

Supplementary Material for “Neural Belief Corrections for MPC: Plan-Consistent Scalar Budgets versus Myopic Directional Constraints”

Thanana Nuchkrua, Sudchai Boonto, Woraphrut Kornmaneesang, and Xiaoqi Liu

This document supports the eight-page Technical Note. Equation, theorem and table numbers without the prefix “S” refer to the main manuscript. Supplementary Video S1 accompanies this document (Secs. S5 and S7). Code and data are available at <https://thanana.github.io/NeuralBeliefMPC.html> (Sec. S7).

Contents

S1 Pendulum protocol	1
S2 Single-IP results and budget traces	2
S3 Controller-first frozen-context study	2
S4 Pooled frozen-context study (Sec. V-C)	3
S5 Closed-loop deployment of the postprocessors	5
S6 Sensitivity and conditional invariance	6
S7 Implementation and reproducibility	7

S1 Pendulum protocol

We use the DeepMind Control Suite cart–pole XML and its official two-pole generator (Tassa et al., 2018), pinned at repository commit 773d5b5d94d8. These specify model assets; the noisy partial-sensing regulation task is ours, not the suite’s reward/reset benchmark. The cart is 1 kg; each pole is a 0.1 kg, 1 m capsule of radius 0.045 m. MuJoCo (Todorov et al., 2012) infers capsule inertias, exported with the compiled masses. Upstream damping, slider range, disabled contacts and actuator settings are retained. RK4 takes two 10 ms steps per 20 ms control interval, and physical force u maps to motor control $u/10$.

States, sensors and profiles. States are $(p, \theta, \dot{p}, \dot{\theta})$ for Single IP and $(p, \theta_1, \theta_{2,r}, \dot{p}, \dot{\theta}_1, \dot{\theta}_{2,r})$ for Double IP, with $\theta_{2,a} = \theta_1 + \theta_{2,r}$. Sensors measure position and absolute angles with independent deviations 0.002 m and 0.2° ; velocities are unobserved. Initial positions/relative angles are Gaussian about $(0.02 \text{ m}, 2^\circ)$ or $(0.02 \text{ m}, 2^\circ, -1^\circ)$ with deviations 0.002 m/ 0.2° and zero velocities; within each seed, methods share the initial draw and noise. Five 20 s profiles are nominal, last-pole mass +50% at 8 s, lower-angle bias ramped to 3° over 8–9 s, all angles missing over 12–12.4 s, and combined. Observers receive masks, not truth or intervention times. Each native simulator step checks $|p| \leq 0.8 \text{ m}$ and each absolute angle within 12° ; applied force must stay within 10 N. An episode succeeds only if it completes without any violation; failures are retained.

Controller. Both plants use horizon-50 certainty-equivalent nominal linear MPC with input weight 0.2 and $Q = T_a^T \text{diag}(10, 100\mathbf{1}_{n_p}, 1, 5\mathbf{1}_{n_p})T_a$, where $n_p \in \{1, 2\}$ is the number of poles and T_a maps relative to absolute coordinates. Planning limits are 0.75 m/ 11° /9.5 N. Terminal sets have 138/312 facets; floating LP support checks give maxima below $1 + 6 \times 10^{-15}$. Invalid plans use shifted feasible plans or saturated LQR recovery, without a safety certificate.

Networks, selection and adaptation. Transition/encoder networks have two width-32 ReLU layers; observation width is 16. Each plant is trained on 64 six-second truth-supervised trajectories (seeds 6001–6064) with 16 validation trajectories (8001–8016). Closed-loop validation uses seeds 8101–8102, nominal/combined profiles and radii $0, 10^{-4}, 10^{-3}, 10^{-2}$; a prespecified lexicographic rule minimizes incomplete runs, violating episodes, then angle-estimation RMSE over the positive radii, giving $10^{-4}/10^{-2}$ for Single/Double IP; for Single IP radius 0 has lower validation RMSE (0.338° versus 0.425°) with equal completion and violation counts. Adaptation uses $\varepsilon_v = 0.002$, $R_v = 0.25$, $B_v = 2$, step 0.02, $\mu = 0.001$, $\beta_J = 10^{-4}$, gradient cap 5, eight backtracking attempts every ten ticks and at most 16 cached pairs. MM variants moment-match an IMM (Blom and Bar-Shalom, 1988) with last-pole mass ratios $\{0.7, 1, 1.5, 2\}$, prior weights $(0.05, 0.85, 0.08, 0.02)$ and mixing $0.99I + 0.01\mathbf{1}\mathbf{1}^T/4$; bias is estimated and the learned transition residual is zero. MM-CF adapts only the encoder; its score uses the fixed trained neural transition/observation surrogate. MM dropout clears replay. MM-CF binding uses coupling budgets $1.1197533 \times 10^{-6}/2.6113003 \times 10^{-5}$, half the median positive accepted coupling at calibration seed 8103, fixed before testing.

Tests and metrics. Seeds 96301–96305 give 300 episodes per plant across twelve variants (five independent seed clusters). Angle-estimation RMSE is $\sqrt{\text{mean}_t \text{mean}_j (\hat{\theta}_{j,a} - \theta_{j,a})^2}$ per episode, in degrees, averaged over seeds; tracking RMSE uses true angles relative to upright. Upward-rounded 100-digit bounds give $\ell_E \leq 8046.614878/10267.124261$. Maximum finite-budget excess over 300,000/298,754 IP rows is $4.04 \times 10^{-15}/1.02 \times 10^{-13}$.

S2 Single-IP results and budget traces

Table S1 gives the Single-IP angle-estimation errors and zero-violation counts summarized in Sec. V-B of the manuscript.

Table S1: Single IP: angle-estimation RMSE in degrees (five-seed means); ZV: complete zero-violation episodes over five profiles; EKF: extended Kalman filter.

Observer	Nominal	Combined	ZV
Oracle	0.000	0.000	25/25
EKF	0.112	2.659	15/25
Bias EKF	0.135	0.438	25/25
Physics IMM	0.135	0.460	25/25
Neural Gaussian	0.152	0.680	25/25
Neural frozen	0.156	0.746	25/25
Neural CF encoder	0.142	0.709	25/25
Neural CF joint	127.772	315.804	0/25
Joint, no projection	1360.330	2588.906	0/25
MM frozen	0.139	0.445	25/25
MM-CF	0.140	0.448	25/25
MM-CF binding	0.139	0.447	25/25

Figure S1 shows the Single-IP combined profile for seed 96301 (MM-CF binding and comparison observers). Panel (b) plots the projected correction relative to its radius, the accepted parameter step relative to $\varepsilon_v = 0.002$, and the accepted binding coupling relative to its budget; all remain at or below one, as Theorems 1–2 require. Panel (b) appears in single-column form as Fig. 2 of the manuscript.

S3 Controller-first frozen-context study

We reuse the 98 combined-seed-96301 contexts per IP; 98/58 Single/Double references lie in the cell. Set $\bar{u}_c \in \{0.001, 0.01, 0.1\}$ N and $\bar{u}_a = 0.1\bar{u}_c$. *Correction-only* uses $\delta_E = \min\{\delta_{\text{nominal}}, 0.99\delta_C, \bar{u}_c/g_M\}$ and $\varepsilon_B = 0.002$; *simultaneous* also caps ε_B at $\bar{u}_a/g_M$; *fixed* uses $\delta_{\text{nominal}} = 10^{-4}/10^{-2}$ and $\varepsilon_B = 0.002$. Identical encoder-gradient proposals use archived frozen caches, norm cap 0.002, scale 0.02 and eight backtracking trials.

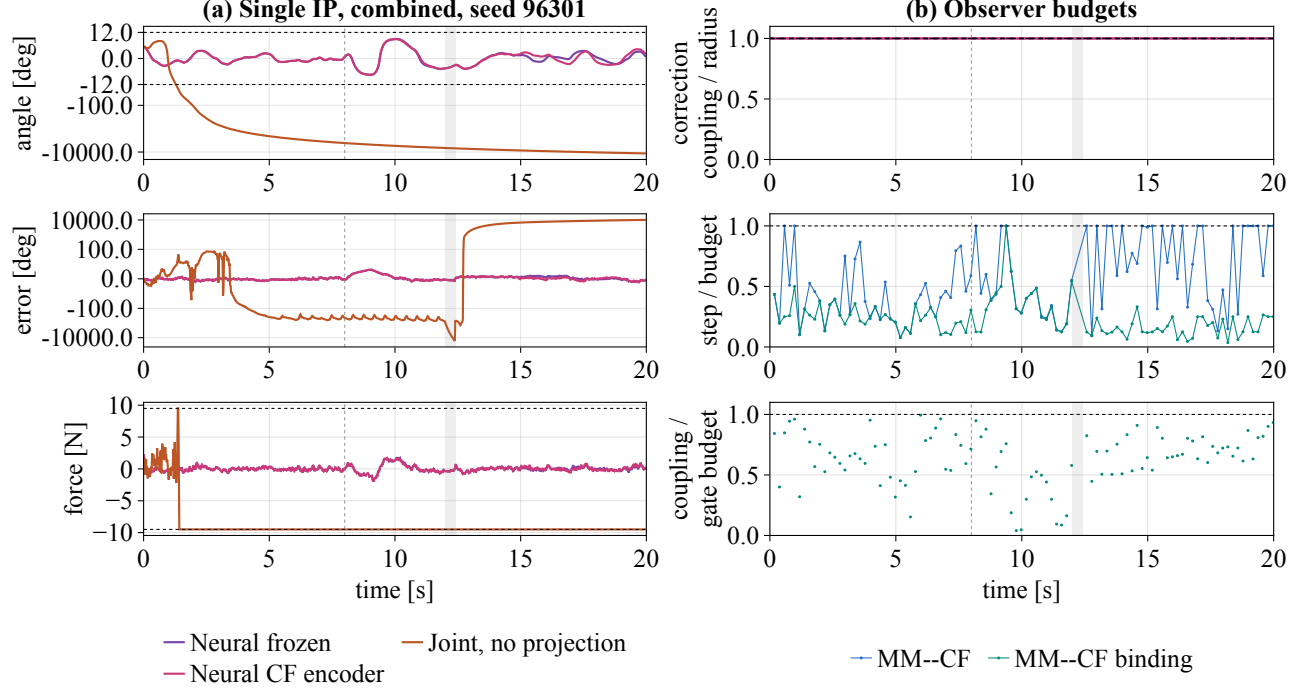


Figure S1: Single-IP combined profile, seed 96301. (a) True angle, estimation error and force; state/error axes are signed-log with physical-value ticks. (b) Projected correction/radius, accepted parameter step/0.002, and accepted binding coupling/budget. Marker: mass/bias onset; shading: dropout. Dashed lines mark physical angle, planning force or observer budgets.

All in-cell cases accept updates without cell exits. Correction-only already meets both input limits; simultaneous rejects two additional tight-tolerance Single-IP candidates before accepting smaller steps. Independent reconstruction checks 1,568 candidate scores and 1,764 decisions. Cold-QP checks at archived indices 001, 011, ..., 091 (ten Single, six Double contexts) give 104 distinct solves and 432 point comparisons with cell-law discrepancy below 3.85×10^{-14} N. Median squared factor fractions of reference coupling are 0.609/0.988. These dependent contexts establish neither an adaptation/covariance benefit nor nonlinear safety.

S4 Pooled frozen-context study (Sec. V-C)

Context generation. The MM-CF binding observer is replayed in all 25 profile/seed test episodes per plant (`revision/julia/replay_seed.jl`), saving every adaptation context exactly as the original `run_guard_replay.jl` does for combined/96301. All 50 replayed episode logs equal the archived test logs in every logged column (maximum absolute difference 0), and the combined/96301 frozen records and raw proposals equal the archived ones bit for bit. `run_study.jl` and `extract_raw.jl` then produce the encoder-gradient proposal, the first cached-score-passing candidate and the raw mean offsets of the old and new parameters for each context.

Problem data. As in the archived study, the incumbent d_0 is the old raw mean offset scaled onto the scalar-budget radius $\delta_s = \min\{\delta_E, 0.99\delta_C(\hat{x}_r), \bar{u}_c/g\}$, the proposal q is the new raw mean offset, $\eta = 0.99$, $\varepsilon_B = \varepsilon_S = 0.002$ and $\bar{u}_a = 0.1\bar{u}_c$. Cell facets that can cut the radius- δ_E ball are retained. The predicted-input map G_U is rebuilt from the stored A, B, K, P, Q and $R = 0.2$ by the pre-stabilized condensing of `mpc_core.jl`; its first row equals $-K_M S_x$, up to the sign convention of the stored gain, within 2×10^{-12} . Row norms peak at $k = 0$ for both plants ($\max_k \|G_k\| = 21.10/79.95$), and $\text{rank}(G_{N_h}) = n$ first at $N_h = 4/6$ (smallest singular values $2.4 \times 10^{-5}/8.0 \times 10^{-8}$, so G_{N_h} is numerically near-singular there). For Double IP, $\gamma_1 = 3.92$, $\max_k \gamma_k = 9.77$ at $k = 5$, $\gamma_A = 3.92$, $\|G_A\| = 86.50$ and $|K_M B| = 0.515$; for Single IP, $\gamma_1 = 0.57$, $\max_k \gamma_k = 2.31$ at $k = 10$.

Table S2: Input-limit exceedances (correction/update) in the controller-first study: 98 Single/58 Double in-cell contexts per tolerance. C: correction-only; S: simultaneous.

Plant	$\bar{u}_c$ [N]	Fixed	C	S
Single	0.001	37/7	0/0	0/0
	0.01	0/0	0/0	0/0
	0.1	0/0	0/0	0/0
Double	0.001	56/57	0/0	0/0
	0.01	45/41	0/0	0/0
	0.1	0/0	0/0	0/0

Solvers and numerical checks. Projections use Clarabel (scaled variables, tolerances 10^{-12}) with constraint violations below 5×10^{-11} . Restricted to combined/96301, the pooled pipeline reproduces the previous $N_h = 1/N_h = N$ results within 1.9×10^{-9} , and $N_h = 1$ reproduces the archived Dykstra distances within 10^{-6} . An independent check on 564 random in-cell contexts (6768 projections) re-solves each problem more tightly, recovers multipliers by NNLS on the near-active set and evaluates an analytical Lagrangian-dual expression to obtain numerical a posteriori estimates of the projection-distance gaps. These estimates are below 3.3×10^{-11} , except for three near-feasible proposals whose directional distances are below 1.1×10^{-5} ; they are numerical checks rather than outward-rounded certificates.

Results by profile. Table S3 breaks down $\bar{u}_c = 0.001$ N by profile; the pooled values for all tolerances are in Table III of the manuscript. Unavailable Double-IP contexts (reference outside the cell) occur mostly after dropout in episodes that later violate limits. The raw proposal is closer to the true state than the reference in only 35–48% of contexts. This observation does not establish why directional projection changes accuracy; the direct comparison shows that directional means are farther from truth than scalar means in 59–64% of Double-IP contexts.

Table S3: Pooled frozen-context study by profile at $\bar{u}_c = 0.001$ N (25 episodes per plant). Red.: median distortion reduction; Plan 1: largest predicted-input change with $N_h = 1$ (rounded up); Worse: share of contexts where the directional mean is farther from truth than the scalar mean; Raw<ref: share where the raw proposal is closer to truth than the reference.

Plant	Profile	In-cell	Red. 1 [%]	Red. N [%]	Plan 1 [$\bar{u}_c$]	Worse 1 [%]	Worse N [%]	Raw<ref [%]
Single	nominal	500/500	1.2	1.2	1.00	48	47	46
	mass	500/500	1.1	1.1	1.00	46	46	47
	bias	500/500	1.1	1.0	1.00	45	45	45
	dropout	490/490	1.2	1.2	1.00	49	49	45
	combined	490/490	1.1	1.1	1.00	42	42	48
Double	nominal	497/500	27.4	9.6	18.68	59	58	40
	mass	495/500	26.1	10.0	18.68	59	59	40
	bias	496/500	28.3	9.9	18.68	63	60	35
	dropout	336/490	27.9	9.4	18.68	57	56	42
	combined	329/490	26.1	9.2	18.68	59	62	39

Across all tolerances, Double-IP directional means are farther from truth in 59–64% of contexts, and by episode median in 24–25 of 25 episodes. Single-IP differences are below 6.4×10^{-5} in normalized units; the directional mean is slightly closer in 54% of contexts at 0.001 N and identical at larger tolerances.

Timing. On the archived contexts, Python Dykstra $N_h = 1$ projections alone (95th percentiles 17.48–20.44 ms) exceed the 20 ms control interval in 4, 4 and 1 of 58 Double contexts.

S5 Closed-loop deployment of the postprocessors

Setup. The MM-CF binding observer of the 600-episode study is run unchanged (adaptation, cached score, coupling gate, IMM bank), except that at every fully observed tick whose reference lies in the cell the deployed physical mean offset is replaced by its projection onto the correction constraints of Sec. III-E at the reference covariance: the scalar budget $\|d\| \leq \delta_s$, or $\{\|d\| \leq \delta_E, \bar{M}d \leq 0.99(b - M\hat{x}_r), \|G_{N_h}d\|_\infty \leq \bar{u}_c\}$ with $N_h = 1$ or $N_h = N = 50$. No incumbent or update constraint is used in closed loop; nuisance coordinates keep the original correction. Out-of-cell or masked ticks keep the original map. Directional runs leave the cell more often: median Double-IP postprocessed ticks per episode (of about 1000) are 997/987 for the scalar budget, 793/676 for $N_h = 1$ and 466/518 for $N_h = N$ ($\bar{u}_c = 0.001/0.01$ N); Single-IP references stay in the cell at every unmasked tick under all three (`post_ticks`, `fallback_ticks` in `revision/analysis/results/closed_loop_results.csv`). We solve the polyhedral least-distance problem using Lawson–Hanson NNLS and enforce the ball constraint by numerical bisection on its multiplier; a check against Clarabel on 672 frozen contexts gave feasibility 2.7×10^{-15} and distances never larger than Clarabel’s. All 25 test episodes per plant (five profiles, seeds 96301–96305) are run for $\bar{u}_c = 0.001$ and 0.01 N and each policy (300 runs, `revision/julia/closed_loop/`).

Table S4: Closed-loop zero-violation episodes by profile (five seeds each), Double IP.

$\bar{u}_c$ [N]	Policy	Nominal	Mass	Bias	Dropout	Combined	Total
0.001	Original radial	5/5	5/5	5/5	1/5	1/5	17/25
	Scalar	5/5	5/5	5/5	0/5	1/5	16/25
	$N_h = 1$	4/5	1/5	3/5	1/5	0/5	9/25
	$N_h = N$	0/5	0/5	0/5	1/5	0/5	1/25
0.01	Original radial	5/5	5/5	5/5	1/5	1/5	17/25
	Scalar	5/5	5/5	5/5	1/5	1/5	17/25
	$N_h = 1$	3/5	1/5	3/5	1/5	0/5	8/25
	$N_h = N$	1/5	0/5	1/5	0/5	0/5	2/25

Table S5: Median normalized estimation RMSE per state over the first 8 s (or until the first violation) of the five nominal Double-IP episodes, $\bar{u}_c = 0.01$ N.

Deployed mean	p	θ_1	$\theta_{2,r}$	$\dot{p}$	$\dot{\theta}_1$	$\dot{\theta}_{2,r}$
Original radial	0.0016	0.0321	0.0387	0.0138	0.0279	0.0556
Scalar	0.0012	0.0440	0.0504	0.0043	0.0389	0.0743
$N_h = 1$	0.0045	0.0530	0.0562	0.0474	0.0416	0.0817
$N_h = N$	0.0061	0.0463	0.0535	0.0534	0.0404	0.0919

Table S6: Closed-loop Single IP, 25 episodes per entry: zero-violation episodes, median tracking and angle-estimation RMSE (deg), and median per-episode median of $\max_k |G_k d|/\bar{u}_c$.

$\bar{u}_c$ [N]	Deployed mean	Zero-viol.	Tracking	Estimation	Plan
0.001	Original radial	25/25	0.343	0.145	–
	Scalar	25/25	0.343	0.148	0.82
	$N_h = 1$	25/25	0.341	0.151	1.00
	$N_h = N$	25/25	0.341	0.151	1.00
0.01	Original radial	25/25	0.343	0.145	–
	Scalar	25/25	0.343	0.150	0.16
	$N_h = 1$	25/25	0.343	0.150	0.16
	$N_h = N$	25/25	0.343	0.150	0.16

In Single IP all policies keep every episode safe and track as the original map does, consistent with the negligible frozen-context differences (Table S3).

Video. Part 2 of Supplementary Video S1 animates the mass-change episode with seed 96301 at $\bar{u}_c = 0.01$ N for all four deployed means: the original map and the scalar budget stay within limits, whereas $N_h = 1$ and $N_h = N$ violate them first at 10.20 s and 11.26 s after the mass change at 8 s.

Interpretation (Double IP). The scalar budget changes the zero-violation outcome of only 3–4 of 25 episodes relative to the original map (it keeps 15 of the 17 originally zero-violation episodes at both tolerances and adds 1–2 others), and its tracking cost comes from its small radius $\delta_s \leq \bar{u}_c/g$. The directional sets keep larger corrections, but the corrections they keep are those the current or predicted inputs barely see (Theorem 3, Proposition 2). Table S5 shows that cart position and velocity are then estimated much worse than under either the scalar budget or the original map, and these episodes later violate the cart or angle limits; the link between the admitted directions and these states is an interpretation, not a proof. With $N_h = N$, at every tick where the postprocessor is applied, all predicted-input correction changes satisfy the tolerance, yet only 1–2 of 25 episodes stay safe: plan consistency of the belief correction does not imply closed-loop safety. These runs are frozen-hyperparameter tests of one observer; they do not certify the scalar budget either.

S6 Sensitivity and conditional invariance

Fixed-context sensitivity. The following result bounds the same-context coupling of Theorem 2 by a Lipschitz constant of a ReLU encoder.

Proposition S1 (fixed-context sensitivity). *Here $\phi_\star$ is the encoder block of $v_\star$ and the covariance bound $pI_{n_z} \preceq P_r \preceq \bar{p}I_{n_z}$ of the main text holds. Let the concatenated encoder output $o_\phi = (o_\phi^m, o_\phi^\sigma)$ be a two-hidden-layer ReLU network $h_1 = \max\{W_1 a + \beta_1, 0\}$, $h_2 = \max\{W_2 h_1 + \beta_2, 0\}$, $o_\phi = W_3 h_2 + \beta_3$ with input norm $\|a\| \leq a_{\text{in}}$, and let ϕ, ϕ' lie in $\{\phi : \|\phi - \phi_\star\| \leq R_v\}$, where the parameter vector concatenates all weight entries and biases. For $i = 1, 2, 3$ put $\bar{W}_i = \|W_{i,\star}\|_2 + R_v$, $\beta_i^+ = \|\beta_{i,\star}\| + R_v$, $\bar{h}_1 = \bar{W}_1 a_{\text{in}} + \beta_1^+$, $\bar{h}_2 = \bar{W}_2 \bar{h}_1 + \beta_2^+$ and*

$$K_\phi = \bar{h}_2 + 1 + \bar{W}_3 [\bar{h}_1 + 1 + \bar{W}_2(a_{\text{in}} + 1)]. \quad (\text{S1})$$

Then, at the same reference and radius,

$$\rho_t(v', v) \leq \min\{2\delta_E, \ell_E \|\phi' - \phi\|\}, \quad \ell_E = K_\phi \max\left\{\max_{i \leq n} a_{E,i}, \kappa_0 \sqrt{\bar{p}}\right\}, \quad (\text{S2})$$

with $\kappa_0 = 2(\sqrt{2} - 1)e^{\sqrt{2}-1}$. Hence the gate (6) is redundant whenever $\varepsilon_B \geq \min\{2\delta_E, \ell_E \varepsilon_v\}$.

Proof. Let $\epsilon = \|\phi' - \phi\|$. Since the parameter vector concatenates all entries, every weight block obeys $\|W'_i - W_i\|_2 \leq \|W'_i - W_i\|_F \leq \epsilon$ and every bias $\|\beta'_i - \beta_i\| \leq \epsilon$; similarly $\|W_i\|_2, \|W'_i\|_2 \leq \bar{W}_i$ and $\|\beta_i\|, \|\beta'_i\| \leq \beta_i^+$, so $\|h_1\|, \|h'_1\| \leq \bar{h}_1$ and $\|h_2\|, \|h'_2\| \leq \bar{h}_2$. As the ReLU is 1-Lipschitz,

$$\begin{aligned} \|h'_1 - h_1\| &\leq \|W'_1 - W_1\|_2 \|a\| + \|\beta'_1 - \beta_1\| \leq (a_{\text{in}} + 1)\epsilon, \\ \|h'_2 - h_2\| &\leq \|W'_2 - W_2\|_2 \|h'_1\| + \|W_2\|_2 \|h'_1 - h_1\| + \|\beta'_2 - \beta_2\| \\ &\leq (\bar{h}_1 + 1)\epsilon + \bar{W}_2(a_{\text{in}} + 1)\epsilon, \\ \|o_{\phi'} - o_\phi\| &\leq \|W'_3 - W_3\|_2 \|h'_2\| + \|W_3\|_2 \|h'_2 - h_2\| + \|\beta'_3 - \beta_3\| \leq K_\phi \epsilon. \end{aligned}$$

At a common reference, (3) gives the raw physical offset differences $\Delta m_x = a_{E,x} \odot (\tanh(o_x^m) - \tanh(o_x^m))$ and $\Delta L_x = (D'_x - D_x)L_{r,x}$. Since $\tanh$ is 1-Lipschitz, $\|\Delta m_x\| \leq \max_{i \leq n} a_{E,i} \|\Delta o_x^m\|$. With $s_i = \tanh(\cdot)$, $\frac{d}{ds} e^{\tanh s} = e^{s_i}(1 - s_i^2)$, whose maximum over $s_i \in (-1, 1)$ is attained at $s_i = \sqrt{2} - 1$ and equals κ_0 ; thus each diagonal entry of $D'_x - D_x$ is at most $\kappa_0 |\Delta o_i^\sigma|$ in modulus. Row i of $L_{r,x}$ has squared norm $(P_r)_{ii} \leq \bar{p}$, so $\|\Delta L_x\|_F^2 \leq \kappa_0^2 \bar{p} \|\Delta o_x^\sigma\|^2$. Therefore

$$(\|\Delta m_x\|^2 + \|\Delta L_x\|_F^2)^{1/2} \leq \max\left\{\max_i a_{E,i}, \kappa_0 \sqrt{\bar{p}}\right\} \|o_{\phi'} - o_\phi\| \leq \ell_E \epsilon.$$

The deployed offsets are the images of the raw offsets under the nonexpansive projection of Theorem 1(c), so $\rho_t(v', v) \leq \ell_E \epsilon$; Theorem 2(c) gives $\rho_t \leq 2\delta_E$. Transition parameters do not affect the current correction. If $\varepsilon_B \geq \min\{2\delta_E, \ell_E \varepsilon_v\}$, every candidate with $\|\phi' - \phi\| \leq \varepsilon_v$ already satisfies the coupling inequality in (6). ■

Upward-rounded 100-digit evaluations give $\ell_E \leq 8046.614878$ (Single IP) and $\ell_E \leq 10267.124261$ (Double IP) for the trained anchors ($R_v = 0.25$, $\bar{p} = 4$); at the tested step budget, $2\delta_E$ is the tighter bound.

Conditional local invariance. The budgets also give a forward-invariant set when a true-state error envelope and a contraction of the nonlinear loop are assumed in addition.

Proposition S2 (conditional local invariance). *Use the same law π at every tick. Let $P_V \succ 0$, $\|x\|_{P_V} = (x^T P_V x)^{1/2}$, $R_x > 0$ and $\mathcal{X} = \{x : \|x\|_{P_V} \leq R_x\}$. Suppose that, for every $x \in \mathcal{X}$, all inputs u, u' generated by π on $\mathcal{X} + \mathbb{B}(\bar{e})$ and all $\|w\| \leq \bar{w}$,*

$$\begin{aligned} \|f^*(x, \pi(x), 0)\|_{P_V} &\leq \lambda_c \|x\|_{P_V}, \quad 0 \leq \lambda_c < 1, \\ \|f^*(x, u, w) - f^*(x, u', 0)\|_{P_V} &\leq \ell_u \|u - u'\| + \ell_w \|w\|, \end{aligned} \quad (\text{S3})$$

where $\mathbb{B}(\bar{e})$ is the closed Euclidean ball of radius $\bar{e} = \bar{e}_r + \|S_x\|_2 \bar{\delta}_E$. Assume $\|\hat{x}_{r,t} - x_t\| \leq \bar{e}_r$, $\delta_{E,t} \leq \bar{\delta}_E$ and $\|w_t\| \leq \bar{w}$ whenever $x_t \in \mathcal{X}$, and let π be ℓ_π -Lipschitz on $\mathcal{X} + \mathbb{B}(\bar{e})$. Put $\Delta_x = \ell_u \ell_\pi \bar{e} + \ell_w \bar{w}$. If $\Delta_x \leq (1 - \lambda_c)R_x$, then for all $x_0 \in \mathcal{X}$ and $t \geq 0$,

$$\|x_t\|_{P_V} \leq \lambda_c^t \|x_0\|_{P_V} + \frac{1 - \lambda_c^t}{1 - \lambda_c} \Delta_x; \quad (\text{S4})$$

in particular, $\mathcal{X}$ is forward invariant.

Proof. Let $x_t \in \mathcal{X}$. By Theorem 1(a), $\|\hat{x}_{g,t} - \hat{x}_{r,t}\| \leq \|S_x\|_2 \delta_{E,t}$, so $\|\hat{x}_{g,t} - x_t\| \leq \bar{e}_r + \|S_x\|_2 \bar{\delta}_E = \bar{e}$, and both x_t and $\hat{x}_{g,t}$ lie in the domain of the Lipschitz bound. With $u_t = \pi(\hat{x}_{g,t})$, (S3) and the triangle inequality give

$$\begin{aligned} \|x_{t+1}\|_{P_V} &\leq \|f^*(x_t, \pi(x_t), 0)\|_{P_V} + \|f^*(x_t, u_t, w_t) - f^*(x_t, \pi(x_t), 0)\|_{P_V} \\ &\leq \lambda_c \|x_t\|_{P_V} + \ell_u \ell_\pi \|\hat{x}_{g,t} - x_t\| + \ell_w \|w_t\| \leq \lambda_c \|x_t\|_{P_V} + \Delta_x. \end{aligned}$$

If $\|x_t\|_{P_V} \leq R_x$, then $\|x_{t+1}\|_{P_V} \leq \lambda_c R_x + (1 - \lambda_c)R_x = R_x$, so $x_{t+1} \in \mathcal{X}$ and all hypotheses hold again. By induction from $x_0 \in \mathcal{X}$, $x_t \in \mathcal{X}$ for all t , and iterating the scalar recursion gives (S4) through the geometric sum. ■

These hypotheses are not verified for the benchmark: uniform nonlinear contraction, true-state reference-error envelopes and controller-branch coverage remain open. For the Double-IP horizon-50 ideal QP, 220-digit outward-rounded checks validate $P_V \succ 0$, a nominal (linear) contraction factor of at most 0.999764 and $\{x : \|x\|_{P_V} \leq 6.718\} \subset \mathcal{C}$; they certify the stored finite-decimal nominal problem only.

S7 Implementation and reproducibility

Code and data. The archive `NeuralBeliefMPC_code.zip` at <https://thanana.github.io/NeuralBeliefMPC.html> contains the pinned per-plant Julia environments (`Project.toml/Manifest.toml`, Julia 1.12.1), trained text checkpoints, MPC data and certificates, the packed frozen contexts of the 50 replays, every pooled context and closed-loop run, and reproduction commands. Raw per-tick records of the 600 IP episodes, the archived frozen studies and the interval and guard audits are available from the authors on request. The IP runner skips existing episodes, so fresh execution needs a clean output folder; timing-dependent solver branches may differ in principle, although all 50 replays in Sec. S4 were bitwise identical.

Its folder `revision/` contains (see `README.md`):

- `julia/run_all.sh`, `replay_seed.jl`, `run_study.jl`, `extract_raw.jl`: replay of 2 plants $\times$ 5 profiles $\times$ 5 seeds and frozen-context extraction (about one minute per episode). Usage: `bash run_all.sh; SEEDS=96301 PROFILES=combined bash run_all.sh` reproduces the archived study.
- `julia/closed_loop/`: closed-loop deployment of the postprocessors (Sec. S5); `bash run_closed_loop.sh`.
- `analysis/seed_study.py`, `seed_summary.py`, `seed_kkt.py`, `propagation.py`: pooled projection study, summaries, numerical primal-dual checks and the constants of Proposition 2 (Python with NumPy, CVXPY/Clarabel, SciPy); `analysis/results/` holds every context and every closed-loop run, and `analysis/inputs/` the packed replay contexts.
- `julia/single_figure/`, `julia/double_figure/`: single-column Figs. 2 and 3 of the main paper from the archived records (CairoMakie).
- `video/`: scripts of Supplementary Video S1, which is submitted as a separate file (94 s, 1920 $\times$ 1200): Part 1 copies, frame for frame, the saved Double-IP dropout replay of seed 96301 at 0.5 $\times$ playback (Fig. 3 of the manuscript); Part 2 animates the closed-loop comparison of Sec. S5 from its saved logs; Part 3 summarizes Tables III–IV. No simulation or training is rerun to make the video (`make_video_v5.py`, `make_cl_frames.py`).