trades a further small amount of accuracy for a large reduction in arithmetic.

The decomposition is the main empirical message of this section. In the MP regime the spline KAN imitates the teacher to 2.43% of the actuator range, and pykan's `auto_symbolic` then degrades that to 7.85% — a factor of 3.2 lost in a single step, and a step that reporting the network's

Table 3
Open-loop imitation of the LTV-MPC policy on the held-out test split. nMAE is the mean absolute command error as a percentage of the full actuator range ($U_{\max} - U_{\min}$ = 12 V); predictions are clipped to the actuator box before scoring, as they are on the target. Every row is trained on the same data, the same split and the same feed-forward/feedback decomposition.

		MP regime		NMP regime	
Surrogate	**Params/terms**	**nMAE (%)**	R^2	**nMAE (%)**	R^2
KAN (spline, before symbolic read-out)	–	2.43	0.986	4.45	0.959
Symbolic KAN, `auto_symbolic` output	–	7.85	0.875	10.97	0.782
Symbolic KAN + coefficient refit	–	5.51	0.935	9.11	0.862
Symbolic KAN + refit + shape constraint	–	5.71	0.929	8.77	0.866
MLP 8–32–32–2	1410	3.19	0.979	4.15	0.958
MLP 8–8–2	90	8.44	0.850	11.20	0.780
DeepONet (branch/trunk)	17282	4.11	0.965	6.82	0.922
Polynomial ridge, degree 2	45	7.27	0.870	9.91	0.805
Polynomial ridge, degree 3	165	5.71	0.922	7.53	0.880
Polynomial ridge, degree 4	495	5.27	0.935	8.92	0.867
Sparse polynomial (OMP, matched budget)	290	5.29	0.937	8.30	0.878
Deployed symbolic law (4 terms)	8	**6.16**	–	**8.58**	–

error alone renders invisible. The NMP regime behaves the same way, 4.45% becoming 10.97%. Re-estimating the coefficients on the KAN-selected support recovers most of the gap, to 5.51% and 9.11% respectively.

The loss is a structural consequence of extracting symbols after training: nothing in the training objective encourages the learned activations to lie close to the declared library. This is visible at the level of individual edges. Across the 41 (MP) and 48 (NMP) surviving activations, the coefficient of determination of the symbolic fit averages 0.78 and 0.77 but falls as low as 0.088 and 0.024, so a handful of activations are essentially unrepresentable in the declared library and can be identified individually — the diagnostic that an aggregate error figure hides. Approaches that fold symbolic structure into training (Bagrow and Bongard, 2025; Faroughi et al., 2026; Howard et al., 2026) attack the same gap at its source and are the natural next step; the convex re-estimation used here is a cheap alternative that requires no change to the training procedure.

In the NMP regime the shape-constrained fit (8.77%) is more accurate than the unconstrained one (9.11%), and the deployed 48-term law (8.58%) is more accurate than the fit on the full 145-monomial support. Neither is a contradiction: the constraint and the term budget both act as regularizers on an estimator that would otherwise overfit, and the network's contribution is the monomial support, not the coefficients on it.

5.2. Accuracy against arithmetic cost, and against direct regression

If the deployed object is a polynomial, the honest baseline is polynomial regression. Fig. 4 sweeps the term budget for two supports: the KAN-selected one, and one chosen by orthogonal matching pursuit over the full degree-4 dictionary of 495 monomials — that is, fitting the polynomial directly. Both use the same data, the same solver and the same shape constraint.

We report this negative result plainly, because it bears directly on what a symbolic KAN is for. Across the eleven budgets and two regimes the two curves are within a few tenths of a percentage point of one another. The KAN-selected support is ahead at four of eleven budgets in the MP regime and one of eleven in the NMP regime; the directly regressed support is ahead at two and five respectively; the rest are ties. On this problem, in other words, the KAN's contribution to support selection is not measurable: orthogonal matching pursuit over the full 495-monomial dictionary finds essentially as good a support without any network at all.

The two supports have to be compared at matched term and arithmetic budget, with the same solver and the same constraint, before the comparison means anything. Meanwhile the dense MLP is at least as accurate as any symbolic form in both regimes (3.19% and 4.15%), at the cost of admitting neither the shape constraint of (12) nor the certified budget selection of Section 4.5. What survives of the case for the symbolic law is its auditability and its fixed arithmetic cost — not its accuracy, and not the network that produced it.

5.3. Closed-loop performance

Open-loop imitation error is not the quantity of interest; closed-loop behavior is. Table 4 compares all controllers on the nominal regulation task and Fig. 5 shows the trajectories in the NMP configuration.

On this single-setpoint regulation task the controllers are close: in the NMP regime the MPC achieves 1.004 cm RMSE, against 1.050 cm for the MLP, 1.098 cm for the gain-scheduled LQR and 1.160 cm for the deployed symbolic law — a spread of 16% between best and worst. The plant

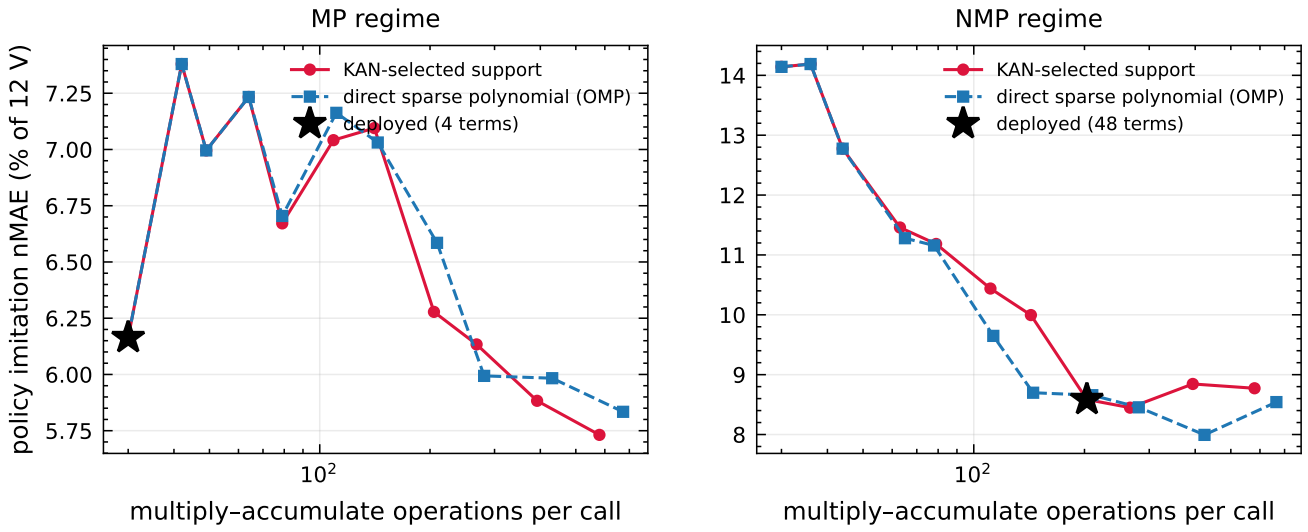

Figure 4: Imitation accuracy against arithmetic cost. The two curves track each other at every budget: the KAN-selected support offers no measurable advantage over choosing the same number of monomials directly by orthogonal matching pursuit over the full degree-4 dictionary. The star marks the deployed operating point, chosen by the stability certificate rather than by accuracy alone.

Table 4
Closed-loop tracking of the reachable equilibrium of $(h_1, h_2) = (10, 10)$ cm from the Johansson initial state, with the EKF and measurement noise in the loop (120 s, $\Delta t = 0.1$ s). RMSE and steady-state error are over the controlled levels h_1, h_2; TV is the total variation of the two pump commands.

	MP regime			NMP regime		
Controller	**RMSE (cm)**	$\lvert e_{ss}\rvert$ **(cm)**	**TV (V)**	**RMSE (cm)**	$\lvert e_{ss}\rvert$ **(cm)**	**TV (V)**
LTV-MPC	0.50	0.04	16.9	1.00	0.67	16.5
Symbolic KAN	0.55	0.04	11.0	1.16	0.69	7.5
MLP (32-32)	0.53	0.04	11.3	1.05	0.67	9.4
MLP (8)	0.54	0.04	11.2	1.16	0.69	7.9
DeepONet	0.54	0.04	11.2	1.09	0.70	8.8
Poly ridge (deg 2)	0.54	0.04	11.2	1.15	0.68	7.5
Poly ridge (deg 3)	0.54	0.04	11.2	1.12	0.68	7.8
Gain-scheduled LQR	0.53	0.04	11.3	1.10	0.66	8.0

is self-regulating, the initial error is modest, and a well-tuned linear controller is already close to the optimizer. Actuator activity separates them more clearly: total variation is 16.5 V for the MPC against 7.5 V for the symbolic law. The residual 0.59 cm steady-state error on h_2 is common to every controller, the MPC included, and comes from the estimator rather than the policy.

5.4. Scenario campaign

Table 5 summarizes three scenarios chosen so that they exercise structurally different mechanisms; the corresponding figures are Figs. 6–9. Every commanded reference lies on the equilibrium manifold of Section 3.1, which is a precondition for the campaign rather than a refinement of it. A scenario that commands an unreachable target settles at the nearest attainable point whatever mechanism it was meant to probe, so two nominally different experiments — an actuator degradation and an unseen reference, say — return near-identical trajectories differing only in the noise realization, and the campaign measures nothing. With reachable references the scenarios separate as they should.

The staircase is the most demanding scenario and, with a correctly tuned teacher, it produces an outstanding result we report as it stands. Over four successive references spanning the operating box the deployed symbolic law reaches 2.419 cm RMSE, marginally better than the MPC it was distilled from (2.461 cm), with 54.9 V of actuator travel against the MPC's 87.0 V. Every distilled policy lands in a narrow band (2.376–2.440 cm) and the gain-scheduled LQR trails at 2.605 cm, a gap of only 7–9%.

The interpretation matters more than the ranking. That a 30-operation explicit law matches a 120 s-horizon MPC on this task does not mean distillation is powerful; it means the optimizer's advantage over a linear controller on this plant is

Table 5
Scenario campaign, NMP regime. S2: four successive, partly asymmetric set-point changes. S3: +3 cm load disturbance on tank 1 at $t = 60$ s; recovery time is the first instant at which $|h_1 - h_1^{\mathrm{ref}}| < 0.1$ cm. S4: pump gains ramped down by 30% between $t = 40$ s and $t = 80$ s with no controller update.

	S2 staircase		S3 disturbance			S4 degradation	
Controller	**RMSE (cm)**	**worst e_{ss} (cm)**	**peak (cm)**	**rec. (s)**	**TV (V)**	**$\lvert e_{ss}\rvert$ (cm)**	**stable**
LTV-MPC	2.46	3.69	2.99	96.4	28.3	0.34	yes
Symbolic KAN	2.42	3.51	2.98	97.7	12.1	0.67	yes
MLP (32-32)	2.44	3.89	2.99	101.4	13.8	0.66	yes
Gain-scheduled LQR	2.61	3.75	2.99	100.5	12.8	0.69	yes
Poly ridge (deg 3)	2.42	3.68	2.98	100.6	12.4	0.68	yes

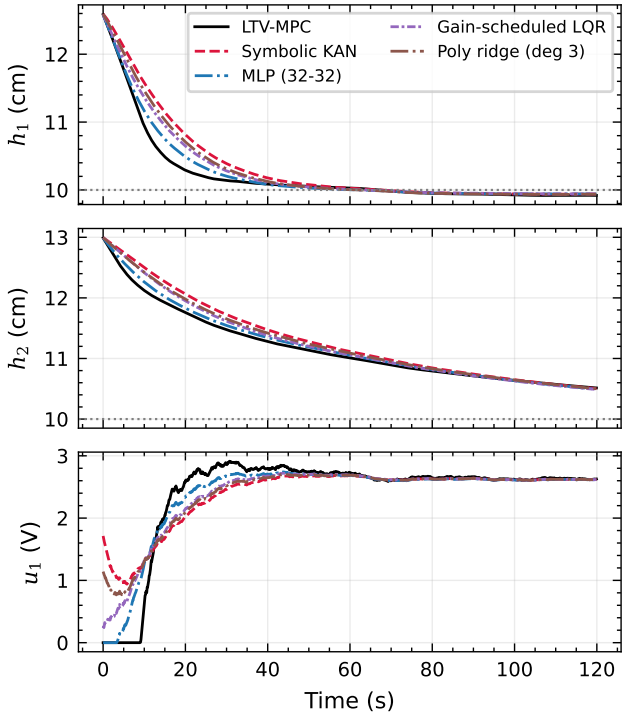

Figure 5: Nominal tracking, NMP regime. All controllers reproduce the inverse response of the non-minimum-phase plant. The residual steady-state error (about 0.6 cm on h_2) is common to every controller including the MPC and comes from the estimator, not from the policy: by property (P1) the explicit law commands exactly $u_{\mathrm{eq}}(r)$ at the setpoint.

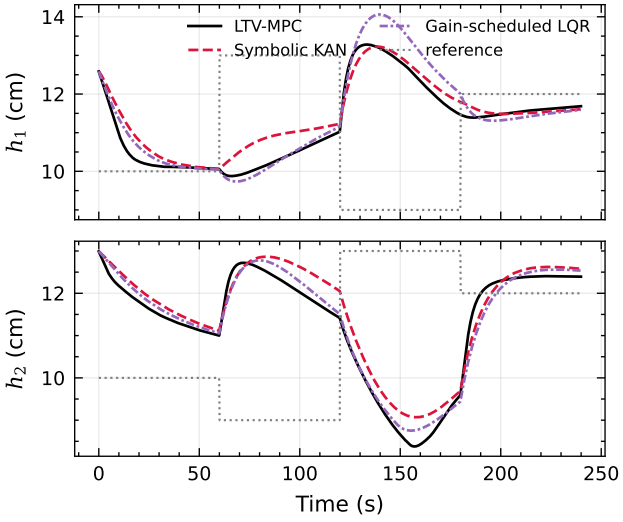

Figure 6: Setpoint staircase, NMP regime. Four successive references, two of them asymmetric ($h_1 \neq h_2$), none of them in the training trajectories. The explicit law follows each transition without retraining.

small to begin with, so there is little for the student to lose. That the symbolic law edges ahead of its teacher is likewise not evidence of superiority: a distilled policy is smoothed relative to the optimizer and, on a scenario dominated by repeated large transitions, less aggressive tracking scores marginally better on root-mean-square error while giving up on peak performance. The tractable conclusion is that this benchmark cannot separate the methods, and that the case for distillation has to be made on plants where online optimization genuinely outperforms a linear controller.

Disturbance rejection separates the controllers less than anticipated. Recovery from the +3 cm step takes 96.4 s for the MPC and 97.7–103.4 s for the explicit policies, the symbolic law being the fastest of them; actuator travel is 28.3 V for the MPC against 12.1 V for the symbolic law. All of them are limited by the same thing: the closed loop's dominant time constant, which for this plant is of the order of the tanks' own drainage time. The certified closed-loop time constant of the deployed law (Table 6) is set by the detuned core, and tightening that core recovers bandwidth at the cost of saturating during transients — which in turn leaves the learned term with nothing to do. That trade-off is the principal design decision in this architecture.

5.5. Stability, robustness and constraints

For a coupled nonlinear MIMO plant a sign condition on one partial derivative is neither necessary nor sufficient for stability; at most it excludes an obviously pathological class of control laws, which is the role it plays in (12). What can be established is set out below and in Table 6, ordered by decreasing strength.

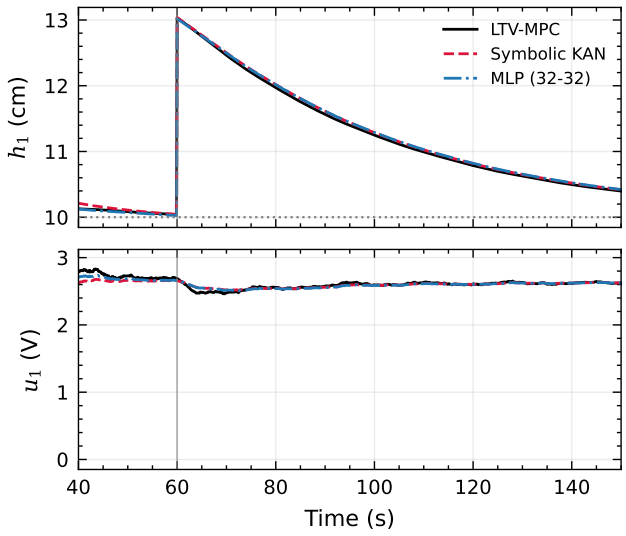

Figure 7: Load disturbance, NMP regime. A +3 cm step is injected on tank 1 at $t = 60$ s. The MPC returns to the setpoint far faster and with a much more aggressive command; every explicit policy — the LQR baseline included — recovers at roughly the plant's own drainage rate.

Local exponential stability (structural, then verified). By property (P2) of Section 4.3, the gate makes the learned correction and its Jacobian vanish at every reachable setpoint, so the closed-loop linearization there is exactly $A(r) - BK$ — independently of what the network learned. Local exponential stability is therefore a property of the parameterization, not of the fit, and Lyapunov's indirect method (Khalil, 2002) applies. We nonetheless recompute the analytic closed-loop Jacobian of the deployed law at six reachable setpoints and report its spectral radius in Table 6. By property (P1) the fixed point coincides with $x_{\mathrm{eq}}(r)$ to numerical precision, which is why the steady-state errors in Table 4 are attributable to noise and estimator bias rather than to policy bias.

Region of attraction. Solving the discrete Lyapunov equation for the certified Jacobian yields $V(x) = (x - x_{\mathrm{fp}})^{\top} P (x - x_{\mathrm{fp}})$. Its decrease condition is then checked on 2×10^4 samples of the operating box, and convergence is checked on a 19×19 grid of initial conditions (Fig. 10).

Parametric robustness (statistical). Valve ratios and pump gains are perturbed by ±20% over a 200-sample Monte-Carlo campaign, without retraining or retuning the controller. The closed loop converges in 97% (MP) and 100% (NMP) of trials, with no overflow in either and a 95th-percentile steady-state error of 0.66 and 1.79 cm respectively.

Invariance of the certificate (numerical check of the structural claim). Property (P2) asserts that the linearization at a setpoint does not depend on what was learned. Substituting for the trained read-out the zero function and random coefficient vectors with standard deviations of 1, 10 and 100 leaves the spectral radius unchanged to 5.5×10^{-11} (MP) and 1.1×10^{-10} (NMP), and $|u(x_{\mathrm{eq}}) - u_{\mathrm{eq}}|$ at machine precision ($< 10^{-15}$ V) in every case. The gate, not the fit, is what bears the guarantee.

Table 6
What is actually verified about the deployed law. ρ is the spectral radius of the closed-loop Jacobian at the fixed point, evaluated at six reachable set-points; $\rho < 1$ certifies local exponential stability (Lyapunov's indirect method), and $\tau = -\Delta t / \ln \rho$ is the corresponding slowest closed-loop decay time constant. The remaining columns are numerical evidence, not proofs: ROA is the fraction of a 19×19 grid of initial conditions that converges; V viol. is the fraction of 2×10^4 box samples violating the decrease condition of the local Lyapunov function; MC stable and p95 $|e|$ are the stable fraction and the 95th-percentile steady-state error over a 200-sample Monte-Carlo (MC) campaign with ±20% perturbation of valve ratios and pump gains.

Regime	$\max \rho$	τ (s)	loc. stable	ROA (%)	V viol. (%)	MC stable (%)	p95 $\lvert e \rvert$ (cm)	overflow (%)
MP	0.9984	62	yes	100.0	0.00	97.0	0.66	0.0
NMP	0.9988	81	yes	89.1	0.13	100.0	1.79	0.0

Constraint satisfaction. The LTV-MPC enforces $0 \le h_i \le 20$ cm as inequality constraints inside (9). The distilled law inherits only the input box, through clipping; it has no mechanism whatsoever for respecting state constraints. Table 7 quantifies what this costs over randomized setpoint changes from randomized initial conditions, measured after a 20 s settling window because a lower tank started near empty drains to zero whatever the pumps do.

Overflow was not observed for any controller. Dry-out of a lower tank does occur: in the MP regime the MPC empties a tank in 2.1% of tasks against 10.4% for the symbolic law and 12.5% for both the MLP and the gain-scheduled LQR. This is the clearest advantage the optimizer shows anywhere in this study, and it is precisely the property that distillation discards. In the NMP regime the difference disappears (0% for the MPC, the symbolic law and the MLP; 4.2% for the LQR), because the transitions there are slower and the lower tanks are held further from empty. For deployments where state constraints are safety-relevant, the explicit law should be wrapped in a supervisory projection or a control-barrier filter; we do not claim that clipping is sufficient. A separate practical safeguard is included in the deployed law: a polynomial extrapolates catastrophically outside the region it was fitted on, so the features are saturated to the box spanned by the training data before the polynomial is evaluated. Without that guard the law overflowed a tank in 6% of a comparable set of tasks; with it, in none.

5.6. Computational cost

The deployed law evaluates 4 monomials per pump in the minimum-phase regime and 48 in the non-minimum-phase regime, which together with the feed-forward and core terms is 30 and 203 multiply–accumulate operations per control step respectively. The count is exact and fixed:

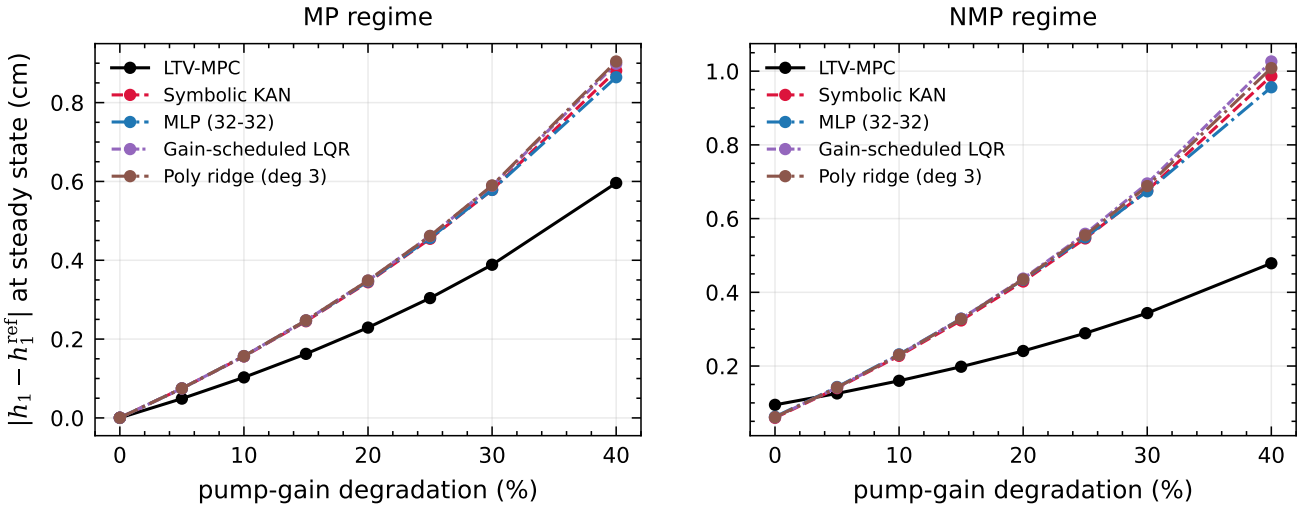

Figure 8: Robustness to unmodelled actuator degradation. Steady-state tracking error of h_1 as both pump gains are reduced, with no controller update. The MPC re-linearizes and re-optimizes online and so absorbs the mismatch almost entirely; the explicit policies accumulate an offset that grows with the mismatch.

Table 7
State-constraint satisfaction over randomized set-point changes from randomized initial conditions. The LTV-MPC enforces $0 \le h_i \le 20$ cm explicitly; every distilled policy inherits only the input box, through clipping. Percentages are the fraction of tasks in which some tank overflowed ($h > 20$ cm) or ran dry ($h \le 0$ cm).

	MP regime			NMP regime		
Controller	**overflow (%)**	**dry-out (%)**	**max h (cm)**	**overflow (%)**	**dry-out (%)**	**max h (cm)**
LTV-MPC	0.0	2.1	14.80	0.0	0.0	17.87
Symbolic KAN	0.0	10.4	17.14	0.0	0.0	18.36
MLP (32-32)	0.0	12.5	15.01	0.0	0.0	17.70
Gain-scheduled LQR	0.0	12.5	14.79	0.0	4.2	18.50

the law contains no loop whose trip count depends on the state, no convergence test and no branch other than the final saturation. Its worst-case cost therefore equals its average cost by construction.

This is the property that distinguishes it from the teacher. An operator-splitting solver such as OSQP (Stellato et al., 2020) performs a number of iterations that depends on the active constraint set and hence on the state, so its worst-case cost exceeds its typical cost by a margin that has to be established empirically for each plant and each tuning, and that is what has to be budgeted on a shared processor (Arnström and Axehill, 2022). Whether the resulting difference in wall-clock time matters depends entirely on the application: for the quadruple tank, with time constants of tens of seconds, it does not. We therefore make the structural claim of fixed arithmetic and no data-dependent control flow.

5.7. Case study: diagnosing the sign anomaly

The law, and the one-coefficient edit. The law examined here is a minimum-phase symbolic read-out obtained the off-the-shelf way: a KAN distilled from a single regulation trajectory and symbolized by `auto_symbolic`, yielding a polynomial in φ scaled by $U_{\max}$ and clipped to the pump range. It diverges in closed loop, and reading the printed expression suggests an obvious remedy — one coefficient carries the wrong sign, therefore flip it. Exhibiting the full expression is unnecessary to follow what happens. The defect lives entirely in the second channel, and within that channel in the two terms that carry level 2 — the coefficient α_2 on $\varphi_2 = h_2/X_n$ and the coefficient β_2 on $\varphi_6 = e_2/X_n$:

$$u_2 \;=\; U_{\max}\big[\;\cdots\;+\;\alpha_2\,\varphi_2\;+\;\beta_2\,\varphi_6\;+\;\cdots\;\big]. \qquad (14)$$

As extracted, $\alpha_2 = -6.2834$ and $\beta_2 = -50.5462$. The repair replaces β_2 by $+50.5462$ and leaves every other coefficient of both channels untouched. That single substitution is the whole of it, and it is what the remainder of this section analyzes: what the edited coefficient controls, why the fit put it there, and what the edit costs.

The edited coefficient is the whole error gain of the channel. The variable φ_6 appears in exactly one monomial of (14): the only quadratic term of this channel is built on φ_7, not on φ_6. Pump 2's response to its own level error is therefore an exact constant, identical at every point of the operating box,

$$\frac{\partial u_2}{\partial e_2} = \frac{U_{\max}}{X_n}\,\beta_2 = \begin{cases} -30.33\ \mathrm{V/cm}, & \text{as extracted,} \\ +30.33\ \mathrm{V/cm}, & \text{after the repair,} \end{cases}$$

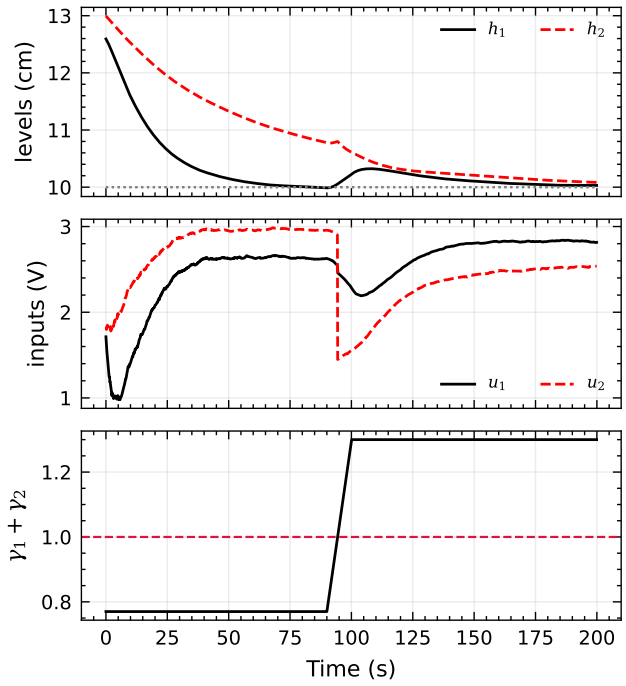

Figure 9: Gain-scheduling handover. The valve ratios are ramped from the NMP to the MP configuration over 10 s starting at $t = 90$ s, as a motorized valve would move. The scheduler switches when the commanded split crosses $\gamma_1 + \gamma_2 = 1$; the command is continuous across the handover and the levels converge to the new equilibrium.

(15)

with $U_{\max} = 12$ V and $X_n = 20$ cm. The extracted law thus answers a one-centimetre shortfall in tank 2 by commanding 30 V *less* on the pump that fills it: positive feedback on the channel's own error, everywhere rather than in some corner of the state space, and exactly the quantity that (12) bounds below by zero. Two further readings of (14) fix the status of the defect. First, the same law's other channel satisfies $\partial u_1/\partial e_1 = +8.14$ V/cm, correctly signed, so what is recorded here is one wrong coefficient rather than a wholesale failure of the extraction. Second, $|\beta_2|$ is comparable to the largest coefficient in the channel (+53.3293 on φ_5), so the offending term dominates the commanded voltage instead of perturbing it — which is why the consequences below are not marginal.

The regressor is rank-deficient. The distillation set featured in Fig. 11 is a single closed-loop trajectory towards the single constant reference $[10, 10, 2, 2]$. With r constant, $e_i = r_i - h_i$, so the eight features plus a bias column span a space of dimension 5, not 9, and the condition number of the design matrix is 2.0×10^{17} (Fig. 11, left). Only the four differences $\alpha_i - \beta_i$ between each level gain and the corresponding error gain — in the notation of (14) — are identifiable; the four remaining directions are a null space, and every member of it fits the training data identically.

The sign is therefore chosen by numerical noise. The fit is performed on inputs perturbed by Gaussian jitter of standard deviation 10^{-4}, a routine numerical safeguard. On a rank-deficient design that jitter is the only thing selecting a point inside the null space. Repeating the fit over 200 jitter seeds (Fig. 11, right) gives a coefficient on e_2 ranging from -36.3 to $+68.3$, with 18% of seeds producing the positive-feedback sign. On the excited designs of Table 2 the same coefficient is 1.292 ± 0.001 (MP) and 0.578 ± 0.003 (NMP), correctly signed for every seed. The "anomaly" is not a rare event to be caught by inspection; it is the expected behavior of an under-determined fit, and a different random seed would have hidden it entirely.

The manual repair is not a relabeling. The substitution made in (14) does not move along that null space — it changes the identifiable combination $\alpha_2 - \beta_2$ from +44.26 to −56.83, a change of 101.1 units. Measured against the MPC labels on the single-trajectory data, the extracted law attains a mean absolute command error of 4.87 V (40.6% of range) and the repaired law 3.07 V (25.6%); in closed loop the extracted law diverges to overflow while the repaired one converges to within 0.16 cm. The repair therefore trades imitation fidelity for stability. That is a defensible trade, but it is a trade, and a sign edit made by reading the expression measures neither side of it — although the interpretability of the symbolic form is exactly what makes both numbers easy to compute.

The systematic version of the problem. The isolated sign error is a special case of a general phenomenon. Even on the well-conditioned datasets of Table 2 an unconstrained symbolic read-out violates $\partial u_j/\partial e_j \geq 0$ somewhere in the operating box at essentially every term budget: up to 30.3% of sampled points in the MP regime and 30.1% in the NMP regime. Solving (12) instead of an unconstrained least-squares problem drives this to zero (MP) or at most 0.4% (NMP). The price is 0.20 percentage points of imitation nMAE on the full support in the MP regime, and in the NMP regime there is no price at all: the constrained fit is 0.34 points more accurate, the physical restriction acting as a regularizer. Detecting bad read-outs after the fact addresses a symptom; constraining the read-out addresses the cause, and does so at a cost that is measured rather than assumed.

6. Discussion

6.1. What the symbolic form does and does not provide

The results support a narrow claim: a symbolic KAN read-out delivers two things: a closed-form control law whose arithmetic cost is fixed and state-independent, and an object on which analytic questions can be asked. The closed-loop Jacobian of Section 5.5 exists because the law is a polynomial; the shape constraint of (12) is a linear inequality for the same reason; the sign diagnosis of Section 5.7 is a statement about identifiable coefficients, which presupposes

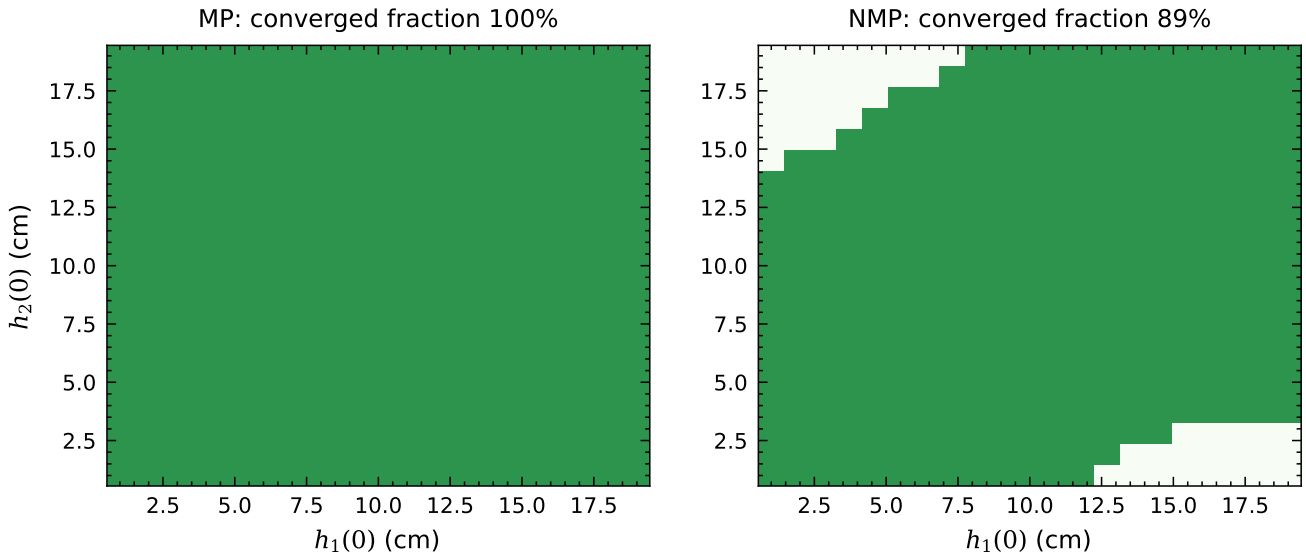

Figure 10: Estimated region of attraction. Initial conditions on a 19×19 grid of $(h_1(0), h_2(0))$, lower tanks at their equilibrium values; shaded cells converge to the closed-loop fixed point within 200 s.

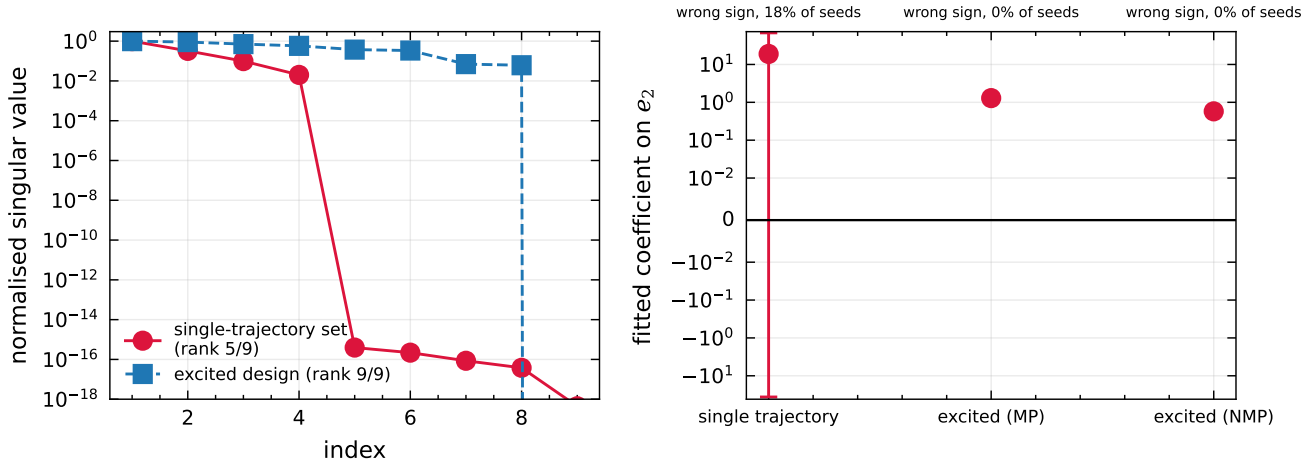

Figure 11: Why the sign flips. Left: normalized singular-value spectrum of the policy regressor. The single-trajectory design collapses to rank 5 of 9; the excited design of Table 2 is full rank. Right: the coefficient on e_2 recovered by an unconstrained fit, over 200 seeds of the 10^{-4} input jitter applied during fitting. On the degenerate design its sign is decided by the seed.

that there are coefficients to identify. None of this is available for an MLP of comparable accuracy.

What it does not provide is accuracy. The MLP imitates the teacher at least as closely at every budget considered, and direct sparse polynomial regression matches or beats the KAN-derived law in the MP regime. The trade-off is auditability and determinism against imitation fidelity.

With a 120 s horizon the LTV-MPC outperforms a gain-scheduled LQR by 7–9% on the most demanding scenario we could construct. When the quantity being distilled is that small, no distillation method can demonstrate much, and the closed-loop ranking between an MLP, a polynomial and a symbolic KAN is within the noise of the experiment (Tables 4–5). We therefore state the boundary of the claim explicitly: what this study establishes is a method validated on a benchmark that cannot by itself justify the method's use. Demonstrating that the method pays for itself requires a plant on which optimization-based control has a large margin over linear control, such as one with active state constraints or short sampling periods.

6.2. Smoothness is approximation error, not a design feature

The MPC's aggressive, near-saturating response is the optimal response to the cost it was given; a low-degree polynomial cannot reproduce it, and the resulting smoothness is a consequence of limited approximation capacity plus a deliberately detuned core. It does reduce the total variation of the command — 7.5 V against the MPC's 16.5 V on the nominal task, 54.9 V against 87.0 V on the staircase —

and reduced actuator activity is genuinely valuable in fluid systems. But the reduction is a factor of two, not an order of magnitude, and it provides slightly slower disturbance rejection (Fig. 7).

The size of this effect depends critically on how the teacher is tuned. Shortening the prediction horizon to three seconds — below the inverse response — raises the MPC's own total variation to 439–818 V, because the myopic controller amplifies estimator noise into command chatter carrying no information about the plant. We therefore report smoothness only against a teacher whose horizon covers the inverse response, where the factor is two. If smooth actuation is the objective, the principled route remains a rate penalty $\|\Delta u\|_R$ in the teacher's cost.

6.3. On function libraries and human control of KAN training

First, sparse identification methods, SINDy (Brunton et al., 2016) among them, were criticized for depending on a predefined function library while KANs were said to avoid the issue. They do not: the symbolic read-out of a KAN draws from exactly such a library, and Section 5.1 shows that its expressiveness is the binding constraint on the final accuracy. The difference is one of stage, not of kind — SINDy commits to the library before fitting, a KAN after.

Second, the learning process can be shaped by constraining the basis functions, by adding terms to the loss, or by making the symbolic selection itself differentiable, and several recent methods do precisely this (Bagrow and Bongard, 2025; Faroughi et al., 2026; Howard et al., 2026). For networks of the size used here none of these is computationally demanding. The post-hoc convex read-out of (12) should be seen as complementary to them: it operates on a network trained by any means, and it is the mechanism by which the physical requirement is guaranteed rather than encouraged.

6.4. Limitations

The teacher is an LTV-MPC on a fourth-order plant with two inputs; the online inference cost of the distilled law is $\mathcal{O}(1)$ in the plant dimension, but the offline symbolic expansion is not, and the number of monomials of degree d in n features grows as $\binom{n+d}{d}$. Scaling beyond roughly ten features will require sparsity to be enforced during training rather than recovered afterwards. The distilled policy has no integral action and therefore retains a steady-state offset under plant–model mismatch that the MPC does not (Fig. 8); an outer integral trim or the online recursive-least-squares adaptation sketched below would address this. State-constraint satisfaction is measured, not guaranteed. Finally, this is a software-in-the-loop study: the cost claims are counts of arithmetic operations, and whether the fixed cost of the explicit law translates into a useful advantage on a given processor is a deployment question that we do not answer here.

7. Conclusion

This paper examined what a symbolic Kolmogorov–Arnold read-out contributes to the distillation of a model predictive controller for a non-minimum-phase plant. Separating the network's approximation error from the symbolic read-out's shows that the second dominates: an off-the-shelf extraction discards most of the fidelity a well-trained KAN had achieved. Re-estimating the coefficients on the KAN-selected support by convex least squares recovers most of it, and imposing elementary negative feedback as a linear inequality inside that same convex problem removes — by construction, and without any manual editing — a failure mode that occurs on up to 31% of the operating box when the read-out is left unconstrained.

Against matched baselines the result is reported as such: across eleven term budgets and two regimes the KAN-selected support is statistically indistinguishable from a support chosen directly by orthogonal matching pursuit over the full monomial dictionary, and a dense MLP is at least as accurate as any symbolic form. The case for the symbolic law is that its arithmetic cost is fixed and state-independent, and that it is an object to which analytic verification can be applied at all: a closed-loop Jacobian certifies local exponential stability at every setpoint tested, and the same closed form makes the shape constraint expressible as an affine inequality. What it does not provide is the constraint satisfaction that motivates MPC in the first place, and we quantify the resulting violations rather than assert their absence.

We also report a limit on what the benchmark can show. Once the teacher's prediction horizon is set correctly, the LTV-MPC's closed-loop advantage over a well-tuned gain-scheduled LQR on this plant is only 7–9%, and every distilled policy lands within a few percent of both. A benchmark on which the optimizer barely beats a linear controller cannot demonstrate the value of approximating that optimizer, whatever the approximator. The methodological contributions above stand on their own, but the case for deploying them belongs on plants where online optimization earns a large margin.

Two directions follow. The first is to move the symbolic restriction from post-processing into training, where recent differentiable-symbolification methods (Bagrow and Bongard, 2025; Faroughi et al., 2026; Howard et al., 2026) make the structure selection itself learnable, removing the fidelity cliff documented here. The second is to restore adaptability: because the symbolic basis is fixed offline and only the coefficients need to change, online adaptation reduces to recursive least squares on a linear-in-parameters model — a convex, deterministic, verifiable update implementable with standard embedded digital signal processor (DSP) routines, in contrast to online backpropagation.

Data and Code Availability

All datasets, training and analysis scripts, and the figure- and table-generation code that produce every number in this paper are available at `https://github.com/cl0sure6/`

QuadTank_KAN_and_LTV-MPC under the Apache 2.0 license. Running the numbered stage scripts in order regenerates the datasets, the trained models, all result files, all figures and all tables.

Acknowledgments

The author thanks the faculty of the School of Information Technology and Engineering at Kazakh-British Technical University for guidance on the theoretical framework and research direction, and the open-source contributors of the PyKAN, CVXPY, OSQP, SciPy (Virtanen et al., 2020) and scikit-learn (Pedregosa et al., 2011) ecosystems, whose software made this comparison possible.

CRediT authorship contribution statement

Adilkhan Salkimbayev: Conceptualization, Methodology, Software, Validation, Formal analysis, Writing - original draft, Writing - review and editing.

References

Alvarado, I., Limon, D., De La Peña, D.M., Maestre, J.M., Ridao, M., Scheu, H., Marquardt, W., Negenborn, R., De Schutter, B., Valencia, F., et al., 2011. A comparative analysis of distributed MPC techniques applied to the HD-MPC four-tank benchmark. Journal of Process Control 21, 800–815. doi:10.1016/j.jprocont.2011.03.003.

Amodei, D., Olah, C., Steinhardt, J., Christiano, P., Schulman, J., Mané, D., 2016. Concrete problems in AI safety. arXiv preprint arXiv:1606.06565. doi:10.48550/arXiv.1606.06565.

Arnold, V.I., 1957. On functions of three variables, in: Doklady Akademii Nauk, Russian Academy of Sciences. pp. 679–681.

Arnström, D., Axehill, D., 2022. Active-set based methods and warm starting in embedded quadratic programming. IEEE Transactions on Automatic Control 67, 2758–2770. doi:10.1109/TAC.2021.3090749.

Bagrow, J., Bongard, J., 2025. Softly symbolifying Kolmogorov–Arnold networks. arXiv preprint arXiv:2512.07875 doi:10.48550/arXiv.2512.07875.

Bemporad, A., Morari, M., Dua, V., Pistikopoulos, E.N., 2002. The explicit linear quadratic regulator for constrained systems. Automatica 38, 3–20. doi:10.1016/S0005-1098(01)00174-1.

Borrelli, F., Bemporad, A., Morari, M., 2017. Predictive control for linear and hybrid systems. Cambridge University Press. doi:10.1017/9781139061759.

Boyd, S., Vandenberghe, L., 2004. Convex optimization. Cambridge University Press.

Brunton, S.L., Proctor, J.L., Kutz, J.N., 2016. Sparse identification of nonlinear dynamics with control (SINDYc). IFAC-PapersOnLine 49, 710–715. doi:10.1016/j.ifacol.2016.10.249.

Chen, S., Saulnier, K., Atanasov, N., Lee, D.D., Kumar, V., Pappas, G.J., Morari, M., 2018. Approximating explicit model predictive control using constrained neural networks, in: 2018 Annual American Control Conference (ACC), IEEE. pp. 1520–1527. doi:10.23919/ACC.2018.8431275.

Diamond, S., Boyd, S., 2016. CVXPY: A Python-embedded modeling language for convex optimization. Journal of Machine Learning Research 17, 1–5.

Faroughi, S.A., Mostajeran, F., Arzani, A., Faroughi, S., 2026. Symbolic-KAN: Kolmogorov–Arnold networks with discrete symbolic structure for interpretable learning. arXiv preprint arXiv:2603.23854 doi:10.48550/arXiv.2603.23854.

Gupta, M., Cotter, A., Pfeifer, J., Voevodski, K., Canini, K., Mangylov, A., Moczydlowski, W., Van Esbroeck, A., 2016. Monotonic calibrated interpolated look-up tables. Journal of Machine Learning Research 17, 1–47.

Hertneck, M., Köhler, J., Trimpe, S., Allgöwer, F., 2018. Learning an approximate model predictive controller with guarantees. IEEE Control Systems Letters 2, 543–548. doi:10.1109/LCSYS.2018.2843682.

Howard, A.A., Zolman, N., Jacob, B., Brunton, S.L., Stinis, P., 2026. SINDy-KANs: Sparse identification of non-linear dynamics through Kolmogorov–Arnold networks. arXiv preprint arXiv:2603.18548 doi:10.48550/arXiv.2603.18548.

Johansson, K.H., 2000. The Quadruple-Tank Process: A multivariable laboratory process with an adjustable zero. IEEE Transactions on Control Systems Technology 8, 456–465.

Kalman, R.E., 1960. A new approach to linear filtering and prediction problems. Transactions of the ASME–Journal of Basic Engineering 82, 35–45.

Karg, B., Lucia, S., 2020. Efficient representation and approximation of model predictive control laws via deep learning. IEEE Transactions on Cybernetics 50, 3866–3878. doi:10.1109/TCYB.2020.2999556.

Khalil, H.K., 2002. Nonlinear Systems. 3 ed., Prentice Hall.

Kolmogorov, A.N., 1957. On the representations of continuous functions of many variables by superposition of continuous functions of one variable and addition, in: Dokl. Akad. Nauk USSR, pp. 953–956.

Liu, Z., Wang, Y., Vaidya, S., Ruehle, F., Halverson, J., Soljačić, M., Hou, T.Y., Tegmark, M., 2024. KAN: Kolmogorov-Arnold Networks. arXiv preprint arXiv:2404.19756. doi:10.48550/arXiv.2404.19756.

Lu, L., Jin, P., Pang, G., Zhang, Z., Karniadakis, G.E., 2021. Learning nonlinear operators via DeepONet based on the universal approximation theorem of operators. Nature Machine Intelligence 3, 218–229. doi:10.1038/s42256-021-00302-5.

Lucia, S., Navarro, D., Lucia, Ó., Zometa, P., Findeisen, R., 2021. Deep learning-based model predictive control for resonant power converters. IEEE Transactions on Industrial Informatics 17, 409–420. doi:10.1109/TII.2020.2969729.

Mayne, D.Q., Rawlings, J.B., Rao, C.V., Scokaert, P.O., 2000. Constrained model predictive control: Stability and optimality. Automatica 36, 789–814. doi:10.1016/S0005-1098(99)00214-9.

Parisini, T., Zoppoli, R., 1995. A receding-horizon regulator for nonlinear systems and a neural approximation. Automatica 31, 1443–1451. doi:10.1016/0005-1098(95)00044-W.

Pati, Y.C., Rezaiifar, R., Krishnaprasad, P.S., 1993. Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition. Proceedings of 27th Asilomar Conference on Signals, Systems and Computers , 40–44doi:10.1109/ACSSC.1993.342465.

Pedregosa, F., Varoquaux, G., Gramfort, A., et al., 2011. Scikit-learn: Machine learning in Python. Journal of Machine Learning Research 12, 2825–2830.

Rudin, C., 2019. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nature Machine Intelligence 1, 206–215. doi:10.1038/s42256-019-0048-x.

Skogestad, S., Postlethwaite, I., 2005. Multivariable feedback control: analysis and design. John Wiley & Sons.

Stellato, B., Banjac, G., Goulart, P., Bemporad, A., Boyd, S., 2020. OSQP: An operator splitting solver for quadratic programs. Mathematical Programming Computation 12, 637–672. doi:10.1007/s12532-020-00179-2.

Virtanen, P., Gommers, R., Oliphant, T.E., et al., 2020. SciPy 1.0: fundamental algorithms for scientific computing in Python. Nature Methods 17, 261–272. doi:10.1038/s41592-019-0686-2.

Wright, S., Nocedal, J., et al., 1999. Numerical optimization. Springer Science 35, 7.