

Logarithmic Regret via Passive Change Detection in Piecewise-Stationary Self-Tuning Regulation

A.Ch. Madhusudanarao

Department of Computer Science and Automation
Indian Institute of Science, Bengaluru, India
madhusudanar@iisc.ac.in

Rahul Singh

Laboratoire de Recherche de l'EPITA, Paris, France
rahulsingh0188@gmail.com

Abstract

We study minimum-variance control of an unknown autoregressive system with exogenous inputs and coefficients that change at unknown times. Under bounded independent disturbances, fixed detection gaps, stability and feasibility conditions, and sufficient time between changes, we prove $O((C + 1) \log((T + 1)/\delta))$ regret with probability at least $1 - \delta$, where T is the horizon and C the number of changes. Unlike switching bandits, where unselected arms can change unobserved, admissible plant changes provide information during exploitation: the correct feasible controller leaves only the disturbance in the output, whereas a detectable change raises output energy under the old controller. PIECE-CD explores initially and after alarms, then uses gated recursive least squares for control. Its energy test compares windowed output power with a threshold above the noise floor; the extension to unstable controller mismatches also monitors the reference controller's input proposal. We control false alarms across the horizon and prove logarithmic detection delay. Inputs are clipped to prescribed bounds. Logarithmic regret also holds under an explicit condition ensuring that clipping becomes inactive after a finite burn-in. Under the stated feasibility conditions, the extended detector covers destabilizing changes with detectable excess energy over a fixed window.

1 Introduction

Self-tuning regulation (STR) seeks minimum-variance control of an unknown autoregressive system with exogenous inputs (ARX). The ideal controller cancels the predictable part of the next output, leaving only the disturbance. Classical work established stability and asymptotic efficiency (Goodwin et al., 1981; Becker et al., 1985; Kumar and Praly, 1987; Lai, 1986; Lai and Wei, 1987); Singh et al. (2024) obtained finite-time logarithmic regret for a stationary plant. We ask whether this rate survives unknown changes in the plant coefficients. A controller must detect these changes and identify the new plant, while avoiding repeated exploration caused by false alarms.

The key observation is that minimum-variance control can supply evidence of a change during exploitation. With the correct feasible controller, $y_{t+1} = w_{t+1}$, so output power equals the disturbance variance σ_w^2 . A change that raises output power under the old controller can therefore be detected from ordinary control observations. This differs from switching bandits, where an unselected arm can change unobserved and even a two-breakpoint class has an $\Omega(\sqrt{T})$ minimax lower bound (Garivier and Moulines, 2011). The distinction is informational: our guarantee requires a fixed detectable energy gap and sufficient time between changes, and does not apply to arbitrary nonstationarity.

Our regret measures the predictable output energy left uncanceled by the learner; its expectation equals the expected excess output energy above the disturbance level ((3)–(4)).

Our algorithm, PIECE-CD, extends Probing Inputs for Exploration in Certainty Equivalence (PIECE). It explores initially and after an alarm, then controls with gated recursive least squares (RLS) around an estimated reference gain. Its principal version combines an output-energy test with a boundary alarm on the reference controller’s requested input. The latter controls clipping errors when a changed plant is driven by the old gain. Outputs remain dependent, and alarm times depend on the observed history; both features matter for the guarantees.

Contributions.

1. **A common detector and logarithmic regret.** Theorem 5 proves no false alarms, logarithmic detection delay, and $\mathcal{R}_T = O((C + 1) \log((T + 1)/\delta))$ with probability at least $1 - \delta$, where C is the number of changes. Under own-regime feasibility, the theorem covers stable mismatches with a stationary power gap and potentially unstable mismatches with a finite-window energy gap, without identifying which condition applies. For the compact minimum-phase ARX class, non-Schur mismatches imply a uniform finite-window gap. The rate holds for fixed model, feasibility, and detectability bounds under the stated dwell-time requirements.
2. **Identification and control after random restarts.** We establish excitation and estimation bounds for complete exploration blocks conditional on the history at each restart. An RLS potential bound controls both accepted RLS actions and reference fallback actions. Chronological failure-event bounds combine these results with detection without assuming independent episodes or conditioning on future alarms.
3. **Supporting guarantees under stable mismatches.** An energy-only baseline isolates the stationary-gap argument (Theorem 4). A separate extension replaces uniform clipping feasibility by a steady-state slack condition and a finite burn-in, preserving logarithmic regret (Theorem 6).

1.1 Related Work

Our closest predecessor is stationary PIECE (Singh et al., 2024). Its covariance analysis treats sparse, recurring exploration episodes. Here the exploration phase is a contiguous block after a data-dependent restart. This requires conditional excitation guarantees and a control analysis that remains valid before the next change or alarm, together with a detector for the dependent observations generated by feedback.

Detect-and-restart methods are well established in piecewise-stationary bandits, including CUSUM-based CD-UCB (Liu et al., 2018), M-UCB (Cao et al., 2019), GLR-klUCB (Besson et al., 2022), and the modular DAB framework (Huang et al., 2026). Their exploration schedules address changes hidden by partial feedback. Additional structural information can permit logarithmic gap-dependent bounds, as when all arms change simultaneously (Mukherjee and Maillard, 2019). PIECE-CD similarly uses a restricted detectable class; its monitoring signal comes from outputs observed during exploitation. Sequential detection provides the statistical background (Veeravalli and Banerjee, 2014; Xie et al., 2021; Tartakovsky et al., 2014).

Fully unknown online linear-quadratic regulation has square-root local minimax regret (Simchowitz and Foster, 2020). STR instead has no positive input penalty, and its ideal controller leaves only the innovation in the output. This property makes the noise floor an observable reference for detecting the changes admitted here.

2 MODEL AND SCOPE

We consider a scalar ARX system of known orders $p, q \geq 1$ whose coefficients change C times over a horizon T . Let $0 = \tau^{(0)} < \tau^{(1)} < \dots < \tau^{(C)} < \tau^{(C+1)} = T$ denote the deterministic segment boundaries, with the C interior boundaries unknown. On segment $c \in \{0, \dots, C\}$,

$$y_{t+1} = \sum_{i=1}^p a_i^{(c)} y_{t-i+1} + \sum_{j=1}^q b_j^{(c)} u_{t-j+1} + w_{t+1}, \quad (1)$$

$$\tau^{(c)} \leq t < \tau^{(c+1)}.$$

Here y_t is the output, u_t the control input, and w_t the disturbance. The unknown parameter vector is constant within each segment:

$$\theta^{(c)} = (a_1^{(c)}, \dots, a_p^{(c)}, b_1^{(c)}, \dots, b_q^{(c)})^\top.$$

Write $\theta_t = \theta^{(c)}$ for $\tau^{(c)} \leq t < \tau^{(c+1)}$.

We use Euclidean vector norms, induced spectral matrix norms, and $\|x\|_M^2 = x^\top M x$. We use two regressors:

$$\psi_t = (y_t, \dots, y_{t-p+1}, u_{t-1}, \dots, u_{t-q+1})^\top,$$

$$\phi_t = (y_t, \dots, y_{t-p+1}, u_t, \dots, u_{t-q+1})^\top.$$

The control regressor $\psi_t \in \mathbb{R}^d$, with $d = p+q-1$, contains quantities available before u_t is chosen. The identification regressor $\phi_t \in \mathbb{R}^{p+q}$ additionally includes u_t , so that $y_{t+1} = (\theta^{(c)})^\top \phi_t + w_{t+1}$. The lagged-input block in ψ_t is empty when $q = 1$.

For a generic parameter vector $\theta = (a_1, \dots, a_p, b_1, \dots, b_q)^\top$ with $b_1 \neq 0$, define the minimum-variance gain

$$\lambda(\theta) := -\frac{1}{b_1} (a_1, \dots, a_p, b_2, \dots, b_q)^\top. \quad (2)$$

On segment c ,

$$y_{t+1} = b_1^{(c)} \{u_t - \lambda(\theta^{(c)})^\top \psi_t\} + w_{t+1}.$$

When feasible, the ideal action $u_t^* = \lambda(\theta^{(c)})^\top \psi_t$ therefore leaves $y_{t+1} = w_{t+1}$. We measure learning by the predictable output energy

$$\mathcal{R}_T := \sum_{t=0}^{T-1} (y_{t+1} - w_{t+1})^2. \quad (3)$$

Let $\mathcal{F}_t$ include the observations and randomization available through the choice of u_t . Then $v_{t+1} := y_{t+1} - w_{t+1}$ is $\mathcal{F}_t$ -measurable and $\mathbb{E}[v_{t+1} w_{t+1} \mid \mathcal{F}_t] = 0$. Hence

$$\mathbb{E} \mathcal{R}_T = \mathbb{E} \sum_{t=0}^{T-1} (y_{t+1}^2 - w_{t+1}^2). \quad (4)$$

This is an expectation identity; the realized excess output cost also contains $2 \sum_t v_{t+1} w_{t+1}$. Our high-probability bounds concern $\mathcal{R}_T$.

We first state the conditions for the stable-mismatch baseline. The extension in Section 5 replaces its mismatch and feasibility conditions; the model, noise, and open-loop switching conditions remain in force. Appendices A–B give the corresponding constants and exploration conditions.

Assumption 1 (Baseline conditions). *The controller knows a compact class $\Theta \subset \mathbb{R}^{p+q}$ containing every $\theta^{(c)}$. Each $\theta \in \Theta$ has a Schur-stable AR polynomial, a minimum-phase input polynomial, and $0 < b_{\min} \leq b_1 \leq \bar{b}$. The autonomous output dynamics are uniformly exponentially stable under admissible switching. The disturbances are i.i.d., independent of the available past, and satisfy $\mathbb{E}w_t = 0$, $\mathbb{E}w_t^2 = \sigma_w^2 > 0$, and $|w_t| \leq B_w$, where σ_w^2 and B_w are known.*

Uniform clipping feasibility holds: $|\lambda(\theta)^\top \psi| \leq B_u$ for every $\theta \in \Theta$ and every regressor ψ reachable under admissible switching, bounded disturbances, and inputs in $[-B_u, B_u]$.

For each change $c \in \{1, \dots, C\}$, let $\theta^- = \theta^{(c-1)}$ and $\theta^+ = \theta^{(c)}$. The unclipped loop obtained by applying the old gain $\lambda(\theta^-)$ to the new plant θ^+ has matrix $A_\Delta := A(\theta^-, \theta^+)$, where $A(\theta^-, \theta^+) := A_{\text{cl}}(\lambda(\theta^-), \theta^+)$, satisfying $\|A_\Delta^k\| \leq C_\Delta \rho_\Delta^k$ for all $k \geq 0$, with known common $C_\Delta < \infty$ and $\rho_\Delta \in (0, 1)$. Its stationary output power satisfies

$$J_y(\lambda(\theta^-); \theta^+) \geq \sigma_w^2 + \Delta \quad (5)$$

for a known fixed $\Delta > 0$. Here $J_y = \sigma_w^2 \sum_{k \geq 0} (e_1^\top A_\Delta^k e_1)^2$, where e_1 is the first coordinate vector in $\mathbb{R}^d$.

For exploration length H and delay bound D defined below, the dwell conditions, when $C \geq 1$, are

$$\begin{aligned} \tau^{(1)} &\geq H, & T - \tau^{(C)} &\geq D, \\ \tau^{(c+1)} - \tau^{(c)} &\geq D + H \quad (1 \leq c < C). \end{aligned} \quad (6)$$

The final exploration may be truncated at T .

Stability and detectability serve different purposes. Stability controls the discrepancy between the learner and the exact old-gain loop; the power gap supplies the detection signal. A parameter change preserving the minimum-variance gain can remain at the noise floor and is excluded by the gap condition. The algorithm uses known model, noise, excitation, and gap bounds for tuning, while the segment parameters and change times are unknown. The alternative feasibility conditions are stated with their respective results in Sections 5 and 6.

Switching stability and clipped inputs give known deterministic bounds $|y_t| \leq B_Y$ and $\|\psi_t\| \leq B_\psi$; their explicit values are given in (40).

3 PIECE-CD

We first specify the energy-only baseline; Section 5 adds the boundary alarm for potentially unstable mismatches. At initialization and after each alarm, PIECE-CD estimates a fixed reference gain from H exploration rounds. During exploitation, a gate keeps RLS actions close to that reference while the detector monitors output energy. Use $\delta_0 = \delta/[32(T+1)]$ for estimation and RLS and $\delta_d = \delta/2$ for detection.

Exploration and the reference controller. Let r denote the start time of the current exploration phase, with $r = 0$ at initialization. After an alarm upon observing y_{t+1} , the next exploration phase starts at $r = t + 1$. For $t = r, \dots, r + H - 1$, draw i.i.d. mean-zero inputs from a prescribed distribution with positive variance and support in $[-B_u, B_u]$, independently of the disturbances. Let $r_0 = \max\{p, q\}$, choose $H > r_0$, and retain the rows $\mathcal{I}_r = \{r + r_0, \dots, r + H - 1\}$, whose lagged measurements all follow the restart.

Fix a ridge parameter $\nu > 0$. Write I_k for the $k \times k$ identity matrix and Π_Θ for a Euclidean

nearest-point projection onto Θ . After completing exploration, compute

$$\begin{aligned} V_r^\theta &= \nu I_{p+q} + \sum_{t \in \mathcal{I}_r} \phi_t \phi_t^\top, \\ \tilde{\theta}_r &= (V_r^\theta)^{-1} \sum_{t \in \mathcal{I}_r} \phi_t y_{t+1}, \\ \hat{\theta}_r &= \Pi_\Theta(\tilde{\theta}_r), \quad \hat{b}_r = b_1(\hat{\theta}_r). \end{aligned} \quad (7)$$

The reference gain $\bar{\lambda}_r = \lambda(\hat{\theta}_r)$ and $\hat{b}_r$ remain fixed until the next restart.

The computable parameter and reference-gain error bounds $E_\theta(H, \delta_0)$ and $E_\lambda(H, \delta_0)$ are defined in (46)–(47). For the prescribed choice of H and a complete exploration block within segment c , with probability at least $1 - 2\delta_0$ conditional on the restart history,

$$\begin{aligned} \|\hat{\theta}_r - \theta^{(c)}\| &\leq E_\theta(H, \delta_0), \\ E_\theta(H, \delta_0) &= O\left(\sqrt{\frac{\log((T+1)/\delta)}{H}}\right). \end{aligned} \quad (8)$$

This reference accuracy sets the gate tolerance: $g_H = 4E_\lambda(H, \delta_0)$. Write $\varepsilon_u = 5E_\lambda(H, \delta_0)B_\psi$ for the resulting reference-tracking tolerance used in the analysis.

Gated control and recursive estimation. Let $\hat{\lambda}_t$ denote the current RLS gain estimate and V_t its regularized $d \times d$ Gram matrix. Initialize both from the exploration data:

$$V_{r+H} = \nu I_d + \sum_{t \in \mathcal{I}_r} \psi_t \psi_t^\top, \quad \hat{\lambda}_{r+H} = \bar{\lambda}_r.$$

At each subsequent control round, form the input proposal z_t and clip it to the prescribed bounds:

$$\begin{aligned} z_t &= \begin{cases} \hat{\lambda}_t^\top \psi_t, & |(\hat{\lambda}_t - \bar{\lambda}_r)^\top \psi_t| \leq g_H \|\psi_t\|, \\ \bar{\lambda}_r^\top \psi_t, & \text{otherwise,} \end{cases} \\ u_t &= \min\{B_u, \max\{-B_u, z_t\}\}. \end{aligned} \quad (9)$$

The gate limits departure from the reference action used for detection; clipping enforces the prescribed input bound.

After observing y_{t+1} , construct the response x_{t+1}^λ and update RLS:

$$\begin{aligned} x_{t+1}^\lambda &= u_t - \frac{y_{t+1}}{\hat{b}_r}, & V_{t+1} &= V_t + \psi_t \psi_t^\top, \\ \hat{\lambda}_{t+1} &= \hat{\lambda}_t + V_{t+1}^{-1} \psi_t \cdot (x_{t+1}^\lambda - \hat{\lambda}_t^\top \psi_t). \end{aligned} \quad (10)$$

RLS uses every exploitation observation, including rounds with a reference fallback or an active clip. The analysis accounts for bias from estimating $b_1^{(c)}$.

Monitoring and restarting. Let the positive integers h and m denote the sampling interval and detection-window size, respectively, chosen below. After exploration, fix the recording times $s_j = r + H + jh$, $j = 1, 2, \dots$. At each s_j , store $y_{s_j}^2$ and retain the most recent m values. Once $j \geq m$, raise an alarm if

$$\hat{J}(s_j) := \frac{1}{m} \sum_{i=j-m+1}^j y_{s_i}^2 \geq \sigma_w^2 + \frac{\Delta}{2}. \quad (11)$$

Control and RLS updates continue every round; only the energy test is subsampled. An alarm at s_j clears the estimates and stored values, sets $r = s_j$, and starts a new exploration phase.

Choosing the exploration and detection lengths. Choose the smallest integer $H > r_0$ satisfying (93), (51), and (52) in Appendix B. These conditions ensure informative exploration and sufficiently accurate reference estimates for control and detection. Choose the smallest $h \geq 1$ satisfying

$$\frac{\sigma_w^2 C_\Delta^2 \rho_\Delta^{2h}}{1 - \rho_\Delta^2} \leq \frac{\Delta}{8}.$$

Recorded outputs remain dependent; their spacing permits concentration around lag- h conditional second moments.

Let B_o be the known common bound in (50) on disturbances, actual outputs, and outputs of the new plant driven by the exact old gain from the state at the change. Set the window size m and associated detection-delay bound D to

$$m = \left\lceil \frac{512B_o^4}{\Delta^2} \log \frac{4T^2}{\delta_d} \right\rceil, \quad D = (m + 1)h. \quad (12)$$

This window size controls fluctuations across the horizon's repeated tests. For fixed problem constants, $h = O(1)$ and $H, m, D = O(\log((T + 1)/\delta))$.

Algorithm 1 PIECE-CD: stable-mismatch baseline

1. Initialize $r = 0$. Stop at T , including during an incomplete phase.
 2. Explore for H rounds; compute (7), the reference gain, and the RLS initialization. Clear the detector buffer and fix the sampling times $s_j = r + H + jh$.
 3. At every exploitation round, apply (9), observe y_{t+1} , and update (10). At sampling times, update the buffer and test (11) once it is full.
 4. If an alarm occurs on observing y_{t+1} , set $r = t + 1$ and return to step 2; otherwise continue step 3.
-

4 FINITE-TIME ANALYSIS

The proof bounds detection delays and false alarms, then controls prediction error on each unchanged exploitation interval. These bounds yield the stable-mismatch baseline in Theorem 4. Section 5 extends the detector analysis to potentially unstable mismatches under alternative input feasibility.

Detection guarantee. Let $\mathcal{E}_{\text{ref}}$ be the event that every reached, complete, unchanged exploration block whose restart precedes the first detector failure produces an accurate reference gain. Here a detector failure is a false alarm or a change not detected by its D -round deadline; the precise reference-event definition is given in (74). Let $\mathcal{E}_{\text{det}}$ be the event of no false alarms and exactly one alarm per change. Writing $\hat{\tau}^{(c)}$ for the output time of the alarm corresponding to change c , this event also requires

$$\tau^{(c)} < \hat{\tau}^{(c)} \leq \tau^{(c)} + D, \quad c = 1, \dots, C. \quad (13)$$

Theorem 2 (Alarm correspondence). *Under the complete assumptions in Appendices A and B and the tuning above,*

$$\mathbb{P}(\mathcal{E}_{\text{ref}} \cap \mathcal{E}_{\text{det}}^c) \leq \delta_d. \quad (14)$$

Proof. See Theorem 18 and Appendix H. □

This joint bound controls detector failure on accurate-reference histories; reference-estimation failures enter Theorem 4 separately.

Why the detector works. Let $\mathcal{F}_t$ denote the information available when u_t is selected, including the controller's randomization and the selected input. Suppose the current reference is accurate and the plant has not changed since its exploration. The gate and input feasibility then give, at each recorded output time s ,

$$y_s = w_s + v_s, \quad |v_s| \leq \varepsilon_y, \quad (15)$$

where $v_s := y_s - w_s$ is $\mathcal{F}_{s-1}$ -measurable and ε_y is the deterministic output-tracking bound in (49). Thus w_s is fresh noise relative to v_s . The tuning makes v_s^2 small enough that concentration of w_s^2 and $2w_s v_s$ keeps unchanged windows below the threshold with high probability.

After a change from θ^- to θ^+ at control time τ , we compare the learner with an auxiliary process that applies the exact old gain $\lambda(\theta^-)$ to the new plant. It starts from the learner's regressor ψ_τ and uses the same future disturbances:

$$\begin{aligned} \psi_\tau^{\text{aux}} &= \psi_\tau, \\ \psi_{t+1}^{\text{aux}} &= A_\Delta \psi_t^{\text{aux}} + e_1 w_{t+1}, \quad t \geq \tau. \end{aligned}$$

Here A_Δ is the closed-loop matrix of the new plant under the old gain, $e_1 \in \mathbb{R}^d$ is the first coordinate vector, and $y_s^{\text{aux}} = e_1^\top \psi_s^{\text{aux}}$ is the auxiliary output; see (68). For a fixed calendar time s with $s - h \geq \tau$, the fresh disturbances over the preceding h steps give

$$\begin{aligned} \mathbb{E}[(y_s^{\text{aux}})^2 \mid \mathcal{F}_{s-h}] &= (e_1^\top A_\Delta^h \psi_{s-h}^{\text{aux}})^2 + \sigma_w^2 \sum_{j=0}^{h-1} (e_1^\top A_\Delta^j e_1)^2 \\ &\geq \sigma_w^2 + \frac{7\Delta}{8}. \end{aligned} \quad (16)$$

The initial-state term is nonnegative, and the choice of h loses at most $\Delta/8$ from the stationary power gap. On an accurate-reference history, the gate and stability also ensure $|y_s^2 - (y_s^{\text{aux}})^2| \leq \Delta/8$ until the next alarm or plant change. With high probability, the average of m fully post-change auxiliary output squares lies at most $\Delta/8$ below its average conditional power. Unless an alarm occurs earlier, the actual average is therefore at least $\sigma_w^2 + 5\Delta/8$, above the threshold. The sampling calendar supplies such a window by $\tau + D$; the dwell-time condition prevents another change before this deadline.

Restart times depend on past observations. For each potential restart, the proof uses a continuation with the same control and recording rules but suppresses subsequent restarts. It agrees with the implemented trajectory through the first alarm. Joint failure bounds over reached histories give the global guarantee without conditioning on future detector success; see Appendices H and I.

With restarts and detection delays controlled, we next bound the RLS regret accumulated during unchanged exploitation.

Theorem 3 (Unchanged exploitation prefixes). *Suppose a full exploration block $[r, r + H)$ lies in one regime and its Gram and estimation events hold. Under the baseline conditions and tuning, conditional on the history at $r + H$, with probability at least $1 - \delta_0$, simultaneously for every exploitation prefix ending no later than the next plant change, restart, or horizon,*

$$\sum_{t=r+H}^{r+H+n-1} (y_{t+1} - w_{t+1})^2 \leq R_E(n, \delta_0). \quad (17)$$

Here R_E is the explicit bound in (78). For fixed model, noise, and exploration constants and the prescribed choice of H , it satisfies $R_E(T, \delta_0) = O(\log((T + 1)/\delta))$.

Theorem 19 and Appendix G give the conditional, time-uniform bound. Its main mechanism is as follows.

RLS argument. Consider an unchanged exploitation interval in segment c after successful exploration, and write $\lambda = \lambda(\theta^{(c)})$ and $b_\star = b_1^{(c)}$. Define the RLS proposal error and the applied action error by

$$a_t = (\hat{\lambda}_t - \lambda)^\top \psi_t, \quad e_t = u_t - \lambda^\top \psi_t.$$

Reference accuracy, the gate, and feasibility of the ideal action imply $e_t^2 \leq a_t^2$. Thus it suffices to control the cumulative RLS proposal error, even on rounds that use the reference input or clipping.

The response used by RLS satisfies

$$\begin{aligned} x_{t+1}^\lambda &= \lambda^\top \psi_t + \alpha e_t + \eta_{t+1}, \\ \alpha &:= 1 - \frac{b_\star}{\widehat{b}_r}, \quad \eta_{t+1} := -\frac{w_{t+1}}{\widehat{b}_r}. \end{aligned}$$

Here αe_t is the predictable bias caused by estimating $b_\star$, whereas η_{t+1} is bounded and conditionally mean zero. To track estimation progress, define

$$\begin{aligned} Q_t &= (\hat{\lambda}_t - \lambda)^\top V_t (\hat{\lambda}_t - \lambda), \\ \ell_t &= \psi_t^\top V_{t+1}^{-1} \psi_t. \end{aligned}$$

Exploration controls both $|\alpha|$ and ℓ_t . Consequently, the exact RLS identity gives

$$Q_{t+1} - Q_t \leq -c_0 a_t^2 + 2a_t \eta_{t+1} + 2\ell_t \eta_{t+1}^2, \quad (18)$$

with $c_0 > 0$ defined in (77).

With probability at least $1 - \delta_0$, a time-uniform martingale bound controls the cumulative mixed term by a fraction of $c_0 \sum_t a_t^2$ plus a confidence term. Bounded noise and the log-determinant bound on $\sum_t \ell_t$ control the final term logarithmically. Summing (18), using $Q_t \geq 0$ and $Q_{r+H} \leq \overline{Q}_H$, therefore bounds $\sum_t a_t^2$ simultaneously over every exploitation prefix. Finally,

$$(y_{t+1} - w_{t+1})^2 = b_\star^2 e_t^2 \leq \widehat{b}^2 a_t^2,$$

which yields Theorem 3. Appendix G gives the identities and concentration details.

Theorem 4 (Stable-mismatch regret guarantee). *Under Assumption 1, the exploration setup in Section 3, and the precise model and exploration bounds in Appendices A–B, choose the smallest feasible H and the detector tuning (12). Then, with probability at least $1 - \delta$, there are no false alarms, every change has exactly one alarm satisfying (13), and*

$$\mathcal{R}_T \leq CK_D D + (C + 1)K_X H + (C + 1)R_E(T, \delta_0), \quad (19)$$

where the fixed per-step delay and exploration bounds K_D, K_X are defined in (84)–(85). Consequently, when all problem constants (model orders, model-class and noise bounds, initial state, exploration and ridge settings, stability constants, clipping threshold, and power gap) are held fixed as T grows,

$$\mathcal{R}_T = O\left((C + 1) \log \frac{T + 1}{\delta}\right). \quad (20)$$

Proof. See Theorem 20 and Appendices I–J. $\square$

Proof sketch of Theorem 4. The regret splits into three parts: each change costs at most $K_D D$ before it is detected (per-step cost K_D times delay D), each of the $C + 1$ exploration blocks costs at most $K_X H$, and Theorem 3 bounds the regret on the remaining unchanged exploitation periods. For the probability, each restart can fail in three ways (insufficient excitation, inaccurate estimate, or RLS failure), each with probability at most δ_0 , and at most $T + 1$ restarts are possible. Adding the detector failure probability δ_d gives a total of $3(T + 1)\delta_0 + \delta_d < \delta$. Because restart times are random, failures are counted in time order, each only if its restart is actually reached; Appendix I gives the details.

5 DETECTION UNDER STABLE AND UNSTABLE MISMATCHES

The baseline theorem in Section 4 assumes that the new plant remains stable under the old controller. We now allow this mismatched closed loop to be unstable. Its stationary power may then be unavailable, and clipping can alter its output-energy signal. We address these issues by converting both detectability conditions below into a common finite-window gap and combining the energy test with an alarm on the reference gain's input proposal. Fixed-window comparisons replace the comparison over the full detection delay used in the stable analysis.

Two sources of detectable energy. Let $\mathcal{P} \subseteq \Theta^2$ be the prescribed class of admissible adjacent parameter pairs. For an actual change $z_c = (\theta^{(c-1)}, \theta^{(c)})$, $c = 1, \dots, C$, write

$$A_c := A_{z_c} := A(\theta^{(c-1)}, \theta^{(c)}),$$

the closed-loop matrix of the new plant under the old gain. We cover $\mathcal{P}$ by two possibly overlapping subfamilies. In the stable subfamily $\mathcal{P}_S$, known common constants $C_S < \infty$, $\rho_S \in (0, 1)$, and $\underline{\Delta}_S > 0$ satisfy

$$\begin{aligned} \|A_c^k\| &\leq C_S \rho_S^k \quad (k \geq 0), \\ J_y(\lambda(\theta^{(c-1)}); \theta^{(c)}) &\geq \sigma_w^2 + \underline{\Delta}_S. \end{aligned} \quad (21)$$

In the finite-visible subfamily $\mathcal{P}_F$, a known common $\underline{\Delta}_F > 0$ instead satisfies

$$\mathcal{V}(z_c) := \sigma_w^2 \sum_{j=1}^d (e_1^\top A_c^j e_1)^2 \geq \underline{\Delta}_F, \quad (22)$$

where $e_1 \in \mathbb{R}^d$ is the first coordinate vector. The coefficients $e_1^\top A_c^j e_1$ describe a disturbance's effect on later outputs, so $\mathcal{V}(z_c)$ measures its excess energy over the first d lags. This condition permits unstable mismatches. Every actual pair belongs to $\mathcal{P} \subseteq \mathcal{P}_S \cup \mathcal{P}_F$; the algorithm uses the class-wide bounds for tuning but never determines which subfamily covers a particular change.

A common finite-window gap. For an integer $h \geq 1$ and a $d \times d$ matrix A , define

$$G_h(A) := \sigma_w^2 \sum_{j=1}^{h-1} (e_1^\top A^j e_1)^2.$$

It measures the excess variance supplied by the preceding $h - 1$ disturbances. For stable pairs, the omitted tail is bounded by $r_S(h)$ below. When both subfamilies are nonempty, choose

$$\begin{aligned} r_S(h) &:= \frac{\sigma_w^2 C_S^2 \rho_S^{2h}}{1 - \rho_S^2}, \\ h_S &:= \min\{h \geq 1 : r_S(h) \leq \underline{\Delta}_S/8\}, \\ h_U &:= \max\{d + 1, h_S\}, \quad \underline{\Delta}_U := \min\{7\underline{\Delta}_S/8, \underline{\Delta}_F\}. \end{aligned} \quad (23)$$

Appendix K gives the convention when a subfamily is empty. Since $h_U \geq h_S$ bounds the stable tail and $h_U \geq d + 1$ includes every term in (22),

$$G_{h_U}(A_c) \geq \begin{cases} 7\underline{\Delta}_S/8, & z_c \in \mathcal{P}_S, \\ \underline{\Delta}_F, & z_c \in \mathcal{P}_F, \end{cases} \geq \underline{\Delta}_U. \quad (24)$$

Stable pairs therefore need not satisfy a separate first- d visibility condition: their stationary gap supplies the required finite-window gap.

The detector and its tuning. Algorithm 3 retains PIECE-CD’s exploration, estimation, gated RLS updates, and restart rule. Its energy test records one output square every h_U rounds and raises an alarm when the average of the latest m_U recorded squares reaches $\sigma_w^2 + \underline{\Delta}_U/2$. Let B_U be the common bound on disturbances, actual outputs, and auxiliary old-gain outputs over an h_U -step comparison, defined in (141). The auxiliary bound is finite even for an unstable mismatch because each comparison lasts only h_U steps. Set the number of recorded outputs and the detection-delay bound to

$$m_U := \left\lceil \frac{512B_U^4}{\underline{\Delta}_U^2} \log \frac{4T^2}{\delta_d} \right\rceil, \quad D_U := (m_U + 1)h_U. \quad (25)$$

To handle clipping, replace uniform feasibility by closed *own-regime feasibility* (Assumption 24). Each ideal gain must respect the input bound on every regressor reached after at least r_0 controls under its own plant, from any initial regressor of norm at most B_ψ , with bounded disturbances and arbitrary inputs in $[-B_u, B_u]$. Equality is allowed. This is an alternative condition: it need not follow from baseline feasibility because it also allows initially unreachable regressors. The old gain need not remain feasible after a change. The additional test monitors the reference gain’s input proposal. After observing y_{t+1} and forming ψ_{t+1} , it raises an alarm if

$$|\bar{\lambda}_r^\top \psi_{t+1}| > B_u + \zeta_H, \quad \zeta_H := E_\lambda(H, \delta_0)B_\psi. \quad (26)$$

With an accurate reference, own-regime feasibility prevents this alarm while the plant is unchanged. After a change, before this alarm fires, the old ideal input exceeds the allowed interval by at most $2\zeta_H$, and the applied input differs from it by at most $7\zeta_H$. These bounds control the effect of clipping in the finite-step comparison.

Choose H to satisfy the exploration and RLS conditions together with (142), which controls finite-window tracking errors. Smaller gaps require longer exploration and detection windows, increasing the required dwell time and regret constant. Either test triggers a restart, with simultaneous crossings counted as one alarm. The following theorem gives the detection and logarithmic-regret guarantees.

Theorem 5 (Detection and regret for the union class). *Retain the compact-model, noise, and open-loop switching-stability conditions of Assumption 1, and impose the own-regime feasibility above. Under the precise union-class conditions in Assumption 26, suppose each actual adjacent pair satisfies at least one of (21) and (22), without revealing which one. Choose H as the smallest feasible integer satisfying (93), (51), and (142); choose m_U, D_U by (25), and impose (6) with D_U in place of D . Then Algorithm 3, with probability at least $1 - \delta$, has no false alarms, and each change has exactly one alarm $\hat{\tau}^{(c)}$ satisfying*

$$\tau^{(c)} < \hat{\tau}^{(c)} \leq \tau^{(c)} + D_U, \quad (27)$$

and

$$\mathcal{R}_T \leq CK_D D_U + (C + 1)K_X H + (C + 1)R_E(T, \delta_0). \quad (28)$$

For fixed model and union-class bounds, including every applicable positive gap bound for the nonempty cover components, this is $O((C + 1) \log((T + 1)/\delta))$. The algorithm never determines which subfamily covers any particular change.

Proof. See Theorem 31 in Appendix L. □

Detection argument for Theorem 5. The proof starts a separate h_U -step auxiliary process before each stored output. If the boundary test has not already crossed, the actual and auxiliary output squares remain close over these finitely many steps. For a pair in $\mathcal{P}_S$, its stationary gap and stable-tail bound give conditional auxiliary energy at least $\sigma_w^2 + 7\underline{\Delta}_S/8$; for a pair in $\mathcal{P}_F$, finite visibility gives at least $\sigma_w^2 + \underline{\Delta}_F$. Equation (24) therefore supplies the same $\underline{\Delta}_U$ to the concentration argument in both cases.

Finite-order visibility of unstable ARX mismatches. Spectral instability alone does not imply visibility for arbitrary matrices. For this compact minimum-phase ARX class, (131) derives a uniform finite-order gap from $\rho(A(\theta^-, \theta^+)) \geq 1$, so no separate visibility assumption is needed on that subclass.

Why input feasibility is needed. Some restriction on saturation is unavoidable for the regret benchmark in (3). For every clipped policy,

$$\mathcal{R}_T \geq b_{\min}^2 \sum_{t=0}^{T-1} (|\lambda(\theta_t)^\top \psi_t| - B_u)_+^2. \quad (29)$$

Appendix K also gives an exploitation bound with this saturation energy as an additive term. Persistent saturation can therefore force linear regret, even when the parameter is known.

6 EVENTUAL CLIPPING INACTIVITY

For stable detectable mismatches, a separate variant replaces uniform clipping feasibility by a steady-state slack condition. Define

$$P_\infty(B_u) := \sqrt{\left(\frac{C_y}{1 - \rho_y}(B_w + B_b B_u)\right)^2 + (q - 1)B_u^2},$$

where C_y, ρ_y are the uniform open-loop switching bounds and B_b bounds the sum of absolute input coefficients. Let $\Lambda = \sup_{\theta \in \Theta} \|\lambda(\theta)\|$. The quantity P_∞ bounds the forced-response contribution to the regressor norm; the slack absorbs the decaying initial-state term.

Theorem 6 (Regret under eventual clipping inactivity). *Retain the baseline model, noise, switching-stability, stable power-gap, and exploration conditions, but replace uniform feasibility by*

$$\mu_0 := B_u - \Lambda P_\infty(B_u) > 0. \quad (30)$$

Choose the smallest baseline-feasible H also satisfying $\varepsilon_u \leq \mu_0/2$. After each full exploration, use the clipped reference controller for τ_{nc} rounds, with RLS and detection disabled, where τ_{nc} is given by (119) with $\mu_{\text{nc}} = \mu_0 - \varepsilon_u$. At $t_0 = r + H + \tau_{\text{nc}}$, initialize RLS from the same exploration data as in Section 3, and record outputs at $t_0 + jh$, $j \geq 1$. Use (6) with $H + \tau_{\text{nc}}$ in place of H . With probability at least $1 - \delta$, there are no false alarms, every change is detected within D rounds, and

$$\mathcal{R}_T \leq CK_D D + (C + 1)K_X H + (C + 1)\{K_D \tau_{\text{nc}} + R_E(T, \delta_0)\}. \quad (31)$$

For fixed problem constants and positive slack, $\tau_{\text{nc}} = O(1)$, so the regret remains $O((C + 1) \log((T + 1)/\delta))$.

Proposition 23 and the burn-in argument in Appendix K.2 establish feasibility after the burn-in, including the continuations needed for detection. Thus only the additional bounded cost per restart enters the regret decomposition.

7 LIMITATIONS

The logarithmic horizon dependence holds for fixed change count, confidence level, and problem constants. The bounds retain the factor $C + 1$ when changes become more frequent. A uniform energy gap and sufficient dwell time are essential: hidden or arbitrarily weak changes are outside

the guarantees. The algorithm also requires supplied bounds on noise, detectability, excitation, and the relevant stability and finite-window sensitivities. Smaller gaps increase the delay and required dwell time. Estimating these bounds, removing knowledge of the horizon, and accommodating heavy-tailed disturbances remain open extensions. The feasibility conditions restrict the nonlinear effects of saturation; they do not cover arbitrary clipped dynamics.

8 CONCLUSION

We establish high-probability regret $O((C + 1) \log((T + 1)/\delta))$ for minimum-variance control of unknown piecewise-stationary ARX systems under fixed detectability gaps and the stated feasibility and dwell-time conditions. PIECE-CD uses output observations during exploitation to detect changes, explores at initialization and after each alarm, and uses gated RLS to control regret between changes. A common detection rule covers both stable controller mismatches with a stationary power gap and potentially unstable mismatches whose disturbance response produces detectable excess energy within a fixed number of steps.

AI Use Statement

Generative AI tools assisted with editing and reorganizing the exposition, checking consistency between main-text statements and the appendices, and LaTeX formatting and compilation checks. The authors take responsibility for the final text, mathematical claims, proofs, and references.

References

- Abbasi-Yadkori, Y., Pál, D., and Szepesvári, C. (2011). Improved algorithms for linear stochastic bandits. In *Advances in Neural Information Processing Systems*, volume 24.
- Audibert, J.-Y. and Bubeck, S. (2009). Minimax policies for adversarial and stochastic bandits. In *Proceedings of the 22nd Annual Conference on Learning Theory*.
- Azuma, K. (1967). Weighted sums of certain dependent random variables. *Tohoku Mathematical Journal*, 19(3):357–367.
- Becker, A., Kumar, P. R., and Wei, C.-Z. (1985). Adaptive control with the stochastic approximation algorithm: Geometry and convergence. *IEEE Transactions on Automatic Control*, 30(4):330–338.
- Besbes, O., Gur, Y., and Zeevi, A. (2014). Stochastic multi-armed-bandit problem with non-stationary rewards. In *Advances in Neural Information Processing Systems*, volume 27, pages 199–207.
- Besson, L., Kaufmann, E., Maillard, O.-A., and Seznec, J. (2022). Efficient change-point detection for tackling piecewise-stationary bandits. *Journal of Machine Learning Research*, 23(77):1–40.
- Cao, Y., Wen, Z., Kveton, B., and Xie, Y. (2019). Nearly optimal adaptive procedure with change detection for piecewise-stationary bandit. In *Proceedings of the Twenty-Second International Conference on Artificial Intelligence and Statistics*, volume 89 of *Proceedings of Machine Learning Research*, pages 418–427. PMLR.
- Garivier, A. and Moulines, E. (2011). On upper-confidence bound policies for switching bandit problems. In *Proceedings of the 22nd International Conference on Algorithmic Learning Theory*, pages 174–188.
- Goodwin, G. C., Ramadge, P. J., and Caines, P. E. (1981). Discrete time stochastic adaptive control. *SIAM Journal on Control and Optimization*, 19(6):829–853.
- Hoeffding, W. (1963). Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58(301):13–30.
- Horn, R. A. and Johnson, C. R. (2013). *Matrix Analysis*. Cambridge University Press, Cambridge, 2 edition.
- Howard, S. R., Ramdas, A., McAuliffe, J., and Sekhon, J. (2020). Time-uniform Chernoff bounds via nonnegative supermartingales. *Probability Surveys*, 17:257–317.
- Huang, Y.-H., Gerogiannis, A., Bose, S., and Veeravalli, V. V. (2026). Detection augmented bandit procedures for piecewise stationary MABs: A modular approach. *IEEE Transactions on Information Theory*.
- Kumar, P. R. and Praly, L. (1987). Self-tuning trackers. *SIAM Journal on Control and Optimization*, 25(4):1053–1071.
- Lai, T. L. (1986). Asymptotically efficient adaptive control in stochastic regression models. *Advances in Applied Mathematics*, 7(1):23–45.
- Lai, T. L. and Robbins, H. (1985). Asymptotically efficient adaptive allocation rules. *Advances in Applied Mathematics*, 6(1):4–22.

- Lai, T. L. and Wei, C.-Z. (1987). Asymptotically efficient self-tuning regulators. *SIAM Journal on Control and Optimization*, 25(2):466–481.
- Liu, F., Lee, J., and Shroff, N. B. (2018). A change-detection based framework for piecewise-stationary multi-armed bandit problem. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 32, pages 3651–3658.
- Mukherjee, S. and Maillard, O.-A. (2019). Distribution-dependent and time-uniform bounds for piecewise i.i.d. bandits. *arXiv preprint arXiv:1905.13159*.
- Simchowitz, M. and Foster, D. (2020). Naive exploration is optimal for online LQR. In *Proceedings of the 37th International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pages 8937–8948. PMLR.
- Singh, R., Mete, A., Kar, A., and Kumar, P. R. (2024). Finite time logarithmic regret bounds for self-tuning regulation. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pages 45571–45636. PMLR.
- Tartakovsky, A., Nikiforov, I., and Basseville, M. (2014). *Sequential Analysis: Hypothesis Testing and Changepoint Detection*. CRC Press.
- Tropp, J. A. (2011). Freedman’s inequality for matrix martingales. *Electronic Communications in Probability*, 16:262–270.
- Veeravalli, V. V. and Banerjee, T. (2014). Quickest change detection. *Academic Press Library in Signal Processing*, 3:209–255.
- Ville, J. (1939). *Étude critique de la notion de collectif*. Gauthier-Villars, Paris.
- Xie, L., Zou, S., Xie, Y., and Veeravalli, V. V. (2021). Sequential (quickest) change detection: Classical results and new directions. *IEEE Journal on Selected Areas in Information Theory*, 2(2):494–514.
- Zhou, H., Wang, L., Varshney, L. R., and Lim, E.-P. (2020). A near-optimal change-detection based algorithm for piecewise-stationary combinatorial semi-bandits. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 34, pages 6933–6940.

APPENDIX CONTENTS

A	MODEL AND ASSUMPTIONS.....	17
B	THE PIECE-CD ALGORITHM	19
C	CHANGE-DETECTION GUARANTEES.....	22
D	REGRET ANALYSIS.....	26
E	DETERMINISTIC BOUNDS AND RESTART CONVENTIONS.....	29
F	EXPLORATION, EXCITATION, AND THE RESTARTED ESTIMATE.....	30
G	THE RLS POTENTIAL ARGUMENT.....	34
H	DETAILED DETECTOR PROOFS.....	37
I	GLOBAL EVENT AND REGRET DECOMPOSITION.....	39
J	TUNING FEASIBILITY AND ORDERS.....	41
K	RELAXATIONS OF CLIPPING AND MISMATCH STABILITY	42
L	A UNIFIED DETECTOR FOR THE UNION CLASS.....	52

ROADMAP OF THE APPENDICES

- A.** Specifies the piecewise-stationary ARX model, regret criterion, and assumptions on stability, noise, clipping feasibility, and detectability of adjacent changes.
- B.** Gives the exploration estimator, gated RLS controller, output-energy detector, tuning and dwell-time requirements, and complete PIECE-CD algorithm.
- C.** Proves action tracking, sampled-square concentration, false-alarm control, detection delay, and correspondence between alarms and true changes.
- D.** Derives exploitation regret on unchanged segments and combines it with exploration and detection-delay costs to obtain the total regret guarantee.
- E.** Establishes deterministic output and regressor bounds, clipping projection inequalities, auxiliary-process bounds, and conditioning conventions at random restarts.
- F.** Proves conditional excitation, exploration Gram-matrix growth, and projected ridge estimation bounds for a complete unchanged exploration block.
- G.** Develops the RLS potential identity, time-uniform noise control, and elliptical-potential bound underlying the segmentwise exploitation guarantee.
- H.** Supplies detailed auxiliary-process power, sampling, coupling, and calendar arguments supporting the detection results in Appendix C.
- I.** Constructs a common success event and partitions the horizon to complete the proof of the global regret bound.
- J.** Verifies feasibility of the tuning requirements and derives the logarithmic orders of exploration length, detection delay, and regret.
- K.** Develops the clipping-feasibility relaxations, eventual nonclipping, boundary alarms, finite-order visibility and union-class detection, and regret bounds with saturation.
- L.** Presents the unified union-class algorithm and common tuning, then proves detection and regret guarantees without identifying the realized mismatch class.

A MODEL AND ASSUMPTIONS

Fix an integer horizon $T \geq 1$, confidence level $\delta \in (0, 1)$, and integer model orders $p, q \geq 1$. The deterministic but unknown segment boundaries are

$$0 = \tau^{(0)} < \tau^{(1)} < \dots < \tau^{(C)} < \tau^{(C+1)} = T.$$

On segment c , the parameter

$$\theta^{(c)} := (a_1^{(c)}, \dots, a_p^{(c)}, b_1^{(c)}, \dots, b_q^{(c)})'$$

is active at control times $\tau^{(c)} \leq t < \tau^{(c+1)}$, and the system obeys

$$y_{t+1} = \sum_{i=1}^p a_i^{(c)} y_{t-i+1} + \sum_{j=1}^q b_j^{(c)} u_{t-j+1} + w_{t+1}, \quad \tau^{(c)} \leq t < \tau^{(c+1)}. \quad (32)$$

Write $\theta_t = \theta^{(c)}$ on this segment and $Y_t = (y_t, \dots, y_{t-p+1})'$. The initial history is fixed, with $\|Y_0\| < \infty$ and $|u_t| \leq B_u$ for $t = -q+1, \dots, -1$. The control and identification regressors are

$$\psi_t := (y_t, \dots, y_{t-p+1}, u_{t-1}, \dots, u_{t-q+1})', \quad (33)$$

$$\phi_t := (y_t, \dots, y_{t-p+1}, u_t, u_{t-1}, \dots, u_{t-q+1})', \quad d := p + q - 1. \quad (34)$$

For $\theta = (a_1, \dots, a_p, b_1, \dots, b_q)$ with $b_1 \neq 0$, define

$$\lambda(\theta) := -\frac{1}{b_1} (a_1, \dots, a_p, b_2, \dots, b_q)'. \quad (35)$$

The corresponding gain representation is

$$y_{t+1} = b_1^{(c)} \{u_t - \lambda(\theta^{(c)})' \psi_t\} + w_{t+1}. \quad (36)$$

The ideal input $u_t = \lambda(\theta^{(c)})' \psi_t$ leaves only the disturbance in the output. Our sample-path learning regret is

$$\mathcal{R}_T := \sum_{t=0}^{T-1} (y_{t+1} - w_{t+1})^2.$$

Let $\mathcal{F}_t$ contain the initial history, the controller randomization revealed through time t , and all inputs and outputs available when u_t is selected. In particular, u_t is $\mathcal{F}_t$ -measurable.

Assumption 7 (Known model class). *The controller knows a compact set $\Theta \subset \mathbb{R}^{p+q}$ containing every $\theta^{(c)}$. For each $\theta \in \Theta$, the autoregressive polynomial is Schur stable, the input polynomial is minimum phase, and*

$$0 < b_{\min} \leq b_1(\theta) \leq \bar{b} < \infty. \quad (37)$$

The gain map in (35) is therefore Lipschitz on Θ ; fix a strictly positive known Lipschitz bound $L_\lambda > 0$ satisfying

$$\|\lambda(\theta) - \lambda(\vartheta)\| \leq L_\lambda \|\theta - \vartheta\|, \quad \theta, \vartheta \in \Theta. \quad (38)$$

Write $\Lambda := \sup_{\theta \in \Theta} \|\lambda(\theta)\|$ and $B_b := \sup_{\theta \in \Theta} \sum_{j=1}^q |b_j(\theta)|$.

Let $A_y(\theta)$ be the companion matrix in (88). A parameter sequence is admissible if it takes values in Θ and is piecewise constant at the specified segment boundaries. Individual Schur stability does not ensure stability under switching, so we impose the following uniform product bound.

Assumption 8 (Uniform switching stability). *There are $1 \leq C_y < \infty$ and $\rho_y \in (0, 1)$ such that, for every admissible parameter sequence and every $k \geq 0$,*

$$\|A_y(\theta_{t+k-1}) \cdots A_y(\theta_t)\| \leq C_y \rho_y^k. \quad (39)$$

Assumption 9 (Noise). *Each w_t is independent of $\mathcal{F}_{t-1}$, and the disturbances are identically distributed. They satisfy*

$$\mathbb{E}w_t = 0, \quad \mathbb{E}w_t^2 = \sigma_w^2 > 0, \quad |w_t| \leq B_w \quad \text{almost surely.}$$

The detector knows σ_w^2 and B_w .

The algorithm always clips its input to $[-B_u, B_u]$. Consequently, Assumptions 7–9 imply the deterministic bounds

$$\begin{aligned} B_Y &:= C_y \|Y_0\| + \frac{C_y}{1 - \rho_y} (B_w + B_b B_u), & |y_t| &\leq B_Y, \\ B_\psi &:= \{p B_Y^2 + (q - 1) B_u^2\}^{1/2}, & \|\psi_t\| &\leq B_\psi. \end{aligned} \quad (40)$$

A proof is given in Appendix E.

For the baseline analysis, the ideal action must also be feasible at every reachable regressor. Here, reachability is from the fixed initial history under admissible parameter sequences, disturbances bounded by B_w , and inputs in $[-B_u, B_u]$.

Assumption 10 (Uniform clipping feasibility). *The ideal action belongs to the closed clipping interval:*

$$|\lambda(\theta)' \psi| \leq B_u \quad \text{for every } \theta \in \Theta \quad \text{and every reachable } \psi \text{ with } \|\psi\| \leq B_\psi.$$

The stronger condition $\Lambda B_\psi \leq B_u$ is sufficient. Equality is allowed; projection nonexpansiveness requires no positive margin.

For a gain $\ell \in \mathbb{R}^d$ and parameter θ , let $A_{\text{cl}}(\ell, \theta)$ be the matrix satisfying

$$\psi_{t+1} = A_{\text{cl}}(\ell, \theta) \psi_t + e_1 w_{t+1} \quad (41)$$

when the unclipped input $u_t = \ell' \psi_t$ is applied. Here e_1 is the first coordinate vector. If the input is $u_t = \ell' \psi_t + e_t$, the regressor dynamics become

$$\psi_{t+1} = A_{\text{cl}}(\ell, \theta) \psi_t + g_u(\theta) e_t + e_1 w_{t+1},$$

where the input-channel vector and its uniform bound are defined by

$$g_u(\theta) := \begin{cases} b_1(\theta) e_1, & q = 1, \\ (b_1(\theta), 0'_{p-1}, 1, 0'_{q-2})', & q \geq 2, \end{cases} \quad G_u := \sup_{\theta \in \Theta} \|g_u(\theta)\| \leq \sqrt{\bar{b}^2 + 1}. \quad (42)$$

Here 0_k denotes the zero vector in $\mathbb{R}^k$; the final block is absent when $q = 2$. When $A_{\text{cl}}(\ell, \theta)$ is Schur stable, define its stationary output power

$$J_y(\ell; \theta) := \sigma_w^2 \sum_{j=0}^{\infty} \{e_1' A_{\text{cl}}(\ell, \theta)^j e_1\}^2. \quad (43)$$

To justify this terminology, freeze (ℓ, θ) and put $A = A_{\text{cl}}(\ell, \theta)$. The zero-mean stationary solution of $\psi_{t+1} = A \psi_t + e_1 w_{t+1}$ is

$$\psi_t^{\text{stat}} = \sum_{j=0}^{\infty} A^j e_1 w_{t-j}.$$

The series converges in mean square because A is Schur. Its first coordinate is the stationary output. Independence, zero mean, and $\mathbb{E}w_t^2 = \sigma_w^2$ eliminate all cross terms, so

$$\begin{aligned}\mathbb{E}[(y_t^{\text{stat}})^2] &= \sum_{j,k \geq 0} (e_1' A^j e_1)(e_1' A^k e_1) \mathbb{E}[w_{t-j} w_{t-k}] \\ &= \sigma_w^2 \sum_{j=0}^{\infty} (e_1' A^j e_1)^2 = J_y(\ell; \theta).\end{aligned}$$

Thus (43) is the stationary second moment (and, because the mean is zero, the variance) of the *unclipped frozen-gain* output process. It is not asserted directly for the implemented clipped trajectory; Lemmas 13 and 16 provide that comparison. In particular, $J_y(\lambda(\theta); \theta) = \sigma_w^2$.

Assumption 11 (Stable and detectable adjacent changes). *There are $C_\Delta < \infty$, $\rho_\Delta \in (0, 1)$, and $\Delta > 0$ such that, for every actual adjacent pair $(\theta^{(c-1)}, \theta^{(c)})$,*

$$\|A_{\text{cl}}(\lambda(\theta^{(c-1)}), \theta^{(c)})^k\| \leq C_\Delta \rho_\Delta^k, \quad k \geq 0, \quad (44)$$

$$J_y(\lambda(\theta^{(c-1)}); \theta^{(c)}) \geq \sigma_w^2 + \Delta. \quad (45)$$

The power-gap condition excludes changes that leave the old controller's output power unchanged. Appendix K gives two extensions: a burn-in variant under eventual nonclipping, and a boundary-alarm variant covering both stable power gaps and finite-order visibility.

Tuning requires valid bounds for C_Δ , ρ_Δ , and Δ , and the model and exploration bounds defining B_ψ , B_o , and κ_G in (40), (50), and (92). Conservative bounds suffice; the analysis does not provide a data-driven calibration rule.

B THE PIECE-CD ALGORITHM

Each PIECE-CD cycle consists of H exploration rounds followed by gated RLS control and output-energy monitoring. An alarm starts a new cycle.

B.1 Exploration and Estimation

Let $s := p + q$, $r_0 := \max\{p, q\}$, and $n_H := H - r_0$. Fix a ridge parameter $\nu > 0$. Exploration uses fresh i.i.d. mean-zero inputs, bounded by $B_e \leq B_u$ and of variance $\sigma_e^2 > 0$, independent of the disturbance sequence. Conditional on $\mathcal{F}_r$, u_r is fixed and the subsequent exploration inputs remain independent with this distribution. For a full unchanged exploration block satisfying (93), Proposition 21 gives

$$\lambda_{\min} \left(\sum_{t=r+r_0}^{r+H-1} \phi_t \phi_t' \right) \geq \kappa_G n_H$$

with conditional probability at least $1 - \delta_0$ given $\mathcal{F}_r$. The constant $\kappa_G > 0$ in (92) is uniform over Θ and the bounded restart states.

Put $B_\phi := (B_\psi^2 + B_u^2)^{1/2}$ and $S_\Theta := \sup_{\theta \in \Theta} \|\theta\|$. For a confidence level δ_0 , define the deterministic estimation bound

$$\mathcal{L}_H(\delta_0) := 2 \log(1/\delta_0) + s \log \left(1 + \frac{H B_\phi^2}{s \nu} \right),$$

$$E_\theta(H, \delta_0) := \frac{2\{B_w \sqrt{\mathcal{L}_H(\delta_0)} + \sqrt{\nu} S_\Theta\}}{\sqrt{\nu} + \kappa_G n_H}, \quad (46)$$

$$E_\lambda(H, \delta_0) := L_\lambda E_\theta(H, \delta_0), \quad g_H := 4E_\lambda(H, \delta_0). \quad (47)$$

For the estimator and reference gain in (55)–(56), Proposition 22 establishes $\|\widehat{\theta}_r - \theta\| \leq E_\theta$ and $\|\bar{\lambda}_r - \lambda(\theta)\| \leq E_\lambda$ on the joint Gram and estimation event.

For the horizon-wide theorem we use

$$\delta_d := \delta/2, \quad \delta_0 := \frac{\delta}{32(T+1)}. \quad (48)$$

This allocation remains valid without knowing C , because no more than $T+1$ restart locations are possible.

B.2 Tuning and Dwell Time

On the successful reference event, Lemma 13 bounds the action and unchanged-regime output errors by

$$\varepsilon_u := \frac{5}{4}g_H B_\psi, \quad \varepsilon_y := \bar{b}\varepsilon_u. \quad (49)$$

For a post-change fixed-old-gain auxiliary trajectory, set

$$\begin{aligned} B_{\text{aux}} &:= C_\Delta B_\psi + \frac{C_\Delta B_w}{1 - \rho_\Delta}, & B_o &:= \max\{B_Y, B_{\text{aux}}, B_w\}, \\ R_\Delta(H) &:= \frac{C_\Delta G_u \varepsilon_u}{1 - \rho_\Delta}. \end{aligned} \quad (50)$$

Choose $H > r_0$ as the smallest integer for which the sample-size condition (93) and the following inequalities hold:

$$\frac{E_\theta(H, \delta_0)}{b_{\min}} \leq \frac{1}{16}, \quad \frac{B_\psi^2}{\nu + \kappa_G n_H} \leq \frac{1}{16}, \quad (51)$$

$$\varepsilon_y^2 \leq \frac{\Delta}{8}, \quad \varepsilon_y \leq B_o, \quad 2B_o R_\Delta(H) \leq \frac{\Delta}{8}. \quad (52)$$

Condition (51) controls the RLS bias and leverage; (52) controls the null bias and coupling error.

Choose the integer sampling interval h , window size m , and delay bound D by

$$h := \min \left\{ k \geq 1 : \frac{\sigma_w^2 C_\Delta^2 \rho_\Delta^{2k}}{1 - \rho_\Delta^2} \leq \frac{\Delta}{8} \right\}, \quad (53)$$

$$m := \left\lceil \frac{512B_o^4}{\Delta^2} \log \frac{4T^2}{\delta_d} \right\rceil, \quad D := (m+1)h. \quad (54)$$

The spacing truncates the stationary-power tail at $\Delta/8$. The window size controls both null and post-change deviations uniformly over potential restart locations and test endpoints.

Assumption 12 (Minimum dwell time). *The first change satisfies $\tau^{(1)} \geq H$ when $C \geq 1$. Every pair of successive changes satisfies*

$$\tau^{(c+1)} - \tau^{(c)} \geq D + H, \quad c = 1, \dots, C-1,$$

and the terminal segment satisfies $T - \tau^{(C)} \geq D$ when $C \geq 1$. When each change is detected within D rounds, these conditions leave a full unchanged exploration block before the next change. The final exploration block may be truncated at T .

B.3 Algorithm

For a vector x , let $\Pi_\Theta(x)$ be a measurable Euclidean projection onto Θ . If Θ is not convex, any measurable nearest-point selection can be used; the only property required below is $\|\Pi_\Theta(x) - \theta\| \leq 2\|x - \theta\|$ for $\theta \in \Theta$. Indeed, since θ itself is an admissible point in the nearest-point minimization,

$$\|\Pi_\Theta(x) - x\| \leq \|\theta - x\|.$$

The triangle inequality therefore gives $\|\Pi_\Theta(x) - \theta\| \leq \|\Pi_\Theta(x) - x\| + \|x - \theta\| \leq 2\|x - \theta\|$. This factor transfers the unprojected ridge error to the feasible estimate in Proposition 22. If Θ is closed and convex, projection is nonexpansive and the factor two can be replaced by one.

Algorithm 2: The PIECE-CD algorithm

Input: the known orders and model class; T, δ ; all tuning quantities defined above; and the exploration distribution.

Initialize the restart time $r \leftarrow 0$ and clear the detector buffer.

Loop. While $r < T$, execute the following cycle.

- (a) **Explore.** For $t = r, \dots, \min\{r + H - 1, T - 1\}$, draw the independent exploratory input, apply it, and observe y_{t+1} . If the horizon ends, stop.
- (b) **Estimate.** Using the post-boundary rows $\mathcal{I}_r := \{r + r_0, \dots, r + H - 1\}$, form

$$V_r^\theta := \nu I_s + \sum_{t \in \mathcal{I}_r} \phi_t \phi_t', \quad \tilde{\theta}_r := (V_r^\theta)^{-1} \sum_{t \in \mathcal{I}_r} \phi_t y_{t+1}, \quad (55)$$

$$\hat{\theta}_r := \Pi_\Theta(\tilde{\theta}_r), \quad \hat{b}_r := b_1(\hat{\theta}_r), \quad \bar{\lambda}_r := \lambda(\hat{\theta}_r). \quad (56)$$

Initialize

$$V_{r+H} := \nu I_d + \sum_{t \in \mathcal{I}_r} \psi_t \psi_t', \quad \hat{\lambda}_{r+H} := \bar{\lambda}_r. \quad (57)$$

Clear the detector buffer and declare the output-time sampling grid

$$\mathcal{S}_r := \{r + H + h + jh : j = 0, 1, \dots\}.$$

- (c) **Exploit and monitor.** For $t = r + H, r + H + 1, \dots, T - 1$, set

$$z_t := \begin{cases} \hat{\lambda}_t' \psi_t, & |(\hat{\lambda}_t - \bar{\lambda}_r)' \psi_t| \leq g_H \|\psi_t\|, \\ \bar{\lambda}_r' \psi_t, & \text{otherwise,} \end{cases} \quad u_t := \text{clip}_{[-B_u, B_u]}(z_t). \quad (58)$$

Before observing y_{t+1} , declare it detector-eligible exactly when $t + 1 \in \mathcal{S}_r$. Observe y_{t+1} and make the standard RLS update with pseudo-response

$$x_{t+1}^\lambda := u_t - \frac{y_{t+1}}{\hat{b}_r}, \quad V_{t+1} := V_t + \psi_t \psi_t', \quad (59)$$

$$\hat{\lambda}_{t+1} := \hat{\lambda}_t + V_{t+1}^{-1} \psi_t (x_{t+1}^\lambda - \hat{\lambda}_t' \psi_t). \quad (60)$$

If y_{t+1} was declared eligible, append y_{t+1}^2 to the buffer. Retain only the most recent m stored values. When the buffer is full, compute

$$\hat{J}(t + 1) := \frac{1}{m} \sum_{z \in \text{buffer}} z.$$

If $\hat{J}(t + 1) \geq \sigma_w^2 + \Delta/2$, raise an alarm, set $r \leftarrow t + 1$, clear the buffer, and begin a new cycle. If no alarm is raised and $t = T - 1$, stop.

C CHANGE-DETECTION GUARANTEES

Throughout this appendix, the assumptions and tuning of Appendices A and B are in force. Fix a restart r with a full *uncontaminated* exploration block: the parameter remains equal to θ^- throughout $[r, r + H)$. Write $\lambda^- := \lambda(\theta^-)$ and define the reference event

$$\mathcal{G}_r^{\text{est}} := \left\{ \left\| \bar{\lambda}_r - \lambda^- \right\| \leq g_H/4 \right\}$$

whose failure probability is controlled by Proposition 22.

Lemma 13 (Gate and action tracking). *On $\mathcal{G}_r^{\text{est}}$, at every exploitation time before the next alarm, including times after an undetected change,*

$$|u_t - (\lambda^-)' \psi_t| \leq \frac{5}{4} g_H \|\psi_t\| \leq \varepsilon_u. \quad (61)$$

If no true change has occurred and $a_t := (\hat{\lambda}_t - \lambda^-)' \psi_t$, then the applied action error $e_t := u_t - (\lambda^-)' \psi_t$ also satisfies

$$e_t^2 \leq a_t^2. \quad (62)$$

Proof. If the recursive branch of (58) is selected, then

$$|z_t - (\lambda^-)' \psi_t| \leq g_H \|\psi_t\| + (g_H/4) \|\psi_t\|.$$

If the reference branch is selected, the same left-hand side is at most $(g_H/4) \|\psi_t\|$. Since the ideal action $(\lambda^-)' \psi_t$ lies in $[-B_u, B_u]$ by Assumption 10, the metric projection onto this interval is nonexpansive and (61) follows after clipping.

For (62), equality holds in the recursive branch before clipping, and nonexpansiveness can only reduce the error. In the reference branch,

$$\begin{aligned} |a_t| &\geq \left| (\hat{\lambda}_t - \bar{\lambda}_r)' \psi_t \right| - \left| (\bar{\lambda}_r - \lambda^-)' \psi_t \right| \\ &> \frac{3}{4} g_H \|\psi_t\|, \end{aligned}$$

whereas the reference action error is at most $(g_H/4) \|\psi_t\|$. Clipping again cannot increase it. $\square$

C.1 Sampled-Square Concentration

The sampling interval h allows concentration around conditional moments given the history h steps earlier.

Lemma 14 (Lagged sampled-square concentration). *Let $h \leq s_1 < \dots < s_m$ be finite lag- h predictable output times, meaning $\{s_k = t\} \in \mathcal{F}_{t-h}$ for every t , and suppose $s_{k+1} - s_k \geq h$. If X_{s_k} is $\mathcal{F}_{s_k}$ -measurable and $|X_{s_k}| \leq B_o$, then, for every $a > 0$,*

$$\mathbb{P} \left(\frac{1}{m} \sum_{k=1}^m \{X_{s_k}^2 - \mathbb{E}[X_{s_k}^2 \mid \mathcal{F}_{s_k-h}]\} \leq -a \right) \leq \exp \left(-\frac{ma^2}{2B_o^4} \right). \quad (63)$$

The analogous upper-tail bound also holds.

Proof. Set $Z_k = X_{s_k}^2 - \mathbb{E}[X_{s_k}^2 \mid \mathcal{F}_{s_k-h}]$. With $\mathcal{H}_{k-1} = \mathcal{F}_{s_k-h}$ and, for $k < m$, $\mathcal{H}_k = \mathcal{F}_{s_{k+1}-h}$, the spacing condition implies that Z_k is $\mathcal{H}_k$ -measurable and has conditional mean zero. Moreover, $\mathcal{H}_m := \mathcal{F}_{s_m}$ contains Z_m . Lag predictability makes the random-indexed sigma-fields well defined as stopped sigma-fields, and $\mathcal{F}_{s_k} \subseteq \mathcal{F}_{s_{k+1}-h}$. Moreover, $|Z_k| \leq B_o^2$. The Azuma–Hoeffding inequality (Azuma, 1967; Hoeffding, 1963) gives (63); the upper tail follows by applying it to $-Z_k$. $\square$

C.2 False Alarms

For each potential restart r , define an *alarm-suppressed continuation* by retaining the control, RLS, grid, and buffer recursions while recording threshold crossings without restarting. It coincides with the implemented trajectory through the first alarm.

A block is *null for restart r* if every stored output is generated under the reference parameter θ^- . If that parameter remains active until control time τ , a block ending at output time τ is null: y_τ is generated by $u_{\tau-1}$. On $\mathcal{G}_r^{\text{est}}$, (36) and Lemma 13 give, for each stored output in a null block,

$$y_s = w_s + v_s, \quad v_s := b_1(\theta^-)\{u_{s-1} - (\lambda^-)'\psi_{s-1}\}, \quad |v_s| \leq \varepsilon_y. \quad (64)$$

The variable v_s is $\mathcal{F}_{s-1}$ -measurable, while w_s is independent of $\mathcal{F}_{s-1}$. Hence, for any null detector block S of m stored samples,

$$\frac{1}{m} \sum_{s \in S} (y_s^2 - \sigma_w^2) = \frac{1}{m} \sum_{s \in S} (w_s^2 - \sigma_w^2) + \frac{2}{m} \sum_{s \in S} w_s v_s + \frac{1}{m} \sum_{s \in S} v_s^2. \quad (65)$$

Lemma 15 (One-block false-alarm bound). *Suppose (52) holds. Conditional on any $\mathcal{F}_{r+H}$ history on which $\mathcal{G}_r^{\text{est}}$ holds, any fixed potential block generated by the predictable sampling grid that is null for restart r crosses the threshold with probability at most $\delta_d/(2T^2)$.*

Proof. By (64), every summand in the last term of (65) satisfies $v_s^2 \leq \varepsilon_y^2$. Consequently,

$$\frac{1}{m} \sum_{s \in S} v_s^2 \leq \varepsilon_y^2 \leq \frac{\Delta}{8},$$

where the final inequality is exactly the first condition in (52). The standard bounded-increment exponential supermartingale for the predictably sampled disturbance squares (Hoeffding, 1963; Howard et al., 2020), applied along the fixed predictable block through its predetermined endpoint, gives

$$\mathbb{P}\left(\frac{1}{m} \sum_{s \in S} (w_s^2 - \sigma_w^2) > \frac{\Delta}{8}\right) \leq \exp\left(-\frac{m\Delta^2}{32B_w^4}\right). \quad (66)$$

The products $w_s v_s$ form bounded martingale differences, so the same Azuma–Hoeffding construction (Azuma, 1967; Hoeffding, 1963) applies. If $\varepsilon_y = 0$, this cross term vanishes identically. Otherwise, the analogous stopped exponential bound gives

$$\mathbb{P}\left(\frac{2}{m} \sum_{s \in S} w_s v_s > \frac{\Delta}{8}\right) \leq \exp\left(-\frac{m\Delta^2}{512B_w^2\varepsilon_y^2}\right). \quad (67)$$

With m from (54) and $B_o \geq \max\{B_w, \varepsilon_y\}$, each right-hand side is at most $\delta_d/(4T^2)$. Outside these two deviations, the left side of (65) is at most $3\Delta/8 < \Delta/2$. $\square$

C.3 Detection after a Change

Suppose the current cycle has not alarmed when the parameter changes from θ^- to θ^+ at control time τ . Starting from $\psi_\tau^{\text{aux}} = \psi_\tau$, define the auxiliary process driven by the same future disturbances and the exact old gain:

$$\psi_{t+1}^{\text{aux}} = A_\Delta \psi_t^{\text{aux}} + e_1 w_{t+1}, \quad A_\Delta := A(\theta^-, \theta^+). \quad (68)$$

Let y_t^{aux} be its first coordinate. By (44), $|y_t^{\text{aux}}| \leq B_{\text{aux}}$.

Lemma 16 (Actual-to-auxiliary coupling). *On $\mathcal{G}_r^{\text{est}}$, until the first of the next alarm and the next true change after τ ,*

$$\|\psi_t - \psi_t^{\text{aux}}\| \leq R_\Delta(H), \quad (69)$$

$$\left| y_t^2 - (y_t^{\text{aux}})^2 \right| \leq 2B_o R_\Delta(H) \leq \Delta/8. \quad (70)$$

Proof. Writing the actual input as $u_t = \lambda(\theta^-)' \psi_t + e_t$ and subtracting (68) yields

$$\psi_{t+1} - \psi_{t+1}^{\text{aux}} = A_\Delta(\psi_t - \psi_t^{\text{aux}}) + g_u(\theta^+)e_t.$$

Put $\delta_t := \psi_t - \psi_t^{\text{aux}}$. Because the two processes start from the same state, $\delta_\tau = 0$. Iterating the displayed recursion gives, for every $t > \tau$ before the stopping endpoint,

$$\delta_t = \sum_{k=\tau}^{t-1} A_\Delta^{t-1-k} g_u(\theta^+) e_k.$$

Lemma 13 gives $|e_k| \leq \varepsilon_u$, while (42) and (44) give $\|g_u(\theta^+)\| \leq G_u$ and $\|A_\Delta^j\| \leq C_\Delta \rho_\Delta^j$. Therefore

$$\begin{aligned} \|\delta_t\| &\leq C_\Delta G_u \varepsilon_u \sum_{k=\tau}^{t-1} \rho_\Delta^{t-1-k} \\ &\leq \frac{C_\Delta G_u \varepsilon_u}{1 - \rho_\Delta} = R_\Delta(H), \end{aligned}$$

which is (69). Since output is the first state coordinate, $|y_t - y_t^{\text{aux}}| \leq \|\delta_t\| \leq R_\Delta(H)$. Both outputs are bounded by B_o , and hence

$$\begin{aligned} |y_t^2 - (y_t^{\text{aux}})^2| &= |y_t - y_t^{\text{aux}}| |y_t + y_t^{\text{aux}}| \\ &\leq 2B_o R_\Delta(H) \leq \frac{\Delta}{8}, \end{aligned}$$

where the last inequality is the second condition in (52). This proves (70). $\square$

For every auxiliary output time s whose preceding h controls are all post-change, independence gives the exact finite-horizon identity

$$\mathbb{E}[(y_s^{\text{aux}})^2 \mid \mathcal{F}_{s-h}] = (e_1' A_\Delta^h \psi_{s-h}^{\text{aux}})^2 + \sigma_w^2 \sum_{j=0}^{h-1} (e_1' A_\Delta^j e_1)^2. \quad (71)$$

Indeed, iterating (68) for h steps and taking the first coordinate gives

$$y_s^{\text{aux}} = e_1' A_\Delta^h \psi_{s-h}^{\text{aux}} + \sum_{j=0}^{h-1} e_1' A_\Delta^j e_1 w_{s-j}.$$

Conditional on $\mathcal{F}_{s-h}$, the first term is fixed. Every future disturbance is centered and independent of that field and of the other future disturbances. Thus the state-noise cross terms and the cross terms with distinct disturbance indices have conditional expectation zero, while each diagonal disturbance term contributes its coefficient squared times σ_w^2 . This proves (71) at deterministic s . The same identity holds at a finite lag- h predictable random output time: on each event $\{s = t\} \in \mathcal{F}_{t-h}$ it is the deterministic-time identity, and summing over t gives the

equality conditional on the stopped field $\mathcal{F}_{s-h}$. The first term is nonnegative. Equations (45) and (53) imply the remaining inequality as follows. With $A_\Delta = A_{\text{cl}}(\lambda(\theta^-), \theta^+)$,

$$\begin{aligned} \sigma_w^2 \sum_{j=0}^{h-1} (e_1' A_\Delta^j e_1)^2 &= J_y(\lambda(\theta^-); \theta^+) - \sigma_w^2 \sum_{j=h}^{\infty} (e_1' A_\Delta^j e_1)^2 \\ &\geq \sigma_w^2 + \Delta - \sigma_w^2 \sum_{j=h}^{\infty} \|A_\Delta^j\|^2 \\ &\geq \sigma_w^2 + \Delta - \frac{\sigma_w^2 C_\Delta^2 \rho_\Delta^{2h}}{1 - \rho_\Delta^2} \\ &\geq \sigma_w^2 + \frac{7\Delta}{8}. \end{aligned}$$

Combining this bound with the nonnegative conditional-state term in (71) proves

$$\mathbb{E}[(y_s^{\text{aux}})^2 \mid \mathcal{F}_{s-h}] \geq \sigma_w^2 + \frac{7\Delta}{8}. \quad (72)$$

Lemma 17 (One-change detection). *Fix a potential restart r and a change time τ with $r+H \leq \tau$ and $\tau+D \leq T$. Let $\mathcal{A}_{r,\tau} \in \mathcal{F}_\tau$ be the event that the cycle at r is reached, θ^- is active throughout $[r, \tau)$, $\mathcal{G}_r^{\text{est}}$ holds, and the cycle remains in exploitation after observing y_τ . Suppose θ^+ is active throughout $[\tau, \tau+D)$. Let $\hat{\tau}_r$ be this cycle's first alarm output time strictly after τ , with value $+\infty$ if none occurs. Then, almost surely on $\mathcal{A}_{r,\tau}$,*

$$\mathbb{P}(\hat{\tau}_r > \tau + D \mid \mathcal{F}_\tau) \leq \frac{\delta_d}{4T^2}. \quad (73)$$

Proof. Condition on an arbitrary $\mathcal{F}_\tau$ history in $\mathcal{A}_{r,\tau}$. In the continuation in which alarms are suppressed, the restart-relative sampling grid, the auxiliary initial state ψ_τ , and the old reference are now fixed, while all future disturbances remain fresh. Within fewer than $2h$ steps, the grid has an output time s_1 with $s_1 - h \geq \tau$. The next $m-1$ grid times are h steps apart, and the last satisfies

$$s_m - \tau < 2h + (m-1)h = (m+1)h = D.$$

Thus the complete block exists within the finite horizon. If any earlier block alarms, the claim is already true. Otherwise the actual and alarm-suppressed paths agree through this complete post-change block.

The conditional version of Lemma 14 is valid because the future innovations are fresh after $\mathcal{F}_\tau$ and $\mathcal{F}_\tau \subseteq \mathcal{F}_{s_1-h}$ by construction. It, together with (72) and $a = \Delta/8$, gives

$$\mathbb{P}\left(\frac{1}{m} \sum_{k=1}^m (y_{s_k}^{\text{aux}})^2 < \sigma_w^2 + \frac{3\Delta}{4}\right) \leq \exp\left(-\frac{m\Delta^2}{128B_o^4}\right) \leq \frac{\delta_d}{4T^2}.$$

Outside this deviation, Lemma 16 applies because no second change occurs before the tested block and gives an actual average at least $\sigma_w^2 + 5\Delta/8$, which crosses the detector threshold. This proves (73). Equivalently, for every $A \in \mathcal{F}_\tau$ with $A \subseteq \mathcal{A}_{r,\tau}$,

$$\mathbb{P}(A \cap \{\hat{\tau}_r > \tau + D\}) \leq \frac{\delta_d}{4T^2} \mathbb{P}(A),$$

which is the eventwise form used below. $\square$

C.4 Alarm Correspondence

Theorem 18 (Alarm correspondence and delay). *Suppose Assumptions 7–12 hold. Let σ_{det} be the first output time at which a false alarm occurs or the first deadline at which a true change remains unmatched, with $\inf \emptyset = +\infty$. For each potential r , let $\mathcal{E}_r^{\text{reach}}$ be the event that r is reached, $r < \sigma_{\text{det}}$, and $[r, r + H)$ is a full, unchanged exploration block. Define*

$$\mathcal{E}_{\text{ref}} := \bigcap_{r=0}^T \{(\mathcal{E}_r^{\text{reach}})^c \cup \mathcal{G}_r^{\text{est}}\}. \quad (74)$$

Let $\mathcal{E}_{\text{det}}$ be the event that PIECE-CD has no false alarms and raises exactly one alarm $\hat{\tau}^{(c)}$ for each true change, with

$$\tau^{(c)} < \hat{\tau}^{(c)} \leq \tau^{(c)} + D, \quad c = 1, \dots, C. \quad (75)$$

Then

$$\mathbb{P}(\mathcal{E}_{\text{ref}} \cap \mathcal{E}_{\text{det}}^c) \leq \delta_d. \quad (76)$$

Proof. There are at most T potential restart locations and at most T potential test endpoints after each location. Fix such a pair. Intersect its crossing event with the events that the restart is reached, its reference event holds, no earlier detector failure has occurred, and every control generating an output in the block uses the same parameter as the reference exploration; that is, the block is null for this restart. Conditional on the history at the end of exploration, apply Lemma 15 to the fixed predictable block in the continuation with alarms suppressed. If the actual test is reached, the continuation and actual trajectory agree through that test. Thus the probability of this *joint gated event*, rather than a probability conditional on survival to the test, is at most $\delta_d/(2T^2)$. A union bound over the restart–endpoint pairs gives a total false-alarm contribution at most $\delta_d/2$.

For a true change $\tau^{(c)}$, intersect missed detection with the successful history up to that change. For each potential r , the event that the past detector history is successful, the cycle at r is active at $\tau^{(c)}$, and its reference succeeds is an $\mathcal{F}_{\tau^{(c)}}$ -measurable subset of $\mathcal{A}_{r, \tau^{(c)}}$ from Lemma 17. These events are disjoint over r . The dwell conditions give $r + H \leq \tau^{(c)}$, $\tau^{(c)} + D \leq T$, and no second change in $[\tau^{(c)}, \tau^{(c)} + D)$. Multiplying (73) by each active-cycle indicator, taking expectations, and summing over the disjoint partition gives $\delta_d/(4T^2)$ for that change, without an additional restart factor. There are at most $C \leq T - 1$ changes, so the total missed-delay contribution is at most $C\delta_d/(4T^2) \leq (T - 1)\delta_d/(4T^2) < \delta_d/4$. A block containing outputs from both sides of a true change is not a null block; if it crosses, it is a valid earlier detection. On the complement of these gated deviations, induction over the change index gives one-to-one alarm correspondence. The deterministic calendar calculation in Appendix H shows that the complete post-change block ends strictly before $\tau^{(c)} + D$; no probability is spent on boundary or integer-rounding conventions. The dwell condition ensures that each alarm occurs before the next change and, when another change remains, leaves an uncontaminated H -step exploration block for the next cycle. The false-alarm contribution is at most $\delta_d/2$ and the missed-delay contribution is less than $\delta_d/4$, so their sum is below δ_d . This proves (76). $\square$

D REGRET ANALYSIS

On the detection event of Theorem 18, the horizon splits into exploration, detection-delay, and unchanged exploitation intervals. The first two have bounded per-round cost; the RLS potential argument controls the cumulative cost of the third.

D.1 Segmentwise Exploitation Regret

Consider an unchanged exploitation interval following a restart at r . Suppress the restart subscript and write $b_\star = b_1(\theta)$, $\lambda = \lambda(\theta)$, $\hat{b} = \hat{b}_r$, and $\bar{\lambda} = \bar{\lambda}_r$. Define

$$\begin{aligned} \alpha &:= 1 - \frac{b_\star}{\hat{b}}, & \bar{\alpha} &:= \frac{E_\theta(H, \delta_0)}{b_{\min}}, \\ \underline{v}_H &:= \nu + \kappa_G n_H, & \bar{\ell} &:= \frac{B_\psi^2}{\underline{v}_H}, \\ c_0 &:= 1 - 2\bar{\alpha} - 2\bar{\ell}(1 + \bar{\alpha})^2, & \bar{Q}_H &:= E_\lambda(H, \delta_0)^2(\nu + n_H B_\psi^2), \\ B_\eta &:= B_w/b_{\min}, & & \\ \Gamma(n) &:= d \log \left(1 + \frac{n B_\psi^2}{d \underline{v}_H} \right). \end{aligned} \tag{77}$$

The tuning condition (51) ensures $c_0 > 0.7$.

Theorem 19 (Logarithmic exploitation regret). *Suppose a single parameter θ is active throughout the full exploration block $[r, r+H)$ and throughout an exploitation interval $[r+H, r+H+N)$ of that same cycle, ending no later than its next alarm. Suppose the Gram and estimation events in Propositions 21 and 22 hold for this same θ . Conditional on the history at the end of exploration, with probability at least $1 - \delta_0$, simultaneously for every prefix length $0 \leq n \leq N$,*

$$\sum_{t=r+H}^{r+H+n-1} (y_{t+1} - w_{t+1})^2 \leq R_E(n, \delta_0), \tag{78}$$

$$R_E(n, \delta_0) := \bar{b}^2 \left[\frac{2\bar{Q}_H}{c_0} + \frac{8B_\eta^2}{c_0^2} \log \frac{1}{\delta_0} + \frac{4B_\eta^2}{c_0} \Gamma(n) \right].$$

Proof. We give the core argument; Appendix G records the matrix identities and conditioning details. Let $d_t = \hat{\lambda}_t - \lambda$, $a_t = d_t' \psi_t$, and $e_t = u_t - \lambda' \psi_t$. Lemma 13 gives $e_t^2 \leq a_t^2$. From (36), the pseudo-response in (59) obeys

$$x_{t+1}^\lambda = \lambda' \psi_t + \alpha e_t + \eta_{t+1}, \quad \eta_{t+1} := -w_{t+1}/\hat{b}.$$

Here η_{t+1} is conditionally mean zero and B_η -sub-Gaussian.

Let $Q_t = d_t' V_t d_t$ and $\ell_t = \psi_t' V_{t+1}^{-1} \psi_t$. To display the algebra, set $\omega_t := \alpha e_t + \eta_{t+1} - a_t$. The RLS update and $V_{t+1} = V_t + \psi_t \psi_t'$ give

$$\begin{aligned} d_{t+1} &= d_t + V_{t+1}^{-1} \psi_t \omega_t, \\ Q_{t+1} &= d_{t+1}' V_{t+1} d_{t+1} + 2\omega_t d_t' \psi_t + \omega_t^2 \psi_t' V_{t+1}^{-1} \psi_t \\ &= Q_t + a_t^2 + 2a_t \omega_t + \ell_t \omega_t^2. \end{aligned}$$

Substituting the definition of ω_t and collecting $a_t^2 - 2a_t^2 = -a_t^2$ gives the exact identity

$$Q_{t+1} - Q_t = -a_t^2 + 2a_t(\alpha e_t + \eta_{t+1}) + \ell_t(\alpha e_t + \eta_{t+1} - a_t)^2. \tag{79}$$

On the exploration event, $|\alpha| \leq \bar{\alpha}$ and $\ell_t \leq \bar{\ell}$. Using $|e_t| \leq |a_t|$ in (79) yields

$$Q_{t+1} - Q_t \leq -c_0 a_t^2 + 2a_t \eta_{t+1} + 2\ell_t \eta_{t+1}^2. \tag{80}$$

For completeness, conditional B_η -sub-Gaussianity and predictability of a_t imply that, for every fixed $\xi > 0$,

$$M_n := \exp \left\{ \xi \sum_{t=r+H}^{r+H+n-1} a_t \eta_{t+1} - \frac{\xi^2 B_\eta^2}{2} \sum_{t=r+H}^{r+H+n-1} a_t^2 \right\}$$

is a nonnegative supermartingale starting at one. Ville's inequality (Ville, 1939; Howard et al., 2020) therefore gives, with probability at least $1 - \delta_0$, simultaneously over every prefix,

$$\sum_t a_t \eta_{t+1} \leq \frac{\xi B_\eta^2}{2} \sum_t a_t^2 + \frac{1}{\xi} \log \frac{1}{\delta_0}.$$

Taking the deterministic value $\xi = c_0/(2B_\eta^2)$ yields

$$\sum_t a_t \eta_{t+1} \leq \frac{c_0}{4} \sum_t a_t^2 + \frac{2B_\eta^2}{c_0} \log \frac{1}{\delta_0}. \quad (81)$$

Appendix G gives the same argument directly from the one-step conditional moment-generating-function inequality. Moreover, the matrix determinant lemma gives

$$\sum_t \ell_t \leq \log \frac{\det V_{r+H+n}}{\det V_{r+H}} \leq \Gamma(n). \quad (82)$$

Summing (80), using $Q_t \geq 0$, $\eta_{t+1}^2 \leq B_\eta^2$, and (81)–(82), gives

$$\sum_t a_t^2 \leq \frac{2Q_{r+H}}{c_0} + \frac{8B_\eta^2}{c_0^2} \log \frac{1}{\delta_0} + \frac{4B_\eta^2}{c_0} \Gamma(n). \quad (83)$$

Finally, $Q_{r+H} \leq \bar{Q}_H$ and, by (36), $(y_{t+1} - w_{t+1})^2 = b_\star^2 e_t^2 \leq \bar{b}^2 a_t^2$. This proves (78). $\square$

D.2 Delay and Exploration Costs

The deterministic state bound gives the per-step delay cost

$$K_D := (B_Y + B_w)^2. \quad (84)$$

During exploration, $|u_t| \leq B_e$ and $\|\psi_t\| \leq B_\psi$, so (36) gives the per-step bound

$$K_X := \bar{b}^2 (B_e + \Lambda B_\psi)^2. \quad (85)$$

Thus a delay interval of length at most D costs at most $K_D D$, and an exploration block costs at most $K_X H$.

D.3 Total Regret

Theorem 20 (Finite-time regret of PIECE-CD). *Let Assumptions 7–12 hold. Let H be the smallest integer greater than r_0 that satisfies (93) and (51)–(52); choose h, m, D by (53)–(54), and run Algorithm 2. Then, with probability at least $1 - \delta$, there are no false alarms, every change has exactly one alarm satisfying (75), and*

$$\mathcal{R}_T \leq CK_D D + (C + 1)K_X H + (C + 1)R_E(T, \delta_0). \quad (86)$$

Consequently, for fixed orders and fixed model, noise, initialization, exploration, ridge, stability, clipping, and gap bounds,

$$\mathcal{R}_T = O\left((C + 1) \log \frac{T + 1}{\delta}\right). \quad (87)$$

Proof. If $T \leq H$, Assumption 12 forces $C = 0$. The algorithm then performs only truncated exploration, and the deterministic exploration bound gives $\mathcal{R}_T \leq K_X T \leq K_X H$, which is bounded by the right-hand side of (86). Hence assume $T > H$ below.

At every potential full, unchanged restart block, gate the estimation and RLS failure events by the event that this restart is reached with a successful history, and then condition on the past. Propositions 21 and 22 cost at most $2\delta_0$, and the time-uniform RLS event in Theorem 19 costs another δ_0 . The tower property and a union bound over at most $T + 1$ locations give $3(T + 1)\delta_0 = 3\delta/32$. Gating the false-alarm and delay events in the same chronological way, Theorem 18 costs less than $\delta_d = \delta/2$. Their intersection therefore has probability greater than $1 - \delta$. Appendix I gives the full induction; in particular, it never conditions an early detector event on future estimator success.

On this intersection, each of the C true changes has a pre-alarm interval of length at most D , giving $CK_D D$. There is one initial exploration block and at most one block after each alarm, giving at most $(C + 1)K_X H$. Removing these intervals leaves at most $C + 1$ unchanged exploitation prefixes. Apply Theorem 19 to each and upper-bound its length by T ; this gives the last term in (86).

Appendix J constructs a feasible $H = O(\log((T + 1)/\delta))$; the smallest feasible value is therefore no larger. Equations (53)–(54) give fixed h and $D = O(\log((T + 1)/\delta))$. Finally, $\bar{Q}_H$, $\Gamma(T)$, and $\log(1/\delta_0)$ are all $O(\log((T + 1)/\delta))$. Substitution proves (87). When $C = 0$, the delay term is absent, leaving the stationary logarithmic bound rather than the meaningless expression $O(C \log T)$. $\square$

E DETERMINISTIC BOUNDS AND RESTART CONVENTIONS

Bounded applied inputs and Assumptions 7–9 yield the pathwise bounds (40).

E.1 Output and Regressor Bounds

Let $Y_t = (y_t, \dots, y_{t-p+1})'$. Equation (32) has the state form

$$Y_{t+1} = A_y(\theta_t)Y_t + e_1\zeta_{t+1}, \quad \zeta_{t+1} := w_{t+1} + \sum_{j=1}^q b_j(\theta_t)u_{t-j+1}. \quad (88)$$

Since $|u_t| \leq B_u$ by construction,

$$|\zeta_{t+1}| \leq B_w + B_b B_u.$$

Unrolling (88) and applying (39) gives

$$\begin{aligned} \|Y_t\| &\leq C_y \rho_y^t \|Y_0\| + C_y (B_w + B_b B_u) \sum_{j=0}^{t-1} \rho_y^j \\ &\leq C_y \|Y_0\| + \frac{C_y}{1 - \rho_y} (B_w + B_b B_u) = B_Y. \end{aligned}$$

Every output coordinate is therefore bounded by B_Y . The definition (33) and clipping give

$$\|\psi_t\|^2 \leq pB_Y^2 + (q - 1)B_u^2 = B_\psi^2,$$

and (34) gives $\|\phi_t\| \leq B_\phi$.

E.2 Projection onto the Clipping Interval

Let $\text{clip}_{[-B, B]}(x) = (-B) \vee x \wedge B$. As the Euclidean projection onto a closed convex set, it is nonexpansive:

$$|\text{clip}_{[-B, B]}(x) - \text{clip}_{[-B, B]}(z)| \leq |x - z|.$$

If $z \in [-B, B]$, this becomes

$$|\text{clip}_{[-B, B]}(x) - z| \leq |x - z|. \quad (89)$$

Assumption 10 permits (89) with $z = \lambda(\theta)' \psi_t$ for $\theta \in \Theta$. For an arbitrary target z , the bound includes its distance from the clipping interval:

$$|\text{clip}_{[-B_u, B_u]}(x) - z| \leq |x - z| + (|z| - B_u)_+.$$

E.3 Auxiliary-Process Bounds

For the fixed mismatch matrix A_Δ in (68), (44) gives, for $t \geq \tau$,

$$\begin{aligned} \|\psi_t^{\text{aux}}\| &\leq C_\Delta \rho_\Delta^{t-\tau} \|\psi_\tau\| + C_\Delta B_w \sum_{j=0}^{t-\tau-1} \rho_\Delta^j \\ &\leq C_\Delta B_\psi + \frac{C_\Delta B_w}{1 - \rho_\Delta} = B_{\text{aux}}. \end{aligned}$$

Thus both actual and auxiliary outputs are bounded in absolute value by B_o .

E.4 Conditioning at a Random Restart

An alarm time r is an $\{\mathcal{F}_t\}$ -stopping time. Conditional on $\mathcal{F}_r$, the state ψ_r is fixed with $\|\psi_r\| \leq B_\psi$, and future disturbances retain the independence, moments, and bound in Assumption 9. The input u_r is already $\mathcal{F}_r$ -measurable; subsequent exploratory inputs are fresh and independent of the disturbances.

Consider a conditional concentration bound that holds uniformly over these restart histories, with failure probability at most $\varepsilon_{\text{fail}}$. For any eligibility event $\mathcal{C}_r \in \mathcal{F}_r$, the tower property gives

$$\begin{aligned} \mathbb{P}(\mathcal{C}_r \cap \{\text{failure at } r\}) &= \mathbb{E}[\mathbf{1}_{\mathcal{C}_r} \mathbb{P}(\text{failure at } r \mid \mathcal{F}_r)] \\ &\leq \varepsilon_{\text{fail}} \mathbb{P}(\mathcal{C}_r) \leq \varepsilon_{\text{fail}}. \end{aligned}$$

A union bound over at most $T + 1$ possible restart locations then controls these failures across the horizon. This is the basis for the allocation (48).

F EXPLORATION, EXCITATION, AND THE RESTARTED ESTIMATE

The covariance analysis of stationary PIECE in Singh et al. (2024) concerns sparse, recurring exploration episodes. Here we establish linear Gram growth and estimation bounds for a contiguous exploration block after a random restart.

F.1 Conditional Excitation

Fix a complete exploration block $[r, r + H)$ during which the parameter remains equal to θ . Put $s := p + q$, $L := r_0 = \max\{p, q\}$, and $\underline{\sigma}^2 := \min\{\sigma_w^2, \sigma_e^2\}$. At any exploration time $t \geq r + L$, collect the most recent innovations in

$$Z_t := (w_{t-p+1}, \dots, w_t, u_{t-q+1}, \dots, u_t)'$$

Because the parameter is fixed within the exploration block, repeated substitution in the ARX recursion gives

$$\phi_t = M(\theta)Z_t + \xi_t. \quad (90)$$

To make the remainder explicit, let $\bar{Z}_t$ collect the members of $(w_{t-L+1}, \dots, w_t, u_{t-L+1}, \dots, u_t)$ that do not already occur in Z_t . Repeated substitution gives $\xi_t = c_t + \bar{M}(\theta)\bar{Z}_t$, where c_t is $\mathcal{F}_{t-L}$ -measurable. Conditional on $\mathcal{F}_{t-L}$, the vector Z_t is centered, is independent of $\bar{Z}_t$, and has covariance

$$\text{Cov}(Z_t | \mathcal{F}_{t-L}) = \text{diag}(\sigma_w^2 I_p, \sigma_e^2 I_q).$$

Thus Z_t and ξ_t are conditionally independent. The remainder depends on the $\mathcal{F}_{t-L}$ -measurable history and the omitted innovations.

To see the triangular structure explicitly, order the output coordinates and their disturbances chronologically as $(y_{t-p+1}, \dots, y_t)$ and $(w_{t-p+1}, \dots, w_t)$. The equation for y_j contains w_j with coefficient one and only disturbances dated at or before j . Thus the output–disturbance block is triangular with unit diagonal. Ordering the current exploration inputs as $(u_{t-q+1}, \dots, u_t)$ makes the input-coordinate block the identity. All remaining post- $\mathcal{F}_{t-L}$ innovations are precisely those collected in $\bar{Z}_t$; older terms belong to c_t . Hence, up to these fixed row and column permutations, $M(\theta)$ has the block-triangular form

$$M(\theta) = \begin{bmatrix} T_w(\theta) & T_u(\theta) \\ 0 & I_q \end{bmatrix},$$

where $T_w(\theta)$ is triangular with unit diagonal: each y_j contains its contemporaneous disturbance w_j with coefficient one. Hence $M(\theta)$ is nonsingular. Continuity and compactness give

$$\chi := \inf_{\theta \in \Theta} \sigma_{\min}(M(\theta)) > 0, \quad \kappa_{\text{inn}} := \underline{\sigma}^2 \chi^2 > 0.$$

For completeness, let $\Sigma_Z := \text{Cov}(Z_t | \mathcal{F}_{t-L})$. Conditional independence removes both cross-covariance terms, and hence

$$\begin{aligned} \text{Cov}(\phi_t | \mathcal{F}_{t-L}) &= M(\theta)\Sigma_Z M(\theta)' + \text{Cov}(\xi_t | \mathcal{F}_{t-L}) \\ &\succeq \underline{\sigma}^2 M(\theta)M(\theta)' \succeq \underline{\sigma}^2 \sigma_{\min}(M(\theta))^2 I_s. \end{aligned}$$

Also, $\mathbb{E}[\phi_t \phi_t' | \mathcal{F}_{t-L}]$ equals this covariance plus the positive semidefinite outer product of the conditional mean. Therefore

$$\mathbb{E}[\phi_t \phi_t' | \mathcal{F}_{t-L}] \succeq \text{Cov}(\phi_t | \mathcal{F}_{t-L}) \succeq \kappa_{\text{inn}} I_s. \quad (91)$$

The bound is uniform over the history in $\mathcal{F}_{t-L}$.

F.2 Exploration Gram Bound

Fix a restart r and the trimmed set $\mathcal{I}_r = \{r + r_0, \dots, r + H - 1\}$. Let

$$\Phi_r := \begin{bmatrix} \phi'_{r+r_0} \\ \vdots \\ \phi'_{r+H-1} \end{bmatrix}, \quad G_r := \Phi_r' \Phi_r = \sum_{t \in \mathcal{I}_r} \phi_t \phi_t'.$$

Set

$$N_H := 1 + \left\lfloor \frac{n_H - 1}{L} \right\rfloor, \quad \kappa_G := \frac{\kappa_{\text{inn}}}{2L}. \quad (92)$$

Consider only the L -spaced rows $t_k = r + L + (k - 1)L$, $k = 1, \dots, N_H$. They are contained in $\mathcal{I}_r$, and $N_H \geq n_H/L$.

The conditioning argument in Appendix E applies at a stopping-time restart r . Although u_r is already $\mathcal{F}_r$ -measurable, the trimming starts at $t = r + L$, so every innovation appearing in Z_t has index at least $r + 1$. Consequently, (90)–(96) apply without treating r as deterministic.

Proposition 21 (Exploration Gram bound). *Let r be a stopping time with $r + H \leq T$, and suppose a single $\theta \in \Theta$ is active at every control time $t = r, \dots, r + H - 1$. If*

$$N_H \geq \frac{10B_\phi^4}{\kappa_{\text{inn}}^2} \log \frac{s}{\delta_0}, \quad (93)$$

then, conditional on $\mathcal{F}_r$,

$$\mathbb{P}\left(\lambda_{\min}(G_r) \geq \kappa_G n_H, \quad \text{tr}(G_r) \leq n_H B_\phi^2 \mid \mathcal{F}_r\right) \geq 1 - \delta_0. \quad (94)$$

The same event implies

$$\lambda_{\min}\left(\sum_{t \in \mathcal{I}_r} \psi_t \psi_t'\right) \geq \kappa_G n_H. \quad (95)$$

Proof. Condition on $\mathcal{F}_r$. For $X_k := \phi_{t_k} \phi_{t_k}'$, use the stopped filtration $\mathcal{H}_0 = \mathcal{F}_r$ and $\mathcal{H}_k = \mathcal{F}_{t_k} = \mathcal{F}_{r+kL}$. Here $t_1 - L = r$, and $t_k - L = t_{k-1}$ for $k \geq 2$, so (91) gives

$$A_k := \mathbb{E}[X_k \mid \mathcal{H}_{k-1}] \succeq \kappa_{\text{inn}} I_s.$$

Also $0 \preceq X_k \preceq B_\phi^2 I_s$. Thus $D_k := X_k - A_k$ is a self-adjoint matrix-martingale difference with $\|D_k\| \leq B_\phi^2$ and predictable quadratic variation bounded by $N_H B_\phi^4 I_s$. Moreover, $\kappa_{\text{inn}} \leq B_\phi^2$ because $A_k \preceq B_\phi^2 I_s$. The lower-tail form of the matrix Freedman inequality (Tropp, 2011), applied to $-\sum_k D_k$, gives

$$\begin{aligned} \mathbb{P}\left(\lambda_{\min}\left(\sum_{k=1}^{N_H} X_k\right) \leq \frac{\kappa_{\text{inn}} N_H}{2} \mid \mathcal{F}_r\right) &\leq s \exp\left\{-\frac{\kappa_{\text{inn}}^2 N_H}{8B_\phi^4 + (4/3)B_\phi^2 \kappa_{\text{inn}}}\right\} \\ &\leq s \exp\left(-\frac{\kappa_{\text{inn}}^2 N_H}{10B_\phi^4}\right). \end{aligned} \quad (96)$$

Condition (93) makes the right-hand side at most δ_0 . Since $G_r \succeq \sum_k X_k$ and $N_H \geq n_H/L$, the claimed eigenvalue bound follows. The trace bound is deterministic because $\|\phi_t\| \leq B_\phi$.

Finally, $\sum_{t \in \mathcal{I}_r} \psi_t \psi_t'$ is the principal block of G_r obtained by deleting the current-input row and column. Eigenvalue interlacing, or direct restriction of the Rayleigh quotient, proves (95). $\square$

F.3 Projected Ridge Estimation

Proposition 22 (Restarted parameter and gain estimate). *Let r be a stopping time with $r + H \leq T$, and suppose a single $\theta \in \Theta$ is active at every control time $t = r, \dots, r + H - 1$. Given $\mathcal{F}_r$,*

event (100) has probability at least $1 - \delta_0$. On its intersection with the event in Proposition 21, the estimator defined in (55)–(56) satisfies

$$\begin{aligned}\|\hat{\theta}_r - \theta\| &\leq E_\theta(H, \delta_0), \\ \|\bar{\lambda}_r - \lambda(\theta)\| &\leq E_\lambda(H, \delta_0) = g_H/4.\end{aligned}\tag{97}$$

Moreover, the RLS initialization satisfies

$$\underline{v}_H I_d \preceq V_{r+H} \preceq (\nu + n_H B_\psi^2) I_d\tag{98}$$

in the sense of its smallest- and largest-eigenvalue bounds. If (93) holds, the Gram event and the two estimates above therefore hold jointly with conditional probability at least $1 - 2\delta_0$.

Proof. Let $V = V_r^\theta$ and $S_r = \sum_{t \in \mathcal{I}_r} \phi_t w_{t+1}$. Since $y_{t+1} = \phi_t' \theta + w_{t+1}$ during the uncontaminated block, write $G_r = \sum_{t \in \mathcal{I}_r} \phi_t \phi_t'$ and calculate

$$\begin{aligned}\sum_{t \in \mathcal{I}_r} \phi_t y_{t+1} &= \sum_{t \in \mathcal{I}_r} \phi_t (\phi_t' \theta + w_{t+1}) \\ &= G_r \theta + S_r = (V - \nu I_s) \theta + S_r.\end{aligned}$$

Multiplication by V^{-1} gives

$$\tilde{\theta}_r = \theta + V^{-1}(S_r - \nu \theta),$$

and therefore

$$\tilde{\theta}_r - \theta = V^{-1}(S_r - \nu \theta).\tag{99}$$

Assumption 9 and Hoeffding's lemma make w_{t+1} conditionally B_w -sub-Gaussian. The fixed-ridge self-normalized inequality (Abbasi-Yadkori et al., 2011, Theorem 1) gives, with conditional probability at least $1 - \delta_0$,

$$\|S_r\|_{V^{-1}} \leq B_w \sqrt{\log \frac{\det V}{\nu^s \delta_0^2}}.\tag{100}$$

Here again $s = p + q$, and $\nu^s = \det(\nu I_s)$ is the determinant of the initial ridge matrix. Also $\|\nu \theta\|_{V^{-1}} \leq \sqrt{\nu} S_\Theta$. By the arithmetic-geometric mean inequality and the trace bound, the determinant step can be written explicitly. Let $\gamma_1, \dots, \gamma_s \geq 0$ be the eigenvalues of G_r . Since $V = \nu I_s + G_r$,

$$\begin{aligned}\frac{\det V}{\nu^s} &= \prod_{i=1}^s \left(1 + \frac{\gamma_i}{\nu}\right) \\ &\leq \left(\frac{1}{s} \sum_{i=1}^s \left(1 + \frac{\gamma_i}{\nu}\right)\right)^s = \left(1 + \frac{\text{tr}(G_r)}{s\nu}\right)^s \\ &\leq \left(1 + \frac{n_H B_\phi^2}{s\nu}\right)^s.\end{aligned}$$

The last inequality uses $\text{tr}(G_r) = \sum_{t \in \mathcal{I}_r} \|\phi_t\|^2 \leq n_H B_\phi^2$. Taking logarithms and using $n_H \leq H$ yields

$$\log \frac{\det V}{\nu^s} \leq s \log \left(1 + \frac{n_H B_\phi^2}{s\nu}\right) \leq s \log \left(1 + \frac{H B_\phi^2}{s\nu}\right).\tag{101}$$

Equations (94), (99)–(101) imply the unprojected Euclidean error bound equal to one half of (46). A nearest-point projection onto a compact set containing θ increases Euclidean error by at most a factor two, proving (97). If Θ is closed and convex, projection is nonexpansive and the factor two may be removed.

The gain bound follows from (38). The lower eigenvalue in (98) follows from (95); the upper eigenvalue follows from $\sum \psi_t \psi_t' \preceq (\sum \|\psi_t\|^2) I_d$. $\square$

The linear Gram bound (94) is the excitation estimate used in Appendix J to obtain $H = O(\log((T+1)/\delta))$.

G THE RLS POTENTIAL ARGUMENT

We establish the potential estimates used in Theorem 19. Consider its unchanged exploration block and exploitation interval, and condition on an $\mathcal{F}_{r+H}$ history on which the Gram and estimation events hold. We use the constants in (77), set $t_0 = r + H$, and let $t_1 \geq t_0$ denote the final control time of an exploitation prefix.

G.1 Control Error and Prediction Error

Let λ be the true gain on the unchanged segment and let $\bar{\lambda}$ be the exploration reference. For a gate of radius $g > 0$, the recursive proposal is accepted when $|(\hat{\lambda}_t - \bar{\lambda})' \psi_t| \leq g \|\psi_t\|$ and the reference proposal is used otherwise. Suppose

$$\|\bar{\lambda} - \lambda\| \leq \epsilon, \quad g > \epsilon.$$

For $a_t = (\hat{\lambda}_t - \lambda)' \psi_t$, the unclipped gated action error $\tilde{e}_t = z_t - \lambda' \psi_t$ obeys

$$\tilde{e}_t^2 \leq K_g a_t^2, \quad K_g := \max \left\{ 1, \left(\frac{\epsilon}{g - \epsilon} \right)^2 \right\}. \quad (102)$$

If $\psi_t = 0$, then $a_t = \tilde{e}_t = 0$ and the claim is immediate. If $\psi_t \neq 0$, then on the accepted branch $\tilde{e}_t = a_t$. On the fallback branch,

$$\begin{aligned} |a_t| &\geq |(\hat{\lambda}_t - \bar{\lambda})' \psi_t| - |(\bar{\lambda} - \lambda)' \psi_t| \\ &> (g - \epsilon) \|\psi_t\|, \end{aligned}$$

while $|\tilde{e}_t| \leq \epsilon \|\psi_t\|$. Since $g - \epsilon > 0$, division gives

$$|\tilde{e}_t| \leq \frac{\epsilon}{g - \epsilon} |a_t|.$$

Squaring this inequality and taking the larger constant over the two branches proves (102). If the ideal action lies in the clipping interval, nonexpansiveness preserves (102). In the algorithm, $g = g_H$ and $\epsilon = E_\lambda = g_H/4$. Therefore $\epsilon/(g - \epsilon) = 1/3$ and $K_g = 1$.

G.2 Exact Regression and Potential Identity

On the segment,

$$y_{t+1} = b_\star(u_t - \lambda' \psi_t) + w_{t+1}. \quad (103)$$

Let $\gamma = b_\star/\hat{b}$, $\alpha = 1 - \gamma$, and $e_t = u_t - \lambda' \psi_t$. The pseudo-response is exactly

$$\begin{aligned} x_{t+1}^\lambda &= u_t - y_{t+1}/\hat{b} \\ &= \lambda' \psi_t + \alpha e_t + \eta_{t+1}, \end{aligned} \quad \eta_{t+1} = -w_{t+1}/\hat{b}.$$

Because $\widehat{\theta}_r \in \Theta$ by projection and hence $\widehat{b} \geq b_{\min}$ by (37), $|\eta_{t+1}| \leq B_\eta = B_w/b_{\min}$ and Hoeffding's lemma gives

$$\mathbb{E}[\exp(\xi\eta_{t+1}) \mid \mathcal{F}_t] \leq \exp(\xi^2 B_\eta^2/2), \quad \xi \in \mathbb{R}. \quad (104)$$

Set $d_t = \widehat{\lambda}_t - \lambda$, $a_t = d'_t \psi_t$, $Q_t = d'_t V_t d_t$, and

$$\ell_t := \psi'_t V_{t+1}^{-1} \psi_t.$$

The update (60) becomes

$$d_{t+1} = d_t + V_{t+1}^{-1} \psi_t (\alpha e_t + \eta_{t+1} - a_t). \quad (105)$$

Using $V_{t+1} = V_t + \psi_t \psi'_t$ in (105) and expanding gives

$$Q_{t+1} - Q_t = -a_t^2 + 2a_t(\alpha e_t + \eta_{t+1}) + \ell_t(\alpha e_t + \eta_{t+1} - a_t)^2. \quad (106)$$

On the exploration event,

$$|\alpha| = \frac{|\widehat{b} - b_\star|}{\widehat{b}} \leq \frac{E_\theta}{b_{\min}} = \bar{\alpha}.$$

Also $V_{t+1} \succeq V_{r+H} \succeq \underline{v}_H I$, so

$$0 \leq \ell_t \leq \frac{\|\psi_t\|^2}{\underline{v}_H} \leq \bar{\ell}.$$

By $e_t^2 \leq a_t^2$ and $(x+y)^2 \leq 2x^2 + 2y^2$,

$$\begin{aligned} 2\alpha a_t e_t &\leq 2\bar{\alpha} a_t^2, \\ \ell_t(\alpha e_t + \eta_{t+1} - a_t)^2 &\leq 2\ell_t(1 + \bar{\alpha})^2 a_t^2 + 2\ell_t \eta_{t+1}^2. \end{aligned}$$

Substitution into (106) proves (80).

G.3 Time-Uniform Noise Bound

For any fixed $\xi > 0$, (104) and predictability of a_t imply that the prefix-indexed process

$$\exp \left\{ \xi \sum_{t=t_0}^{t_1} a_t \eta_{t+1} - \frac{\xi^2 B_\eta^2}{2} \sum_{t=t_0}^{t_1} a_t^2 \right\}$$

is a nonnegative supermartingale starting at one, conditionally on $\mathcal{F}_{t_0}$. Its one-step multiplicative increment satisfies

$$\mathbb{E} \left[\exp \left\{ \xi a_t \eta_{t+1} - \frac{\xi^2 B_\eta^2 a_t^2}{2} \right\} \middle| \mathcal{F}_t \right] \leq 1,$$

because a_t is $\mathcal{F}_t$ -measurable and (104) applies with argument ξa_t .

Ville's inequality (Ville, 1939; Howard et al., 2020) says that, with probability at least $1 - \delta_0$, this process never exceeds $1/\delta_0$. Taking logarithms on that event and dividing by the fixed $\xi > 0$ gives, simultaneously for every exploitation prefix,

$$\sum_t a_t \eta_{t+1} \leq \frac{\xi B_\eta^2}{2} \sum_t a_t^2 + \frac{1}{\xi} \log \frac{1}{\delta_0}.$$

Choosing the deterministic value $\xi = c_0/(2B_\eta^2)$ gives

$$\sum_t a_t \eta_{t+1} \leq \frac{c_0}{4} \sum_t a_t^2 + \frac{2B_\eta^2}{c_0} \log(1/\delta_0). \quad (107)$$

G.4 Elliptical Potential

For $x_t = \psi_t' V_t^{-1} \psi_t$, the Sherman–Morrison identity $V_{t+1}^{-1} = V_t^{-1} - V_t^{-1} \psi_t \psi_t' V_t^{-1} / (1 + x_t)$ and the matrix determinant lemma give

$$\ell_t = \frac{x_t}{1 + x_t}, \quad \frac{\det V_{t+1}}{\det V_t} = 1 + x_t.$$

Since $x/(1+x) \leq \log(1+x)$ for $x \geq 0$,

$$\sum_{t=t_0}^{t_1} \ell_t \leq \log \frac{\det V_{t_1+1}}{\det V_{t_0}}.$$

The determinants are positive since $V_{t_0} \succeq \underline{v}_H I \succ 0$. Set

$$S_n := \sum_{t=t_0}^{t_1} \psi_t \psi_t', \quad A_n := V_{t_0}^{-1/2} S_n V_{t_0}^{-1/2}, \quad n := t_1 - t_0 + 1.$$

Then $A_n \succeq 0$ and

$$\begin{aligned} \log \frac{\det V_{t_1+1}}{\det V_{t_0}} &= \log \det(I + A_n) \\ &\leq d \log \left(1 + \frac{\text{tr}(A_n)}{d} \right) \\ &\leq d \log \left(1 + \frac{n B_\psi^2}{d \underline{v}_H} \right) = \Gamma(n). \end{aligned} \tag{108}$$

The first inequality follows by applying Jensen's inequality to $\log(1+x)$ over the nonnegative eigenvalues of A_n . For the second,

$$\text{tr}(A_n) = \sum_{t=t_0}^{t_1} \psi_t' V_{t_0}^{-1} \psi_t \leq \frac{1}{\underline{v}_H} \sum_{t=t_0}^{t_1} \|\psi_t\|^2 \leq \frac{n B_\psi^2}{\underline{v}_H}.$$

For the standard elliptical-potential argument, see Abbasi-Yadkori et al. (2011, Lemma 11).

G.5 Completion of the Segment Bound

Sum (80). Since $Q_{t_1+1} \geq 0$, $\eta_{t+1}^2 \leq B_\eta^2$, and (107)–(108) hold,

$$\begin{aligned} c_0 \sum_t a_t^2 &\leq Q_{t_0} + 2 \sum_t a_t \eta_{t+1} + 2 B_\eta^2 \Gamma(n) \\ &\leq Q_{t_0} + \frac{c_0}{2} \sum_t a_t^2 + \frac{4 B_\eta^2}{c_0} \log(1/\delta_0) + 2 B_\eta^2 \Gamma(n). \end{aligned}$$

Rearranging yields (83). At initialization,

$$\begin{aligned} Q_{t_0} &= (\bar{\lambda} - \lambda)' V_{t_0} (\bar{\lambda} - \lambda) \\ &\leq E_\lambda^2 \lambda_{\max}(V_{t_0}) \leq E_\lambda^2 (\nu + n_H B_\psi^2) = \bar{Q}_H. \end{aligned}$$

Finally, (103) and the gate charging inequality give

$$\sum_t (y_{t+1} - w_{t+1})^2 = b_\star^2 \sum_t e_t^2 \leq \bar{b}^2 \sum_t a_t^2,$$

which proves Theorem 19.

H DETAILED DETECTOR PROOFS

This appendix supplies the conditional power estimates, concentration arguments, and calendar bounds used in Appendix C. The assumptions and tuning of Appendices A and B remain in force.

H.1 Finite-Horizon Power of the Auxiliary Process

Fix a change at control time τ and write $A = A_\Delta$. For $s - h \geq \tau$, iterating (68) for h steps gives

$$\psi_s^{\text{aux}} = A^h \psi_{s-h}^{\text{aux}} + \sum_{j=0}^{h-1} A^j e_1 w_{s-j}.$$

Taking the first coordinate and squaring, all cross terms involving distinct disturbances have conditional expectation zero. Thus

$$\mathbb{E}[(y_s^{\text{aux}})^2 \mid \mathcal{F}_{s-h}] = (e_1' A^h \psi_{s-h}^{\text{aux}})^2 + \sigma_w^2 \sum_{j=0}^{h-1} (e_1' A^j e_1)^2. \quad (109)$$

The stationary power is

$$J_y(\lambda^-; \theta^+) = \sigma_w^2 \sum_{j=0}^{\infty} (e_1' A^j e_1)^2.$$

Since the first term of (109) is nonnegative,

$$\begin{aligned} \mathbb{E}[(y_s^{\text{aux}})^2 \mid \mathcal{F}_{s-h}] &\geq J_y(\lambda^-; \theta^+) - \sigma_w^2 \sum_{j=h}^{\infty} \|A^j\|^2 \\ &\geq J_y(\lambda^-; \theta^+) - \frac{\sigma_w^2 C_\Delta^2 \rho_\Delta^{2h}}{1 - \rho_\Delta^2}. \end{aligned}$$

Equations (45) and (53) yield (72) for every initial state at time $s - h$.

The calculation also holds when s is a finite lag- h predictable random time. Indeed, $\{s = t\} \in \mathcal{F}_{t-h}$; multiplying the deterministic-time identity by $\mathbf{1}\{s = t\}$ and summing over t gives (109) with s and the stopped sigma-field $\mathcal{F}_{s-h}$.

H.2 Lag-Predictable Sampling

Let $h \leq s_1 < \dots < s_m$ be finite lag- h predictable output times with $s_{k+1} - s_k \geq h$. Define

$$\mathcal{H}_{k-1} := \mathcal{F}_{s_k-h}, \quad k = 1, \dots, m, \quad \mathcal{H}_m := \mathcal{F}_{s_m}.$$

Lag predictability makes these stopped sigma-fields well defined. The spacing condition gives $\mathcal{F}_{s_k} \subseteq \mathcal{H}_k$ for $k < m$, and the inclusion also holds for $k = m$ by definition. Hence $y_{s_k}^2 - \mathbb{E}[y_{s_k}^2 \mid \mathcal{F}_{s_k-h}]$ is $\mathcal{H}_k$ -measurable and has conditional mean zero given $\mathcal{H}_{k-1}$. This is the martingale structure used in Lemma 14; it need not hold for consecutive outputs centered at lag h , or for samples selected after observing their values.

H.3 Null Windows

Fix a potential restart r with a full uncontaminated exploration block and a potential test endpoint e . In the alarm-suppressed continuation from r , let $S_{r,e}$ be the corresponding block of m calendar samples. The block is null for r when each sampled output is generated under

the reference parameter. If that parameter remains active throughout $[r, \tau)$, a block ending at output time τ is null, since y_τ is generated at control time $\tau - 1$.

Condition on an $\mathcal{F}_{r+H}$ history on which $\mathcal{G}_r^{\text{est}}$ holds. The calendar fixes the block before its outputs are observed, and (64) applies to every stored output of a null block. Each $w_s^2 - \sigma_w^2$ is centered and lies in an interval of length B_w^2 . The Hoeffding exponential process (Hoeffding, 1963; Howard et al., 2020) therefore gives, for $x > 0$,

$$\mathbb{P}\left(\sum_{s \in S_{r,e}} (w_s^2 - \sigma_w^2) \geq mx \mid \mathcal{F}_{r+H}\right) \leq \exp\left(-\frac{2mx^2}{B_w^4}\right). \quad (110)$$

For the cross term, v_s is predictable and $|w_s v_s| \leq B_w \varepsilon_y$. If $\varepsilon_y = 0$, the cross term is identically zero. Otherwise, the corresponding Azuma–Hoeffding exponential process (Azuma, 1967; Hoeffding, 1963) gives

$$\mathbb{P}\left(\sum_{s \in S_{r,e}} w_s v_s \geq mx \mid \mathcal{F}_{r+H}\right) \leq \exp\left(-\frac{mx^2}{2B_w^2 \varepsilon_y^2}\right). \quad (111)$$

Taking $x = \Delta/8$ in (110) and $x = \Delta/16$ in (111) gives (66) and (67). The deterministic term in (65) is at most $\varepsilon_y^2 \leq \Delta/8$. Outside the two deviation events, the excess empirical power is therefore at most $3\Delta/8$, below the threshold $\Delta/2$.

If the actual test at e is reached, the alarm-suppressed continuation agrees with the implemented trajectory through that test. The conditional bound thus controls the joint event that the reference succeeds, the test is reached, its block is null for that reference, and the threshold is crossed. There are at most T^2 potential restart–endpoint pairs; applying the $\delta_d/(2T^2)$ one-block bound separately to each pair and taking a union bound gives a horizon-wide false-alarm contribution of at most $\delta_d/2$. Overlapping windows require no independence, and the argument does not condition on survival to a test.

H.4 Trajectory Coupling

Let $\delta_t = \psi_t - \psi_t^{\text{aux}}$ and use the input-channel bound G_u from (42). On $\mathcal{G}_r^{\text{est}}$, for control times after the change at τ and before the next alarm or true change, Lemma 13 gives

$$u_t = (\lambda^-)' \psi_t + e_t, \quad \|g_u(\theta^+) e_t\| \leq G_u \varepsilon_u.$$

Consequently,

$$\delta_{t+1} = A\delta_t + g_u(\theta^+) e_t, \quad \delta_\tau = 0.$$

Unrolling and using (44),

$$\|\delta_t\| \leq \sum_{j=0}^{t-\tau-1} \|A^j\| G_u \varepsilon_u \leq \frac{C_\Delta G_u \varepsilon_u}{1 - \rho_\Delta} = R_\Delta(H).$$

The gate bound uses feasibility of the old-gain action $(\lambda^-)' \psi_t$ at the actual state. The auxiliary action is unclipped and need not be feasible. The comparison holds up to the first alarm or next true change, while the actual plant parameter remains the fixed θ^+ used in A .

H.5 Calendar and Delay

Consider a change at control time τ under the hypotheses of Lemma 17: the cycle is still in exploitation after observing y_τ , $r+H \leq \tau$, and $\tau+D \leq T$. In its alarm-suppressed continuation,

stored output times form an arithmetic grid with gap h . Let s_1 be the first grid output with $s_1 - h \geq \tau$. The distance from $\tau + h$ to the next grid point is less than h , so

$$s_1 - \tau < 2h.$$

The m th grid output is $s_m = s_1 + (m - 1)h$, hence

$$s_m - \tau < (m + 1)h = D.$$

Every interval $[s_k - h, s_k)$ lies after the change, so the auxiliary outputs at these times satisfy (72). The post-change parameter remains active throughout the block by the hypotheses of Lemma 17. If any earlier block alarms, the change has already been detected. Otherwise, the actual and alarm-suppressed trajectories agree through the complete post-change test at s_m . Since times are integer valued, $s_m \leq \tau + D - 1$; the test is therefore available within the horizon.

For the induction in Theorem 18, the dwell condition places the first change after the initial exploration. If change $c < C$ is detected by $\tau^{(c)} + D$, the next exploration ends by $\tau^{(c)} + D + H \leq \tau^{(c+1)}$. Thus each subsequent change is preceded by a full uncontaminated reference block on the successful detection history.

I GLOBAL EVENT AND REGRET DECOMPOSITION

We combine the restart and detector guarantees on a common event and partition the control times to prove Theorem 20.

I.1 Global Success Event

Assume $T > H$; the case $T \leq H$ is handled deterministically in the proof of Theorem 20. Record each failure at its first adapted decision time: at the end of a reached full, unchanged exploration block if its Gram or self-normalized event fails; at the first same-parameter exploitation prefix violating its time-uniform RLS event; at a false detector crossing; or at the deadline of a missed change. Ties at an exploration endpoint are ordered as Gram, self-normalized, and then RLS. Since the true parameter sequence is deterministic in the analysis, these decisions are measurable with respect to the history at their decision times. Thus

$$\mathcal{E}_{\leq t}^{\text{good}} := \{\text{no failure has been recorded by time } t\} \in \mathcal{F}_t.$$

For each potential restart $r \in \{0, \dots, T\}$, define

$$\mathcal{C}_r := \{r \text{ is reached}\} \cap \mathcal{E}_{\leq r}^{\text{good}} \cap \{[r, r + H) \text{ is full and has one parameter}\}.$$

The first two events are $\mathcal{F}_r$ -measurable, and the last is fixed by the change sequence, so $\mathcal{C}_r \in \mathcal{F}_r$. Let $\mathcal{E}_r^G$ and $\mathcal{E}_r^\theta$ denote the Gram and self-normalized events. Let $\mathcal{E}_r^R$ be the time-uniform RLS event on the subsequent exploitation prefix with the same parameter as in $[r, r + H)$, ending at the first of the next true change, the next alarm output time, and the horizon. Set $\mathcal{E}_r^R = \Omega$ when this prefix is empty. The controller-failure events are

$$\begin{aligned} \mathcal{B}_r^G &:= \mathcal{C}_r \cap (\mathcal{E}_r^G)^c, \\ \mathcal{B}_r^\theta &:= \mathcal{C}_r \cap \mathcal{E}_r^G \cap (\mathcal{E}_r^\theta)^c, \\ \mathcal{B}_r^R &:= \mathcal{C}_r \cap \mathcal{E}_r^G \cap \mathcal{E}_r^\theta \cap (\mathcal{E}_r^R)^c. \end{aligned}$$

Conditional on $\mathcal{F}_r$, Propositions 21 and 22 bound the Gram and self-normalized failure probabilities by δ_0 each. On their joint success event, Theorem 19 bounds the conditional RLS failure

probability given $\mathcal{F}_{r+H}$ by δ_0 . Multiplying by the corresponding history indicators and applying the tower property gives

$$\mathbb{P}(\mathcal{B}_r^G) \leq \delta_0, \quad \mathbb{P}(\mathcal{B}_r^\theta) \leq \delta_0, \quad \mathbb{P}(\mathcal{B}_r^R) \leq \delta_0.$$

Consequently, for $\mathcal{B}_{\text{ctrl}} := \bigcup_{r=0}^T (\mathcal{B}_r^G \cup \mathcal{B}_r^\theta \cup \mathcal{B}_r^R)$,

$$\mathbb{P}(\mathcal{B}_{\text{ctrl}}) \leq 3(T+1)\delta_0 = \frac{3\delta}{32}. \quad (112)$$

On $\mathcal{B}_{\text{ctrl}}^c$, every reached full, unchanged reference whose restart precedes the first detector failure is accurate: otherwise the first controller failure would belong to one of the events above. Hence $\mathcal{B}_{\text{ctrl}}^c \cap \mathcal{E}_{\text{det}}^c \subseteq \mathcal{E}_{\text{ref}} \cap \mathcal{E}_{\text{det}}^c$, and Theorem 18 gives

$$\mathbb{P}(\mathcal{B}_{\text{ctrl}}^c \cap \mathcal{E}_{\text{det}}^c) \leq \delta_d.$$

This bounds a joint event and does not condition an early detector test on the success of a later estimator.

Define $\mathcal{E}_{\text{good}} := \mathcal{B}_{\text{ctrl}}^c \cap \mathcal{E}_{\text{det}}^c$. Then

$$\mathbb{P}(\mathcal{E}_{\text{good}}) \geq 1 - \frac{3\delta}{32} - \frac{\delta}{2} = 1 - \frac{19\delta}{32} > 1 - \delta.$$

On this event, the initial reference is valid, every change is detected, and the dwell condition ensures that each nonterminal post-alarm exploration block is full and uncontaminated.

I.2 Regret Decomposition

On $\mathcal{E}_{\text{good}}$, let $\hat{\tau}^{(c)}$ be the alarm corresponding to change c . Its delay interval is

$$\mathcal{D}_c := \{\tau^{(c)}, \dots, \hat{\tau}^{(c)} - 1\}, \quad |\mathcal{D}_c| \leq D, \quad c = 1, \dots, C.$$

The initial and post-alarm exploration intervals are

$$\begin{aligned} \mathcal{X}_0 &:= \{0, \dots, \min\{H, T\} - 1\}, \\ \mathcal{X}_c &:= \{\hat{\tau}^{(c)}, \dots, \min\{\hat{\tau}^{(c)} + H, T\} - 1\}, \quad c = 1, \dots, C. \end{aligned}$$

An interval is empty when its lower endpoint exceeds its upper endpoint. The dwell condition makes the delay and exploration intervals disjoint, and $|\mathcal{X}_c| \leq H$. Their complement consists of at most $C+1$ unchanged exploitation prefixes $\mathcal{P}_c$, one in each true segment, with empty prefixes allowed. Therefore

$$\{0, \dots, T-1\} = \left(\bigcup_{c=1}^C \mathcal{D}_c \right) \dot{\cup} \left(\bigcup_{c=0}^C \mathcal{X}_c \right) \dot{\cup} \left(\bigcup_{c=0}^C \mathcal{P}_c \right).$$

For $t \in \mathcal{D}_c$, the bound $|y_{t+1} - w_{t+1}| \leq B_Y + B_w$ gives cost at most K_D . For $t \in \mathcal{X}_c$, (36) gives cost at most K_X . Set $n_c := |\mathcal{P}_c|$. Every nonempty $\mathcal{P}_c$ follows a full, unchanged exploration block, so Theorem 19 bounds its regret by $R_E(n_c, \delta_0)$, uniformly over the resulting prefix length. An empty prefix has zero regret, also bounded by $R_E(0, \delta_0)$. Consequently,

$$\begin{aligned} \mathcal{R}_T &\leq \sum_{c=1}^C K_D |\mathcal{D}_c| + \sum_{c=0}^C K_X |\mathcal{X}_c| + \sum_{c=0}^C R_E(n_c, \delta_0) \\ &\leq CK_D D + (C+1)K_X H + (C+1)R_E(T, \delta_0), \end{aligned}$$

which is (86).

J TUNING FEASIBILITY AND ORDERS

We verify that the exploration and detector requirements in Appendix B admit the logarithmic choices used in Theorem 20. The asymptotic statements keep the model orders and the model, noise, initialization, exploration, ridge, stability, clipping, and gap bounds fixed.

J.1 Exploration Length

Equation (49) gives

$$\varepsilon_y = 5\bar{b}L_\lambda B_\psi E_\theta(H, \delta_0).$$

It is sufficient to require

$$E_\theta(H, \delta_0) \leq \epsilon_\star,$$

where

$$\epsilon_\star := \min \left\{ 1, \frac{b_{\min}}{16}, \frac{\sqrt{\Delta/8}}{5\bar{b}L_\lambda B_\psi}, \frac{\Delta(1 - \rho_\Delta)}{80B_o C_\Delta G_u L_\lambda B_\psi}, \frac{B_o}{5\bar{b}L_\lambda B_\psi} \right\}.$$

The four accuracy terms enforce the leading-coefficient bound, the null-bias margin, the coupling margin, and $\varepsilon_y \leq B_o$, respectively. The cap at one ensures $\log(1/\epsilon_\star) \geq 0$. All denominators are positive under the standing assumptions, including $L_\lambda > 0$ from Assumption 7.

The deterministic bound (46) satisfies

$$E_\theta(H, \delta_0) \leq \frac{K_0 \{1 + \sqrt{\log(1/\delta_0) + s \log(1 + K_1 H)}\}}{\sqrt{H - r_0}} \quad (113)$$

for constants K_0, K_1 independent of H, T, δ . Thus

$$\mathcal{H} := \left\{ H \in \mathbb{N} : H > r_0, \quad N_H \geq \frac{10B_\phi^4}{\kappa_{\text{inn}}^2} \log \frac{s}{\delta_0}, \quad E_\theta(H, \delta_0) \leq \epsilon_\star, \right. \\ \left. \frac{B_\psi^2}{\nu + \kappa_G n_H} \leq \frac{1}{16} \right\}$$

is nonempty: as $H \rightarrow \infty$, N_H grows linearly, while $E_\theta(H, \delta_0)$ and the leverage bound tend to zero. Every member satisfies (93) and (51)–(52). The smallest feasible integer selected in Appendix B is therefore at most $\min \mathcal{H}$.

For $a, b \geq 1$, the condition $x \geq 2a \log(2ab)$ implies

$$x \geq a \log(bx). \quad (114)$$

Indeed, setting $z = x/(2a)$ gives $\log(bx) = \log(2ab) + \log z \leq \log(2ab) + z \leq 2z = x/a$. Squaring (113), restricting to $H \geq 2r_0$, and applying (114) shows that the accuracy condition is satisfied with

$$H = O\left(\epsilon_\star^{-2} \{\log(1/\delta_0) + s \log(1/\epsilon_\star) + 1\}\right).$$

The Gram-size and leverage conditions require only $O(\log(1/\delta_0) + 1)$ additional lower bounds on H and are absorbed in this order. Since $\delta_0 = \delta/[32(T + 1)]$,

$$H = O(\log((T + 1)/\delta)). \quad (115)$$

J.2 Detector Spacing and Delay

Since $\rho_\Delta \in (0, 1)$, the spacing in (53) is finite. Writing $\log_+(x) := \max\{0, \log x\}$, its explicit form is

$$h = \max \left\{ 1, \left\lceil \frac{\log_+ \left(\frac{8\sigma_w^2 C_\Delta^2}{(1 - \rho_\Delta^2)\Delta} \right)}{2 \log(1/\rho_\Delta)} \right\rceil \right\}.$$

Thus h is independent of T . Equation (54) gives

$$m = O\left(\frac{B_o^4}{\Delta^2} \log \frac{T+1}{\delta}\right), \quad D = O\left(\frac{hB_o^4}{\Delta^2} \log \frac{T+1}{\delta}\right). \quad (116)$$

J.3 Regret Constants

Under (51),

$$c_0 \geq 1 - \frac{1}{8} - \frac{1}{8} \left(\frac{17}{16}\right)^2 > 0.73.$$

Equation (46) also yields

$$\begin{aligned} \bar{Q}_H &= E_\lambda^2(\nu + n_H B_\psi^2) = O(\log(1/\delta_0) + s \log(1 + H)) \\ &= O(\log((T+1)/\delta)), \end{aligned}$$

while

$$\Gamma(T) = d \log \left(1 + \frac{TB_\psi^2}{d\underline{v}_H} \right) = O(d \log T).$$

Therefore $R_E(T, \delta_0) = O(\log((T+1)/\delta))$. Substituting this bound together with (115) and (116) into (86) proves (87).

K RELAXATIONS OF CLIPPING AND MISMATCH STABILITY

This appendix establishes the feasibility and detection results used in Sections 6 and 5. The extensions use either eventual nonclipping or own-regime feasibility with a boundary alarm.

K.1 Closed Clipping Feasibility

Let $P(x) = \text{clip}_{[-B_u, B_u]}(x)$. Projection onto a closed interval is nonexpansive. If $x \in [-B_u, B_u]$, including $x = \pm B_u$, then

$$|P(z) - x| = |P(z) - P(x)| \leq |z - x|. \quad (117)$$

Equation (117) is the only consequence of the clipping condition used in Lemma 13. That lemma and every later detector, RLS, and global-regret argument therefore hold under the closed feasibility condition in Assumption 10; no positive margin enters a proof constant.

For the stable-mismatch theorem, Assumption 10 can be replaced by feasibility of the true reference gain along every unchanged exploitation prefix and every complete alarm-suppressed continuation used in the detector analysis.

K.2 Eventual Inactivity of Clipping

A quantitative steady-state slack condition yields eventual inactivity of clipping and a burn-in time uniform over restarts. This gives the variant of PIECE-CD in Theorem 6.

Define the forced-response output bound and its corresponding regressor bound by

$$\begin{aligned} Y_\infty(B_u) &:= \frac{C_y}{1 - \rho_y}(B_w + B_b B_u), \\ P_\infty(B_u) &:= \left\{ Y_\infty(B_u)^2 + (q - 1)B_u^2 \right\}^{1/2}, \\ \mu_{\text{nc}} &:= B_u - \varepsilon_u - \Lambda P_\infty(B_u). \end{aligned} \tag{118}$$

Here $\varepsilon_u = \frac{5}{4}g_H B_\psi$, defined in (49), bounds the raw-proposal error on $\mathcal{G}_r^{\text{est}}$. Suppose $\mu_{\text{nc}} > 0$. With the convention $\tau_{\text{nc}} = 0$ when $\Lambda C_y B_Y = 0$, put otherwise

$$\tau_{\text{nc}} := \max \left\{ 0, \left\lceil \frac{\log(\Lambda C_y B_Y / \mu_{\text{nc}})}{\log(1/\rho_y)} - H \right\rceil \right\}. \tag{119}$$

Proposition 23 (Uniform eventual inactivity of clipping). *Fix a reached restart r whose full exploration block has one parameter θ^- , and suppose $\mathcal{G}_r^{\text{est}}$ holds. At every exploitation time before the next restart satisfying*

$$t \geq r + H + \tau_{\text{nc}},$$

the raw gated proposal belongs to the clipping interval:

$$|z_t| \leq B_u, \quad u_t = z_t. \tag{120}$$

The conclusion remains valid after a plant change and does not require the old-gain/new-plant mismatch matrix to be Schur.

Proof. The implemented input is clipped at every time, so $|u_t| \leq B_u$ regardless of whether the conclusion has already become true. Starting the state recursion (88) at r , uniform switching stability therefore gives, for every $t \geq r$ and for any admissible changes after r ,

$$\begin{aligned} \|Y_t\| &\leq C_y \rho_y^{t-r} \|Y_r\| + Y_\infty(B_u) \\ &\leq C_y \rho_y^{t-r} B_Y + Y_\infty(B_u). \end{aligned}$$

The $q - 1$ stored inputs in ψ_t each have magnitude at most B_u . Consequently,

$$\begin{aligned} \|\psi_t\| &\leq \left[\{Y_\infty(B_u) + C_y \rho_y^{t-r} B_Y\}^2 + (q - 1)B_u^2 \right]^{1/2} \\ &\leq P_\infty(B_u) + C_y \rho_y^{t-r} B_Y. \end{aligned}$$

On $\mathcal{G}_r^{\text{est}}$, the raw (pre-clipping) part of the gate calculation in Lemma 13 does not use feasibility. On the recursive branch its distance from $(\lambda^-)^\top \psi_t$ is at most $(g_H + g_H/4) \|\psi_t\|$, and on the reference branch it is at most $(g_H/4) \|\psi_t\|$. Hence, on either branch and also after a change,

$$|z_t - (\lambda^-)^\top \psi_t| \leq \varepsilon_u.$$

Since $\|\lambda^-\| \leq \Lambda$, the preceding two displays imply

$$\begin{aligned} |z_t| &\leq \varepsilon_u + \Lambda P_\infty(B_u) + \Lambda C_y B_Y \rho_y^{t-r} \\ &= B_u - \mu_{\text{nc}} + \Lambda C_y B_Y \rho_y^{t-r}. \end{aligned}$$

Definition (119) makes the last term at most μ_{nc} whenever $t - r \geq H + \tau_{\text{nc}}$, proving (120). It also gives $|(\lambda^-)^\top \psi_t| \leq B_u - \varepsilon_u$. The calculation is pathwise on the reference event and uniform over admissible switching sequences. $\square$

Proposition 22 and the restart union in (74) give $\mathbb{P}(\mathcal{E}_{\text{ref}}) \geq 1 - 2(T + 1)\delta_0 = 1 - \delta/16$. On this event, the conclusion holds simultaneously for every reached full, unchanged exploration block whose restart precedes the first detector failure.

Both P_∞ and ε_u depend on B_u , so a positive μ_{nc} is not guaranteed by increasing the clipping threshold. An explicit sufficient condition is obtained by setting

$$a_0 := \frac{C_y B_w}{1 - \rho_y}, \quad a_1 := \frac{C_y B_b}{1 - \rho_y}, \quad \kappa_u := 1 - \Lambda\{a_1 + \sqrt{q - 1}\}.$$

If $\kappa_u > 0$, choose

$$B_u > \frac{2\Lambda a_0}{\kappa_u} \quad \text{and then choose } H \text{ so that } \varepsilon_u \leq \frac{\kappa_u B_u}{2}. \quad (121)$$

Indeed, $P_\infty(B_u) \leq a_0 + (a_1 + \sqrt{q - 1})B_u$, so (121) implies $\mu_{\text{nc}} > 0$. For fixed B_u the required inequality for ε_u is achieved by a sufficiently large feasible H (when the horizon permits) under the exploration bound in (46). A direct sufficient check for immediate nonclipping is $(\Lambda + g_H)B_\psi \leq B_u$: since $\|\bar{\lambda}_r\| \leq \Lambda$, every raw exploitation proposal then lies in the clipping interval and no burn-in is needed.

Bounded noise alone does not ensure eventual nonclipping. Consider the known scalar system

$$y_{t+1} = \frac{1}{2}y_t + u_t + w_{t+1}, \quad B_u = \frac{1}{4}, \quad \mathbb{P}(w_t = 1) = \mathbb{P}(w_t = -1) = \frac{1}{2}.$$

The exact minimum-variance gain is $\lambda = -1/2$. If the exact minimum-variance law were unclipped from some finite time onward, then $y_{t+1} = w_{t+1}$ from that time onward, and every following raw action would have magnitude $1/2 > B_u$, a contradiction. Thus no finite τ_{nc} exists for this example under that law.

Burn-in variant. Before the burn-in ends, clipping may invalidate gate charging and the detector estimates. After each full exploration block, apply reference control for τ_{nc} rounds with detection and RLS updates disabled and the detector buffer empty. Use $z_t = \bar{\lambda}_r^\top \psi_t$ and $u_t = \text{clip}_{[-B_u, B_u]}(z_t)$. At $t_0 = r + H + \tau_{\text{nc}}$ set $V_{t_0} = \nu I_d + \sum_{t \in \mathcal{I}_r} \psi_t \psi_t^\top$ and $\hat{\lambda}_{t_0} = \bar{\lambda}_r$, then use the shifted output grid $\mathcal{S}_r^{\text{nc}} = \{t_0 + h + jh : j \geq 0\}$. Truncate exploration or burn-in at T . Require the first change no earlier than $H + \tau_{\text{nc}}$, successive changes to be separated by at least $D + H + \tau_{\text{nc}}$, and the terminal segment to have length at least D . From t_0 onward clipping is inactive by (120), so the baseline gate, false-alarm, coupling, and RLS proofs apply without Assumption 10. The regret bound gains only the deterministic term $(C + 1)K_D \tau_{\text{nc}}$, and the logarithmic order is unchanged whenever $\tau_{\text{nc}} = O(\log((T + 1)/\delta))$ (in particular, when it is a fixed problem-dependent constant).

K.3 Own-Regime Feasibility and a Boundary Alarm

A boundary alarm controls the post-change action error under feasibility of each gain only on regressors generated by its own plant.

Call a regressor an *own-regime successor* for θ if it can occur after at least r_0 consecutive controls generated by plant θ , starting from any regressor with norm at most B_ψ , under disturbances bounded by B_w and arbitrary inputs in $[-B_u, B_u]$. This definition includes the successors produced during exploration and the terminal state at a segment boundary; it is not restricted to states generated by the ideal policy.

Assumption 24 (Closed own-regime clipping feasibility). *For every $\theta \in \Theta$ and every own-regime successor ψ for θ ,*

$$|\lambda(\theta)^\top \psi| \leq B_u. \quad (122)$$

Equality is allowed, and no positive clipping margin is required.

Assumption 24 is an alternative feasibility condition: it checks each gain only on its own-regime successors and does not require the old gain to remain feasible on states produced by the new regime. Because the successor definition permits any norm-bounded starting regressor, including unreachable states, this condition is not asserted to follow from Assumption 10.

Define the deterministic reference uncertainty and safe tracking bounds

$$\begin{aligned}\zeta_H &:= E_\lambda(H, \delta_0)B_\psi, \\ \bar{\varepsilon}_u &:= \varepsilon_u + 2\zeta_H = 7L_\lambda B_\psi E_\theta(H, \delta_0), \\ \bar{\varepsilon}_y &:= \bar{b}\bar{\varepsilon}_u.\end{aligned}\tag{123}$$

The branch-free union-class version of PIECE-CD adds the following test. After an exploitation round t , once y_{t+1} has been observed and ψ_{t+1} has been formed, raise one alarm and restart at control time $t + 1$ if

$$|\bar{\lambda}_r^\top \psi_{t+1}| > B_u + \zeta_H.\tag{124}$$

The strict inequality is part of the test: under the null, equality with the upper bound is possible for discrete disturbances. If both the boundary and energy tests cross at the same output, they count as one alarm. Algorithm 3 states the complete composite procedure.

Lemma 25 (Boundary-alarm dichotomy). *Under Assumption 24, fix a reached cycle whose full exploration block has parameter θ^- , and suppose $\mathcal{G}_r^{\text{est}}$ holds. Then the boundary test cannot cross while the reference parameter remains active. On such an unchanged exploitation interval, with $a_t := (\hat{\lambda}_t - \lambda(\theta^-))^\top \psi_t$ and $e_t := u_t - \lambda(\theta^-)^\top \psi_t$, one also has $e_t^2 \leq a_t^2$. After a change within this exploitation cycle, at every control time before the cycle's next alarm,*

$$(|\lambda(\theta^-)^\top \psi_t| - B_u)_+ \leq 2\zeta_H,\tag{125}$$

$$|u_t - \lambda(\theta^-)^\top \psi_t| \leq \bar{\varepsilon}_u.\tag{126}$$

These inequalities also hold in the alarm-suppressed continuation at control states satisfying $|\bar{\lambda}_r^\top \psi_t| \leq B_u + \zeta_H$.

Proof. While the reference parameter θ^- remains active, every exploitation regressor is an own-regime successor. On $\mathcal{G}_r^{\text{est}}$, $\|\bar{\lambda}_r - \lambda(\theta^-)\| \leq E_\lambda(H, \delta_0)$, so (122) gives

$$\begin{aligned}|\bar{\lambda}_r^\top \psi_t| &\leq |\lambda(\theta^-)^\top \psi_t| + \|\bar{\lambda}_r - \lambda(\theta^-)\| \|\psi_t\| \\ &\leq B_u + \zeta_H.\end{aligned}$$

The strict test in (124) therefore cannot give a boundary false alarm. Since the old ideal action is feasible on the unchanged interval, the gate-charging and projection argument of Lemma 13 applies verbatim and gives $e_t^2 \leq a_t^2$.

For a change occurring during the exploitation cycle in the lemma, at the first post-change control the regressor is still an own-regime successor for θ^- , so its saturation deficit is zero. At every later post-change control time with no previous boundary alarm,

$$|\lambda(\theta^-)^\top \psi_t| \leq |\bar{\lambda}_r^\top \psi_t| + E_\lambda(H, \delta_0)B_\psi \leq B_u + 2\zeta_H,$$

which proves (125). The raw gated proposal satisfies $|z_t - \lambda(\theta^-)^\top \psi_t| \leq \varepsilon_u$. For $P(x) = \text{clip}_{[-B_u, B_u]}(x)$ and any $x, z \in \mathbb{R}$,

$$\begin{aligned}|P(z) - x| &\leq |P(z) - P(x)| + |P(x) - x| \\ &\leq |z - x| + (|x| - B_u)_+.\end{aligned}$$

Combining this inequality with (125) gives (126). The same calculation applies pointwise in the alarm-suppressed continuation whenever $|\bar{\lambda}_r^\top \psi_t| \leq B_u + \zeta_H$. $\square$

K.4 Detection under the Union Condition

The boundary tracking bound (126) permits a common finite-window detector for stable, directly visible, and non-Schur ARX mismatches.

For an admissible adjacent pair $z = (\theta^-, \theta^+)$, write

$$A_z := A(\theta^-, \theta^+), \quad G_h(A_z) := \sigma_w^2 \sum_{j=1}^{h-1} (e_1^\top A_z^j e_1)^2.$$

The supplied admissible adjacent-pair class $\mathcal{P} \subseteq \Theta^2$ has a cover $\mathcal{P} \subseteq \mathcal{P}_S \cup \mathcal{P}_F$. The finite-window component may itself be covered by $\mathcal{P}_F \subseteq \mathcal{P}_D \cup \mathcal{P}_A$. The set $\mathcal{P}_D$ uses a directly supplied finite-order gap, whereas $\mathcal{P}_A$ uses only the ARX non-Schur condition $\rho(A_z) \geq 1$. These sets may overlap. These classes specify uniform tuning bounds.

Define the finite-order innovation visibility

$$\mathcal{V}(z) := \sigma_w^2 \sum_{j=1}^d \{e_1^\top A_z^j e_1\}^2. \quad (127)$$

For $\theta \in \Theta$, define the monic old-input polynomial

$$\bar{B}_\theta(\xi) := \xi^{q-1} + \frac{b_2(\theta)}{b_1(\theta)} \xi^{q-2} + \dots + \frac{b_q(\theta)}{b_1(\theta)},$$

with $\bar{B}_\theta \equiv 1$ when $q = 1$. Under the minimum-phase convention in Assumption 7, every root of $\bar{B}_\theta$ lies in the open unit disk. Continuity of polynomial roots and compactness therefore give a safe uniform number

$$r_B \geq \sup_{\theta \in \Theta} \max\{|\beta| : \bar{B}_\theta(\beta) = 0\}, \quad 0 \leq r_B < 1, \quad (128)$$

where the maximum over the empty set is zero. Also let

$$R_A \geq \max \left\{ 1, \sup_{(\theta^-, \theta^+) \in \Theta^2} \|A_{\text{cl}}(\lambda(\theta^-), \theta^+)\|_2 \right\} < \infty, \quad K_A := \sum_{k=1}^d R_A^{d-k} \sum_{j=0}^{k-1} \binom{d}{j} R_A^j. \quad (129)$$

These are offline model-class bounds. The uniform root bound follows from continuity of polynomial zeros and compactness (Horn and Johnson, 2013, Appendix D, Theorem D1, and Appendix E).

Finite-order visibility of non-Schur mismatches. Define

$$\underline{\Delta}_A := \sigma_w^2 \left\{ \frac{(1 - r_B)^{q-1}}{K_A} \right\}^2 > 0. \quad (130)$$

Then every ARX mismatch matrix A_z satisfying $\rho(A_z) \geq 1$ obeys

$$\mathcal{V}(z) \geq \underline{\Delta}_A. \quad (131)$$

This implication uses the ARX and minimum-phase structure; it is false for a general square matrix.

To prove it, put $h_j = e_1^\top A_z^j e_1$ and

$$D_z(\xi) := \det(\xi I - A_z) = \xi^d + \sum_{k=1}^d d_k \xi^{d-k},$$

$$Q_-(\xi) := \xi^p \bar{B}_{\theta^-}(\xi) = \xi^d + \sum_{k=1}^d q_k \xi^{d-k}.$$

A cofactor expansion using the shift rows of A_z gives

$$e_1^\top \text{adj}(\xi I - A_z) e_1 = \xi^{p-1} \bar{B}_{\theta^-}(\xi). \quad (132)$$

Indeed, deleting the first row and column leaves a block-lower-triangular minor: its output-shift block has determinant ξ^{p-1} and its old-input companion block has determinant $\bar{B}_{\theta^-}(\xi)$. The adjugate formula and the Laurent expansion at infinity then give the exact disturbance-to-output identity, for $|\xi| > \rho(A_z)$ (equivalently, as a formal Laurent identity at infinity),

$$\xi e_1^\top (\xi I - A_z)^{-1} e_1 = 1 + \sum_{j \geq 1} h_j \xi^{-j} = \frac{Q_-(\xi)}{D_z(\xi)}. \quad (133)$$

Comparing the coefficient of ξ^{d-k} in $D_z(\xi)(1 + \sum_{j \geq 1} h_j \xi^{-j}) = Q_-(\xi)$ gives, with $d_0 = 1$,

$$q_k - d_k = \sum_{i=1}^k d_{k-i} h_i, \quad k = 1, \dots, d. \quad (134)$$

Let $\eta := \max_{1 \leq j \leq d} |h_j|$ and choose an eigenvalue $\lambda_\star$ with $|\lambda_\star| = \rho(A_z) \geq 1$. Since $|d_j| \leq \binom{d}{j} R_A^j$ and $|\lambda_\star| \leq R_A$ (the spectral radius is bounded by every induced matrix norm (Horn and Johnson, 2013, Theorem 5.6.9)), equations (134) and $D_z(\lambda_\star) = 0$ imply

$$|Q_-(\lambda_\star)| \leq \eta K_A.$$

On the other hand, Q_- has p zeros at zero and its other $q-1$ zeros have modulus at most r_B , so

$$|Q_-(\lambda_\star)| \geq |\lambda_\star|^p (|\lambda_\star| - r_B)^{q-1} \geq (1 - r_B)^{q-1}.$$

Combining the last two displays and using $\mathcal{V}(z) \geq \sigma_w^2 \eta^2$ proves (131). If a stronger class-wide margin $\rho(A_z) \geq c_a > 1$ is available, the same proof gives the sharper constant $\underline{\Delta}_A(c_a) := \sigma_w^2 \{c_a^p (c_a - r_B)^{q-1} / K_A\}^2$, which may replace $\underline{\Delta}_A$ in the tuning. Such a margin is not required for the theorem below.

Assumption 26 (Union stable/direct/non-Schur detectability). *Every realized adjacent pair $z_c = (\theta^{(c-1)}, \theta^{(c)})$, $c = 1, \dots, C$, belongs to $\mathcal{P}$. For a nonempty stable component there are supplied constants $C_S \geq 1$, $\rho_S \in (0, 1)$, and $\underline{\Delta}_S > 0$; for a nonempty direct-visibility component there is a supplied $\underline{\Delta}_D > 0$; and the possibly empty ARX non-Schur component satisfies*

$$\begin{aligned} z \in \mathcal{P}_S &\implies \|A_z^k\| \leq C_S \rho_S^k \quad (k \geq 0), & J_y(\lambda(\theta^-); \theta^+) - \sigma_w^2 &\geq \underline{\Delta}_S, \\ z \in \mathcal{P}_D &\implies \mathcal{V}(z) \geq \underline{\Delta}_D, \\ z \in \mathcal{P}_A &\implies \rho(A_z) \geq 1. \end{aligned} \quad (135)$$

Each admitted pair satisfies at least one component condition. A stable pair requires no separate finite-order gap; for a non-Schur pair, that gap is supplied by (130).

Set $\Gamma_D = \underline{\Delta}_D$ when $\mathcal{P}_D \neq \emptyset$ and $\Gamma_D = +\infty$ otherwise. Set $\Gamma_A = \underline{\Delta}_A$ when $\mathcal{P}_A \neq \emptyset$ and $\Gamma_A = +\infty$ otherwise. If $\mathcal{P}_F \neq \emptyset$, define

$$\underline{\Delta}_F := \min\{\Gamma_D, \Gamma_A\} > 0. \quad (136)$$

Equations (127) and (131) then give the derived common bound

$$z \in \mathcal{P}_F \implies \mathcal{V}(z) \geq \underline{\Delta}_F. \quad (137)$$

This is the finite-visible condition used in Section 5.

For a nonempty stable component, define

$$h_S := \min \left\{ k \geq 1 : \frac{\sigma_w^2 C_S^2 \rho_S^{2k}}{1 - \rho_S^2} \leq \frac{\underline{\Delta}_S}{8} \right\}, \quad \Gamma_S := \frac{7\underline{\Delta}_S}{8}.$$

If $\mathcal{P}_S = \emptyset$, set $(h_S, \Gamma_S) = (1, +\infty)$. If $\mathcal{P}_F \neq \emptyset$, set $(h_F, \Gamma_F) = (d+1, \underline{\Delta}_F)$; otherwise set $(h_F, \Gamma_F) = (1, +\infty)$. If $\mathcal{P} = \emptyset$, set $h_U = 1$ and choose any fixed finite $\underline{\Delta}_U > 0$ for null calibration; every missed-change statement is then vacuous. Otherwise at least one cover component is nonempty, and set

$$h_U := \max\{h_S, h_F\}, \quad \underline{\Delta}_U := \min\{\Gamma_S, \Gamma_F\} > 0. \quad (138)$$

The resulting common finite-horizon power separation is

$$G_{h_U}(A_z) \geq \underline{\Delta}_U, \quad z \in \mathcal{P}. \quad (139)$$

Indeed, if $z \in \mathcal{P}_S$, then

$$\begin{aligned} G_{h_U}(A_z) &= J_y(\lambda(\theta^-); \theta^+) - \sigma_w^2 - \sigma_w^2 \sum_{j=h_U}^{\infty} (e_1^\top A_z^j e_1)^2 \\ &\geq \underline{\Delta}_S - \frac{\sigma_w^2 C_S^2 \rho_S^{2h_U}}{1 - \rho_S^2} \geq \frac{7\underline{\Delta}_S}{8} \geq \underline{\Delta}_U. \end{aligned}$$

If $z \in \mathcal{P}_F$, then $h_U \geq d+1$ and the nonnegative summands give $G_{h_U}(A_z) \geq \mathcal{V}(z) \geq \underline{\Delta}_F \geq \underline{\Delta}_U$. This proves (139) without identifying which route holds for the realized change.

The first d impulse coefficients exhaust finite-order visibility by Cayley–Hamilton (Horn and Johnson, 2013, Theorem 2.4.3.2). Let $h_j = e_1^\top A^j e_1$, so $h_0 = 1$, and write the characteristic polynomial as $p_A(z) = z^d + \kappa_{d-1}z^{d-1} + \dots + \kappa_0$. If $h_1 = \dots = h_d = 0$, sandwiching $p_A(A) = 0$ between $e_1^\top$ and e_1 gives $\kappa_0 = 0$. Sandwiching $A^k p_A(A) = 0$ for $k \geq 1$ and using strong induction then gives $h_j = 0$ for every $j \geq 1$. Thus every change visible through a noise-driven impulse coefficient is visible by order d .

If $\mathcal{P} = \emptyset$, set $L_{0,U} = L_{w,U} = L_{u,U} = 0$. Otherwise, the same model-class bound R_A gives the explicit finite-horizon choices

$$\begin{aligned} L_{0,U} &:= R_A^{h_U}, \\ L_{w,U} &:= \sum_{j=0}^{h_U-1} R_A^j, \quad L_{u,U} := G_u \sum_{j=0}^{h_U-1} R_A^j. \end{aligned} \quad (140)$$

Indeed, $|x^\top A^j y| \leq \|x\|_2 \|A\|_2^j \|y\|_2$. These bounds are finite because h_U is finite, even when some matrices in $\mathcal{P}_A$ are non-Schur. Sharper supplied bounds may replace them without changing the proof.

Define

$$\begin{aligned} B_{\text{aux},U} &:= L_{0,U} B_\psi + L_{w,U} B_w, \\ B_U &:= \max\{B_Y, B_{\text{aux},U}, B_w\}, \quad R_U(H) := L_{u,U} \bar{\varepsilon}_u. \end{aligned} \quad (141)$$

In addition to the exploration sample-size condition and (51), choose the smallest H satisfying

$$\bar{\varepsilon}_y^2 \leq \frac{\underline{\Delta}_U}{8}, \quad \bar{\varepsilon}_y \leq B_U, \quad 2B_U R_U(H) \leq \frac{\underline{\Delta}_U}{8}. \quad (142)$$

Use spacing h_U , the energy threshold $\sigma_w^2 + \underline{\Delta}_U/2$, and set

$$m_U := \left\lceil \frac{512B_U^4}{\underline{\Delta}_U^2} \log \frac{4T^2}{\delta_d} \right\rceil, \quad D_U := (m_U + 1)h_U. \quad (143)$$

Assumption 12 is read with D_U in place of D .

Fix $z = (\theta^-, \theta^+) \in \mathcal{P}$ and a potential restart r whose exploration reference estimates $\lambda(\theta^-)$. Continue the clipped controller on the same restart-relative calendar with both alarms suppressed. Its regressor remains bounded by B_ψ on every path. For a stored output time s with $s - h_U \geq r + H$, define the h_U -step auxiliary

$$\begin{aligned}\tilde{\psi}_{s-h_U}^{(s)} &:= \psi_{s-h_U}, \\ \tilde{\psi}_{t+1}^{(s)} &= A_z \tilde{\psi}_t^{(s)} + e_1 w_{t+1}, \quad s - h_U \leq t < s.\end{aligned}\tag{144}$$

Let $\tilde{y}_s^{(s)}$ be its first coordinate. This auxiliary is defined on every continuation path, including paths on which the boundary test would cross.

Lemma 27 (Sample-local union power and comparison). *For the pair $z \in \mathcal{P}$ and continuation defined above, suppose θ^+ is active throughout $[s - h_U, s)$. Then*

$$|\tilde{y}_s^{(s)}| \leq B_{\text{aux}, U}, \tag{145}$$

$$\mathbb{E}[(\tilde{y}_s^{(s)})^2 \mid \mathcal{F}_{s-h_U}] \geq \sigma_w^2 + \underline{\Delta}_U. \tag{146}$$

On $\mathcal{G}_r^{\text{est}}$ and on the event that the boundary test has not crossed at the states used for the controls in $[s - h_U, s)$,

$$\begin{aligned}|y_s - \tilde{y}_s^{(s)}| &\leq R_U(H), \\ |y_s^2 - (\tilde{y}_s^{(s)})^2| &\leq 2B_U R_U(H) \leq \frac{\Delta_U}{8}.\end{aligned}\tag{147}$$

The conditional statements remain valid at finite lag- h_U predictable output times.

Proof. Iterating (144) gives

$$\tilde{\psi}_s^{(s)} = A_z^{h_U} \psi_{s-h_U} + \sum_{j=0}^{h_U-1} A_z^j e_1 w_{s-j}.$$

The definitions of $L_{0,U}$ and $L_{w,U}$ give (145). Conditional on $\mathcal{F}_{s-h_U}$, the starting state is fixed and $w_{s-h_U+1}, \dots, w_s$ are fresh, independent, and centered. Hence

$$\begin{aligned}\mathbb{E}[(\tilde{y}_s^{(s)})^2 \mid \mathcal{F}_{s-h_U}] &= (e_1^\top A_z^{h_U} \psi_{s-h_U})^2 + \sigma_w^2 \sum_{j=0}^{h_U-1} (e_1^\top A_z^j e_1)^2 \\ &\geq \sigma_w^2 + G_{h_U}(A_z) \geq \sigma_w^2 + \underline{\Delta}_U,\end{aligned}$$

which proves (146).

On the no-boundary event in the statement, Lemma 25 gives $u_t = \lambda(\theta^-)^\top \psi_t + e_t$ with $|e_t| \leq \bar{\varepsilon}_u$ throughout $[s - h_U, s)$. Subtracting the two state recursions and unrolling from their common state at $s - h_U$ gives

$$|y_s - \tilde{y}_s^{(s)}| \leq \bar{\varepsilon}_u \sum_{j=0}^{h_U-1} |e_1^\top A_z^j g_u(\theta^+)| \leq L_{u,U} \bar{\varepsilon}_u.$$

Both output magnitudes are at most B_U , so $|x^2 - z^2| \leq |x - z|(|x| + |z|)$ proves (147).

For a finite lag- h_U predictable time, multiply the deterministic-time identities by $\mathbf{1}\{s = t\} \in \mathcal{F}_{t-h_U}$ and sum over t . This gives the same relations with the stopped field $\mathcal{F}_{s-h_U}$. $\square$

Lemma 28 (One-change detection under the union condition). *Adopt the stopping-time, active-cycle, and horizon hypotheses of Lemma 17, use Algorithm 3, replace D by D_U , and suppose Assumptions 24 and 26 hold. Almost surely on the corresponding $\mathcal{A}_{r,\tau}$,*

$$\mathbb{P}(\hat{\tau}_r > \tau + D_U \mid \mathcal{F}_\tau) \leq \frac{\delta_d}{4T^2}. \quad (148)$$

Proof. Condition on an arbitrary $\mathcal{F}_\tau$ history in $\mathcal{A}_{r,\tau}$ and use the alarm-suppressed continuation just defined. The deterministic calendar argument in Appendix H gives stored output times $s_1, \dots, s_{m_U}$ such that $s_1 - h_U \geq \tau$, $s_{k+1} - s_k = h_U$, and $s_{m_U} - \tau < D_U$. Conditionally on $\mathcal{F}_\tau$, the restart-relative grid is lag- h_U predictable, $\mathcal{F}_\tau \subseteq \mathcal{F}_{s_1 - h_U}$, and the disturbances used by the sample-local auxiliaries remain fresh.

Apply the conditional form of Lemma 14 given $\mathcal{F}_\tau$ to the sample-local variables $\tilde{y}_{s_k}^{(s_k)}$, using the lagged filtration $\mathcal{H}_{k-1} := \mathcal{F}_{s_k - h_U}$. Lemma 27 gives the required conditional means, and $|\tilde{y}_{s_k}^{(s_k)}| \leq B_U$. Therefore

$$\begin{aligned} \mathbb{P}\left(\frac{1}{m_U} \sum_{k=1}^{m_U} (\tilde{y}_{s_k}^{(s_k)})^2 < \sigma_w^2 + \frac{7\Delta_U}{8} \mid \mathcal{F}_\tau\right) &\leq \exp\left(-\frac{m_U \Delta_U^2}{128 B_U^4}\right) \\ &\leq \frac{\delta_d}{4T^2}. \end{aligned} \quad (149)$$

This concentration bound is not conditioned on survival of the boundary test.

If either test raises an alarm at or before s_{m_U} , the change has already been detected. Otherwise, the actual and alarm-suppressed trajectories agree through the complete block, and (147) holds at every sample. Outside the deviation in (149), the actual empirical energy is at least $\sigma_w^2 + 3\Delta_U/4$, which crosses the threshold $\sigma_w^2 + \Delta_U/2$. Consequently,

$$\{\hat{\tau}_r > \tau + D_U\} \subseteq \left\{ \frac{1}{m_U} \sum_{k=1}^{m_U} (\tilde{y}_{s_k}^{(s_k)})^2 < \sigma_w^2 + \frac{7\Delta_U}{8} \right\}.$$

Taking probabilities proves (148). $\square$

A smaller Δ_U increases the exploration and window requirements in (142)–(143). Uniform finite delay requires every admitted Schur pair to satisfy either the stable-gap or direct-visibility condition.

A scalar non-Schur mismatch. Take $p = q = 1$, $B_u = 10$, zero initial history, and Rademacher noise with $B_w = \sigma_w = 1$. Let $\Theta = \{\theta^-, \theta^+\}$ and consider the one-change sequence $\theta^- \rightarrow \theta^+$, where

$$\theta^- = (-0.2, 0.2), \quad \theta^+ = (0.2, 1).$$

Both open-loop coefficients have magnitude 0.2, and the deterministic output bound is $B_Y = 11/(1 - 0.2) = 13.75$. The interval $[-B_Y, B_Y]$ is invariant under either plant with $|u_t| \leq B_u$ and $|w_t| \leq 1$. The two ideal gains are $\lambda^- = 1$ and $\lambda^+ = -0.2$. Every one-step own-regime successor satisfies

$$\begin{aligned} |\lambda^- y^+| &\leq 0.2 B_Y + 0.2 B_u + B_w = 5.75, \\ |\lambda^+ y^+| &\leq 0.2 B_Y = 2.75. \end{aligned}$$

Thus Assumption 24 holds, while the old gain need not be feasible on every state produced by the new plant. Nevertheless, $A(\theta^-, \theta^+) = 0.2 + 1 \cdot 1 = 1.2$. Here $d = 1$, so $h_F = 2$ and $\mathcal{V}(\theta^-, \theta^+) = 1.2^2 = 1.44$. Since $r_B = 0$ here, the non-Schur formula itself supplies the valid positive bound $\Delta_A = 1$; no separate visibility assumption is needed.

K.5 Regret with Saturation

Without clipping feasibility, exploitation regret can be controlled by the realized saturation energy. For a true gain λ , define

$$d_t^{\text{sat}} := (|\lambda^\top \psi_t| - B_u)_+. \quad (150)$$

On an unchanged exploitation segment with a successful reference and $g_H = 4E_\lambda$, the unclipped gate-charging inequality from Appendix G and the projection distance inequality give

$$|u_t - \lambda^\top \psi_t| \leq |a_t| + d_t^{\text{sat}}, \quad a_t = (\hat{\lambda}_t - \lambda)^\top \psi_t. \quad (151)$$

Let $\beta_s := 1/8$ and define

$$\begin{aligned} c_s &:= 1 - 2\bar{\alpha} - \beta_s - 4\bar{\ell}(1 + \bar{\alpha})^2, \\ k_s &:= \bar{\alpha}^2(\beta_s^{-1} + 4\bar{\ell}). \end{aligned} \quad (152)$$

Under (51), $c_s > 0.467$.

Proposition 29 (Exploitation regret with saturation energy). *Under the estimation, Gram, and unchanged-segment hypotheses of Theorem 19, but without clipping feasibility, with conditional probability at least $1 - \delta_0$, simultaneously for every prefix of length n ,*

$$\begin{aligned} \sum_t (y_{t+1} - w_{t+1})^2 &\leq 2\bar{b}^2 \left[\frac{2\bar{Q}_H}{c_s} + \frac{8B_\eta^2}{c_s^2} \log \frac{1}{\delta_0} + \frac{4B_\eta^2}{c_s} \Gamma(n) \right. \\ &\quad \left. + \left(1 + \frac{2k_s}{c_s} \right) \sum_t (d_t^{\text{sat}})^2 \right]. \end{aligned} \quad (153)$$

Proof. Start from the exact identity (79). From (151), $|e_t| \leq |a_t| + d_t^{\text{sat}}$. Young's inequality gives

$$\begin{aligned} 2\alpha a_t e_t &\leq (2\bar{\alpha} + \beta_s) a_t^2 + \frac{\bar{\alpha}^2}{\beta_s} (d_t^{\text{sat}})^2, \\ (\alpha e_t + \eta_{t+1} - a_t)^2 &\leq 4(1 + \bar{\alpha})^2 a_t^2 + 4\bar{\alpha}^2 (d_t^{\text{sat}})^2 + 2\eta_{t+1}^2. \end{aligned}$$

Using $\ell_t \leq \bar{\ell}$ yields

$$Q_{t+1} - Q_t \leq -c_s a_t^2 + 2a_t \eta_{t+1} + 2\ell_t \eta_{t+1}^2 + k_s (d_t^{\text{sat}})^2.$$

The same time-uniform exponential supermartingale as in (81), now evaluated with c_s , and the same elliptical-potential bound give

$$\sum_t a_t^2 \leq \frac{2\bar{Q}_H}{c_s} + \frac{8B_\eta^2}{c_s^2} \log \frac{1}{\delta_0} + \frac{4B_\eta^2}{c_s} \Gamma(n) + \frac{2k_s}{c_s} \sum_t (d_t^{\text{sat}})^2.$$

Finally, $(y_{t+1} - w_{t+1})^2 \leq 2\bar{b}^2 (a_t^2 + (d_t^{\text{sat}})^2)$, which proves (153). $\square$

Proposition 29 gives logarithmic unchanged-segment regret when $\sum_t (d_t^{\text{sat}})^2 = O(\log((T+1)/\delta))$. A global detection guarantee additionally requires a blockwise power gap for the clipped loop or control of its saturation residual on every potential detector block. The non-Schur gap (130) applies to the unclipped auxiliary; Lemma 27 transfers it to the actual output using the boundary test.

Proposition 30 (Saturation lower bound). *For every policy with $|u_t| \leq B_u$,*

$$\mathcal{R}_T \geq b_{\min}^2 \sum_{t=0}^{T-1} (|\lambda(\theta_t)^\top \psi_t| - B_u)_+^2. \quad (154)$$

Consequently, if the realized saturation energy grows faster than $\log T$, logarithmic regret under (3) is impossible, even when the parameter sequence is known.

Proof. For $x_t = \lambda(\theta_t)^\top \psi_t$ and every feasible action,

$$|u_t - x_t| \geq \text{dist}(x_t, [-B_u, B_u]) = (|x_t| - B_u)_+.$$

The gain form (36) gives $(y_{t+1} - w_{t+1})^2 = b_1(\theta_t)^2 (u_t - x_t)^2$. Use $b_1(\theta_t) \geq b_{\min}$ and sum over t . $\square$

L A UNIFIED DETECTOR FOR THE UNION CLASS

Algorithm 3 implements the union-class detector introduced in Section 5. Its two tests use constants supplied offline; the algorithm neither receives the realized pair's class nor computes its spectral radius. The terms Schur and non-Schur refer to the old-gain/new-plant matrix $A_{\text{cl}}(\lambda(\theta^-), \theta^+)$. The plant retains Assumptions 7, 8, and 9; bounded applied inputs therefore preserve the deterministic state bound.

L.1 Complete Algorithm

Algorithm 3: Branch-free union-class PIECE-CD

Input: $p, q, \Theta, T, \delta, B_u, \nu, \sigma_w^2, B_w$; the exploration distribution and the safe model/excitation bounds used in $B_\psi, \kappa_G, E_\theta$; H satisfying (93), (51), and (142); the offline union-class constants defining h_U and $\underline{\Delta}_U$ in (138); the model-class bounds r_B, R_A in (128)–(129) when the non-Schur route is used; the finite-horizon constants in (140); and m_U, D_U from (143). Set $\zeta_H = E_\lambda(H, \delta_0) B_\psi$.

Initialize the restart time $r \leftarrow 0$ and clear the energy buffer.

Loop. While $r < T$, execute one cycle.

- (a) **Explore.** For $t = r, \dots, \min\{r + H - 1, T - 1\}$, draw a fresh exploratory input, apply it, and observe y_{t+1} . If the horizon is reached during this step, stop.
- (b) **Estimate and initialize.** From $\mathcal{I}_r = \{r + r_0, \dots, r + H - 1\}$, compute

$$\begin{aligned} V_r^\theta &:= \nu I_s + \sum_{t \in \mathcal{I}_r} \phi_t \phi_t^\top, & \tilde{\theta}_r &:= (V_r^\theta)^{-1} \sum_{t \in \mathcal{I}_r} \phi_t y_{t+1}, \\ \hat{\theta}_r &:= \Pi_\Theta(\tilde{\theta}_r), & \hat{b}_r &:= b_1(\hat{\theta}_r), & \bar{\lambda}_r &:= \lambda(\hat{\theta}_r). \end{aligned}$$

Initialize

$$V_{r+H} := \nu I_d + \sum_{t \in \mathcal{I}_r} \psi_t \psi_t^\top, \quad \hat{\lambda}_{r+H} := \bar{\lambda}_r.$$

Clear the energy buffer and declare the predictable output-time grid

$$\mathcal{S}_r := \{r + H + h_U + j h_U : j = 0, 1, \dots\}.$$

(c) **Exploit and monitor.** For $t = r + H, r + H + 1, \dots, T - 1$, form

$$z_t := \begin{cases} \hat{\lambda}_t^\top \psi_t, & |(\hat{\lambda}_t - \bar{\lambda}_r)^\top \psi_t| \leq g_H \|\psi_t\|, \\ \bar{\lambda}_r^\top \psi_t, & \text{otherwise,} \end{cases} \quad u_t := \text{clip}_{[-B_u, B_u]}(z_t).$$

Before observing y_{t+1} , mark that output as energy-eligible exactly when $t + 1 \in \mathcal{S}_r$. Apply u_t , observe y_{t+1} , and form ψ_{t+1} . Make the RLS update

$$\begin{aligned} x_{t+1}^\lambda &:= u_t - \frac{y_{t+1}}{\hat{b}_r}, & V_{t+1} &:= V_t + \psi_t \psi_t^\top, \\ \hat{\lambda}_{t+1} &:= \hat{\lambda}_t + V_{t+1}^{-1} \psi_t (x_{t+1}^\lambda - \hat{\lambda}_t^\top \psi_t). \end{aligned}$$

Set the boundary flag

$$B_{t+1} := \mathbf{1}\left\{|\bar{\lambda}_r^\top \psi_{t+1}| > B_u + \zeta_H\right\}.$$

The strict inequality is part of the algorithm. If y_{t+1} was marked eligible, append y_{t+1}^2 to the energy buffer and retain only the most recent m_U stored values. Set

$$E_{t+1} := \begin{cases} \mathbf{1}\left\{|\text{buffer}| = m_U, \frac{1}{m_U} \sum_{z \in \text{buffer}} z \geq \sigma_w^2 + \frac{\underline{\Delta}_U}{2}\right\}, & t + 1 \in \mathcal{S}_r, \\ 0, & t + 1 \notin \mathcal{S}_r. \end{cases}$$

If $B_{t+1} = 1$ or $E_{t+1} = 1$, raise one alarm at output time $t + 1$, discard the current estimates, set $r \leftarrow t + 1$, clear the buffer, and return to step (a). If both flags equal one, still record only one alarm. If neither flag crosses and $t = T - 1$, stop; otherwise continue exploitation.

Eligibility is declared before y_{t+1} is observed, so the sampling grid is predictable. The boundary flag is checked after every exploitation output, not only at energy-sampling times. It is therefore available before the next control is chosen.

L.2 Common Tuning

Under Assumption 26, (138) specifies a common lag h_U and separation $\underline{\Delta}_U$. For the excess innovation power G_h defined in Appendix K, (139) gives

$$G_{h_U}(A) \geq \underline{\Delta}_U$$

for every admitted mismatch $A = A_{\text{cl}}(\lambda(\theta^-), \theta^+)$. The finite-horizon bounds in (140)–(141), the accuracy conditions (142), and the window in (143) therefore provide one calibration for all three detection conditions.

L.3 Detection and Regret Guarantee

Theorem 31 (Branch-free union-class detector and regret guarantee). *Retain Assumptions 7, 8, and 9. Replace global cross-regime clipping feasibility and stable mismatch by closed own-regime feasibility (Assumption 24) and the union stable/direct/non-Schur condition (Assumption 26), with the common separation (139) and finite-horizon bounds (140). Choose H as the smallest feasible integer satisfying (93), (51), and (142); choose m_U, D_U by (143); and read Assumption 12 with D_U in place of D .*

Then Algorithm 3 satisfies, with probability at least $1 - \delta$, no false alarms and exactly one alarm $\hat{\tau}^{(c)}$ for every true change, with

$$\tau^{(c)} < \hat{\tau}^{(c)} \leq \tau^{(c)} + D_U, \quad (155)$$

and

$$\mathcal{R}_T \leq CK_D D_U + (C + 1)K_X H + (C + 1)R_E(T, \delta_0). \quad (156)$$

For fixed problem constants and fixed C , $\mathcal{R}_T = O(\log((T + 1)/\delta))$. More generally,

$$\mathcal{R}_T = O\left((C + 1) \log \frac{T + 1}{\delta}\right). \quad (157)$$

Proof. If $T \leq H$, the dwell condition forces $C = 0$. Algorithm 3 performs only truncated exploration and stops, so $\mathcal{R}_T \leq K_X T \leq K_X H$, which is bounded by the right-hand side of (156). Hence assume $T > H$ below.

Read the restart, reference, and detector events from Theorem 18 and Appendix I using the first composite crossing and D_U in place of D . Fix a potential restart r and test endpoint e for which the reference parameter remains active throughout $[r, e)$. Condition on an $\mathcal{F}_{r+H}$ history with a full exploration block and $\mathcal{G}_r^{\text{est}}$. Since $H > r_0$, the action and boundary-test regressors are own-regime successors throughout the unchanged exploitation interval. Lemma 25 therefore excludes boundary crossings and gives $e_t^2 \leq a_t^2$, as required by Theorem 19. This includes a test at output time $e = \tau$ when the first subsequent change occurs at control time τ .

For each fixed null energy block, the proof of Lemma 15 applies with (Δ, B_o, m) replaced by $(\underline{\Delta}_U, B_U, m_U)$. Indeed, $|v_s| \leq \varepsilon_y \leq \bar{\varepsilon}_y$, $\bar{\varepsilon}_y^2 \leq \underline{\Delta}_U/8$, and $B_U \geq \max\{B_w, \bar{\varepsilon}_y\}$. The conditional crossing probability is bounded by

$$\exp\left(-\frac{m_U \underline{\Delta}_U^2}{32 B_w^4}\right) + \exp\left(-\frac{m_U \underline{\Delta}_U^2}{512 B_w^2 \bar{\varepsilon}_y^2}\right) \leq \frac{\delta_d}{2T^2},$$

where the second term is omitted when $\bar{\varepsilon}_y = 0$. If the actual test is reached, its trajectory agrees with the alarm-suppressed continuation through that test. The bound therefore controls the joint event that the reference succeeds, the test is reached with the stated null history, and the threshold is crossed. It does not condition on survival to the test. Before the first detector failure, every false-alarm candidate has such an unchanged null history.

After a change, any earlier crossing of either test completes detection. Otherwise, the absence of a boundary crossing bounds the old ideal action's saturation deficit by $2\zeta_H$ and the applied-action error by $\bar{\varepsilon}_u$. Lemma 27 then compares each output in the complete post-change block with its h_U -step auxiliary. These auxiliaries are defined on the full alarm-suppressed continuation, so Lemma 14 applies without conditioning on survival of either test. The event inclusion in Lemma 28 gives a composite alarm by $\tau + D_U$ except with conditional probability $\delta_d/(4T^2)$ on each eligible $\mathcal{F}_\tau$ history.

The null bounds are unioned over at most T^2 restart–endpoint pairs. At a fixed true change, the successful-history active-cycle events are disjoint $\mathcal{F}_\tau$ -measurable subsets of the corresponding $\mathcal{A}_{r,\tau}$. Multiplying the conditional missed-detection bound by these indicators and summing over r costs at most $\delta_d/(4T^2)$ for that change. The chronological induction in Theorem 18 and Appendix I therefore applies with total detector failure probability at most δ_d ; boundary crossings add no null failure probability on successful-reference histories. Controller failures cost at most $3(T + 1)\delta_0$, so the allocation in (48) is below δ . The delay intervals cost at most $CK_D D_U$, exploration costs at most $(C + 1)K_X H$, and unchanged exploitation prefixes cost at most $(C + 1)R_E(T, \delta_0)$. This proves (155)–(156).

It remains to check the order. Because $\bar{\varepsilon}_u = 7L_\lambda B_\psi E_\theta(H, \delta_0)$ and $\bar{\varepsilon}_y = 7\bar{b}L_\lambda B_\psi E_\theta(H, \delta_0)$, the detector conditions follow from

$$E_\theta(H, \delta_0) \leq \epsilon_{\mathcal{U}}^{\text{uni}} := \min \left\{ \frac{b_{\min}}{16}, \frac{\sqrt{\Delta_{\mathcal{U}}/8}}{7\bar{b}L_\lambda B_\psi}, \frac{B_{\mathcal{U}}}{7\bar{b}L_\lambda B_\psi}, \frac{\Delta_{\mathcal{U}}}{112B_{\mathcal{U}}L_{u,\mathcal{U}}L_\lambda B_\psi} \right\}, \quad (158)$$

Omit the last entry when $L_{u,\mathcal{U}} = 0$. Appendix J and (143) then prove (157). $\square$