CRITICAL DECISION CAPACITY: BOTTLENECK VS. SUBSTITUTABLE DEGRADATION UNDER ACTION AVAILABILITY

Anonymous authors
Paper under double-blind review

ABSTRACT

Action availability is a structural resource: systems whose useful behaviour depends on conjunctive bottlenecks can collapse abruptly under random loss of that resource, while substitutable structure degrades smoothly. Critical Decision Capacity (CDC) is not merely a name for a failure threshold f_c: it is the CAC-linked critical capacity used as an *intervention* objective, with estimators $\mathrm{CDC}_{\mathrm{struct}} = 1 - f_c$, $\mathrm{CDC}_{\mathrm{value}}$, and optional $\mathrm{CDC}_{\mathrm{norm}}$ (coinciding only under stated assumptions). We prove a *sufficient-condition* dichotomy for path MDPs: sparse digraph bottlenecks admit a directed-bond percolation threshold (sharp-class mechanism), while submodular structure with bounded marginals is equi-Lipschitz (no discontinuous jump). Path-MDP hardening and Leduc blind transfer (frozen $G_\star$ rankings under degree-biased / vs-random failure) show the design chain; poker targeted>random fragility remains Tier B, not a competitive-IIG theorem.

1 INTRODUCTION

Self-play multi-agent systems concentrate strategic value at a small number of pivotal information states (Kujur, 2025). When action availability is structurally constrained—by actuator faults, capability sandboxing, API rate limits, or regulatory restrictions—performance degrades far more under *targeted* removal than under random unavailability of equal support. Prior work introduced contingent action capacity (CAC_w, CAC_v) as a diagnostic of that fragility and proved a parameter-class-independent damage lower bound in terms of $\mathrm{CAC}_v^{(k)}$ (Kujur, 2025).

That paper established a *phenomenon*. It did not ask what structural property separates abrupt vs. smooth degradation under random action retention, whether a critical-capacity metric can serve as an *intervention* objective, or whether imperfect-information probes show analogous fragility. Measuring a performance-collapse threshold f_c alone does not answer those questions: the scientific object here is the chain local $\mathrm{CAC}_v \to$ critical structure $\to$ CDC estimators $\to$ intervene $\to$ OOD transfer (Fig. 3). We answer on two evidence tiers: theorems for path MDPs; empirics and a conjecture for poker.

Contributions.

1. **(Proved) Sufficient-condition dichotomy.** CDC is a *concept*—the CAC-linked critical capacity for design—not a relabelling of f_c. Estimators $\mathrm{CDC}_{\mathrm{struct}} = 1 - f_c$ (topology), $\mathrm{CDC}_{\mathrm{value}}$ (ties criticality to local gaps), and optional $\mathrm{CDC}_{\mathrm{norm}}$ coincide only under stated assumptions (Definition 1; Prop. 1). Bottleneck path MDPs admit a directed-bond percolation threshold under random retention (Theorem 1; percolation is the sharp-class *mechanism*, not the sole novelty). Submodular structure with bounded initial marginals is equi-Lipschitz—a *sufficient* no-jump condition (Theorem 2; Remark 1).
2. **(Tier A+B design) CDC as intervention objective.** Path-MDP hardening raises f_c vs. random/centrality (Fig. 30; Proposition 5). On Leduc, threat-aligned protect is a controlled check (Fig. 33); *blind transfer* freezes those rankings and still beats baselines under degree-biased deletion + vs-random (Fig. 3)—not Theorem 1.

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

3. **(Empirical) Poker fragility.** Leduc-$N \in \{5, 6, 7\}$ with QL/NFSP/CFR/DQN; targeted>random at matched budget (`algo_multiseed.json`).
4. **(Conjecture) $G_\star$ protocol.** Pre-registered predictive $G_\star$: held-out Spearman≈ 0.77; opp abstraction load-bearing (Fig. 21).

2 Related Work

Adversarial / robust RL typically perturbs observations, rewards, or executed actions within an intact action set (Huang et al., 2017; Gleave et al., 2020; Tessler et al., 2019); invalid-action masking is usually a safety aid (Huang & Ontañón, 2022).

Availability literatures vs. this paper. Three related lines ask different questions. *Sleeping* / stochastic action-set learning asks how to *learn or plan* when the legal set A_t varies over time (Kanade et al., 2009; Boutilier et al., 2018). Constrained MDPs ask for feasible policies under long-run resource limits (Altman, 1999; Zahavy et al., 2021). *This work* asks when *structural* loss of available decisions causes abrupt vs. bounded value collapse, and whether that critical capacity *selects interventions* (CDC as design, not a learning algorithm under availability). Independent-bond Theorem 1 is an idealized abstraction of motivating actuator/API/sandbox failures; correlated actuator-group probes (Appendix A.10) address non-i.i.d. structure without claiming the theorem covers those processes.

Network interdiction and reliability treat edge/action removal as an optimization or survivability problem (Wood, 1993; Albert et al., 2000; Callaway et al., 2000) rather than a sharp-vs-smooth dichotomy plus intervention objective. Directed bond percolation and network attack theory (Newman et al., 2001; Karp, 1990; Callaway et al., 2000; Albert et al., 2000; Hinrichsen, 2000) ground our path-MDP $G_\star \mapsto f_c$ mapping (Tier A mechanism); targeted-vs-random poker probes are a separate adversarial-fragility axis (Tier B). Finite-size scaling (Fisher, 1967; Cardy, 1996) and susceptibility early-warning (Scheffer et al., 2009) appear as diagnostics. Contingent action capacity (Kujur, 2025) is the local fragility currency; CDC is its critical / design-level counterpart—not a new name for f_c alone. Deep imperfect-information solvers (e.g. NFSP/CFR lines) optimise equilibrium quality; our probes are *orthogonal*—fragility under action masking, not SOTA exploitability. Extended comparisons (parametric MDPs, shielding, curriculum) are in Appendix A.13; Hanabi-scale cooperative IIGs are future work.

3 Problem Formulation

Game and availability. Let Γ be a two-player extensive-form (or single-controller path) game with information sets $\mathcal{I}$ and legal actions $A(h)$ at $h \in \mathcal{I}$. An availability constraint M maps each h to a nonempty $M(h) \subseteq A(h)$.

Two threat models. We distinguish the process that generates M, because theory and experiment must match:

- **Random retention** M_p^{rand}: each action-labelled bond on the transition graph is retained independently with probability p (equivalently removed with $f = 1 - p$). This is the object of Theorem 1 and Proposition 7.
- **Targeted budget-f** M_f^{tgt}: at most $k = \lfloor f\,|\mathcal{I}_0| \rfloor$ information states are constrained, chosen by a CAC_v-greedy (or critical-bond / centrality) rule, removing one action at each (Kujur, 2025). This is an empirical stress test; percolation f_c is only a conservative bound (Proposition 3).

Empty-action handling. Structural Φ uses pure directed-bond deletion (empty out-sets allowed). Policy eval retains one residual action by default; ablations with `terminate`/`allow_empty` leave conclusions unchanged (Appendix A.11) and do not bias Theorem 1 or Proposition 7.

Control and order parameters. The control parameter is the capability-removal fraction $f = 1 - p$ (or budget fraction for targeted masks). The order parameter is the normalised achievable

value

$$\Phi(f) \;=\; \frac{V_0(M_f) - V_{\min}}{V_0(\emptyset) - V_{\min}} \in [0,1], \tag{1}$$

where V_0 is expected return for player 0. Unless stated, primary plots use M^{rand}; $V_{\min}$ is the empirical sweep floor (percolation convention $\Phi \to 0$ at collapse), ablated in Section 5. $\Phi = 1$ is unconstrained play; $\Phi \to 0$ is strategic collapse.

Contingent capacity and CDC. Following Kujur (2025),

$$\mathrm{CAC}_w(M) = \sum_{h \in \mathcal{I}_0} \rho(h)\, \mathbf{1}\big[|M(h)| > 1\big], \tag{2}$$

$$\mathrm{CAC}_v(M) = \sum_{h \in \mathcal{I}_0} \rho(h)\, \delta(h)\, \mathbf{1}\big[|M(h)| > 1\big], \tag{3}$$

with reach $\rho(h)$ and local gap $\delta(h) = Q(h, a^\star) - Q(h, a_{\mathrm{forced}})$.

Definition 1 (Critical Decision Capacity)**. CDC** is the conceptual critical capacity of action availability under a stated threat model—*not* a single scalar invariant and *not* merely a renaming of the failure threshold f_c. Measuring where Φ crosses a collapse level yields f_c; CDC adds the link from local CAC_v to that critical regime and onward to *intervention* selection (Section 5). Let f_c be the critical removal fraction at which Φ crosses a fixed collapse level (we use $1/2$; ablations use $\{0.3, 0.5, 0.7\}$ and susceptibility / $|d\Phi/df|$ peaks). Operational estimators:

$$\mathrm{CDC}_{\mathrm{struct}} := 1 - f_c \qquad \text{(topology estimator)}, \tag{4}$$

$$\mathrm{CDC}_{\mathrm{value}} := \mathbb{E}_{M \sim M^{\mathrm{rand}}_{f_c}}\big[\mathrm{CAC}_v(M)\big] \qquad \text{(value-weighted; CAC link)}, \tag{5}$$

and optionally $\mathrm{CDC}_{\mathrm{norm}} = \mathrm{CDC}_{\mathrm{value}}/\mathrm{CAC}_v(\emptyset) \in [0,1]$. **Do not conflate them:** $\mathrm{CDC}_{\mathrm{struct}}$ and $\mathrm{CDC}_{\mathrm{norm}}$ coincide *only* under Assumption 1 with uniform $\rho\delta$ on the cut; otherwise report which estimator and why (Prop. 1; Section 6). Multi-action nodes use Monte-Carlo $\mathrm{CDC}_{\mathrm{value}}$. (ρ, δ) from the unmasked agent (Section 5).

Two remarks are needed before the equivalence gap can be stated cleanly, because CAC_v in (3) is an indicator on the *multi-action* event $\{|M(h)| > 1\}$, not a per-bond survival variable, and in general $\mathbb{E}\,\mathbf{1}[|M(h)| > 1] \neq \rho_{\mathrm{ret}}$.

Assumption 1 (Single-critical-bond cut)**.** On the critical cut, each contributing information state h has exactly one *critical* action $b(h)$ whose removal collapses h from $|M(h)| > 1$ to $|M(h)| \leq 1$ (a bottleneck decision), and the residual actions at h are retained. Equivalently, the collapse indicator equals the survival of a single bond: $\mathbf{1}[|M(h)| > 1] = Z_h$, where $Z_h = \mathbf{1}[b(h)\text{ retained}] \sim \mathrm{Bernoulli}(\rho_{\mathrm{ret}})$.

Proposition 1 (Quantified equivalence gap under Assumption 1)**.** *Let the critical cut carry m information states with value weights $w_h = \rho(h)\delta(h) > 0$ satisfying Assumption 1, and retain each critical bond independently with probability $\rho_{\mathrm{ret}} = 1 - f_c$. Write $\mathrm{CV}(w) = \mathrm{std}(w)/\mathrm{mean}(w)$. Then $\mathbb{E}[\mathrm{CDC}_{\mathrm{norm}}] = \mathbb{E}[\mathrm{CDC}_{\mathrm{struct}}] = \rho_{\mathrm{ret}}$ (both unbiased), and the realized discrepancy obeys*

$$\sqrt{\mathbb{E}\big(\mathrm{CDC}_{\mathrm{norm}} - \mathrm{CDC}_{\mathrm{struct}}\big)^2} \;=\; \mathrm{CV}(w)\,\sqrt{\frac{\rho_{\mathrm{ret}}(1-\rho_{\mathrm{ret}})}{m}}. \tag{6}$$

Hence the two forms coincide exactly iff $\mathrm{CV}(w) = 0$ (uniform $\rho\delta$), and the discrepancy is controlled by the weight spread otherwise.

Proof sketch. Under Assumption 1 the collapse indicator is a single i.i.d. Bernoulli bond, so both the structural and normalised-value estimators have mean ρ_{ret} and their difference has variance $\rho_{\mathrm{ret}}(1-\rho_{\mathrm{ret}})\mathrm{CV}(w)^2/m$, giving (6) exactly (full derivation in Appendix A.9). □

General multi-action nodes. When Assumption 1 fails (a state needs $k(h) \geq 2$ simultaneous deletions), $\mathbb{E}\,\mathbf{1}[|M(h)| > 1] \neq \rho_{\mathrm{ret}}$, so Proposition 1 does not apply verbatim; there we report $\mathrm{CDC}_{\mathrm{value}} = \mathbb{E}_M[\mathrm{CAC}_v(M)]$ directly by Monte-Carlo (exact for the indicator definition). The single-critical-bond case is precisely the path-MDP bottleneck regime where the percolation mapping is sharpest, so the load-bearing case is covered analytically and the rest empirically. We validate (6) under uniform/skewed/Pareto weights in Section 5.

Directed transition graph. Associate to Γ a *directed* transition graph $G = (V, E)$ whose nodes are physical states (path MDPs) or information states, and whose edges are labelled by actions. Capability removal is directed bond deletion. Reachability of a winning/goal set W from the start s is the structural analogue of Φ in the path-MDP setting where Theorem 1 applies exactly (Assumption 2). For imperfect-information games we construct an approximate $G_\star$ empirically under a frozen protocol with held-out checks (Appendix A.10); mask-aligned predictors yield moderate Spearman correlation (tune≈ 0.64, held-out≈ 0.77) with aligned damage, giving construction-sensitive support for Conjecture 1 without elevating it to a Leduc theorem.

4 THEORY

We inherit the additive damage bounds of Kujur (2025) (Propositions 2–4 therein): for a budget-k mask with support S, $\Delta V_0(M) \gtrsim \sum_{h \in S} \rho(h)\delta(h)$, and $\mathrm{CAC}_v^{(k)}$ lower-bounds worst-case damage independent of the adversary class. Here we lift those bounds to a critical-transition law under *random* retention, with a separate proposition for targeting.

4.1 REACHABILITY–VALUE LINK

Lemma 1 (Lipschitz link)**.** *Consider a single-controller path MDP on directed graph G with bounded terminal reward $r \in \{0, R_{\max}\}$ that equals $R_{\max}$ iff the trajectory visits W and 0 otherwise, and zero intermediate reward. Let π be any (possibly randomised) policy using only open bonds, and let q be the probability that s reaches W under π on the open subgraph. Then the value satisfies $V = R_{\max}\, q$, hence $|V - V'| \le R_{\max}\,|q - q'|$. In particular, with $R_{\max} = 1$ and Φ normalised by unconstrained value, Φ coincides with normalised reach probability.*

Proof. Every trajectory either hits W (payoff $R_{\max}$) or not (payoff 0); expectation yields $V = R_{\max}\, q$. The Lipschitz claim is immediate. □

Proposition 2 (Bounded / shaped rewards)**.** *Drop the binary assumption and let the return of a trajectory be any bounded measurable $g \in [0, R_{\max}]$ (discounted stepwise rewards, shaping, etc.) with the single structural property that reaching W is necessary for the maximum attainable return, i.e. $g \le R_{\max}\,\mathbf{1}[\text{hit } W]$ pointwise and $\mathbb{E}[g \mid \text{hit } W] \ge \underline{r} > 0$. Let q be the $s \to W$ reach probability on the open subgraph. Then*

$$\underline{r}\, q \;\le\; V \;\le\; R_{\max}\, q, \tag{7}$$

so V is sandwiched between two constant multiples of reachability. Consequently the normalised order parameter $\Phi = V/V(\emptyset)$ obeys $(\underline{r}/R_{\max})\,\Phi_{\mathrm{reach}} \le \Phi \le (R_{\max}/\underline{r})\,\Phi_{\mathrm{reach}}$. Multiplicative bounds transfer threshold location *and the* dichotomy class *(existence of a percolation threshold vs. equi-Lipschitz retention dependence) from Φ_{reach} to Φ; they do* not *transfer critical exponents, finite-N nonanalyticity, or every notion of "phase transition."*

Proof. The lower bound is $V = \mathbb{E}[g] \ge \mathbb{E}[g\,\mathbf{1}[\text{hit } W]] \ge \underline{r}\, q$; the upper bound is $V \le R_{\max}\, q$ from $g \le R_{\max}\mathbf{1}[\text{hit } W]$. A discontinuity (or Lipschitz continuity) of q in p transfers to V up to the constant factors, since both bounds are exactly proportional to q. □

Recommended order parameter (shaped rewards). When rewards are shaped/discounted we use $\Phi = \big(V - V_{\min}\big)/\big(V(\emptyset) - V_{\min}\big)$ with V the expected discounted return; (7) guarantees that the location of the transition (its dichotomy class) is inherited from reachability, even though the *shape* of Φ is reward-dependent. We confirm this empirically on discounted synthetic games in Section 5.

Remark 1 (Sharp vs. no-jump in this paper)**.** *Sharp:* under Assumption 2, existence of $p_c \in (0, 1)$ for $s \to W$ reachability / Φ under i.i.d. bond retention (Theorem 1)—*not* finite-N nonanalyticity (Section 6). *No-jump ("smooth" in our prose):* equi-Lipschitz continuity of $\mathbb{E}[V]$ (and of Φ under Assumption 4) in retention p under Assumption 3 (Theorem 2)—*no discontinuous jump*, not "no critical phenomenon" and not "visually gradual"; when C grows with width, finite-size slopes can steepen (Fig. 26).

4.2 EXISTENCE OF A SHARP THRESHOLD (PATH MDPS)

Assumption 2 (Giant $s \to W$ conditions)**.** Throughout Theorem 1 we assume:

1. **Locally tree-like digraph.** $G_\star$ is drawn from a sparse directed ensemble with finite mean out-degree $\langle k \rangle > 1$ and vanishing density of short directed cycles (configuration-model / ER-like digraphs in the large-N limit).
2. **Absorbing terminal set.** W is an absorbing winning set: once visited, play ends with reward $R_{\max}$.
3. **Reachability order parameter.** The structural order parameter is the probability that the open subgraph admits a directed $s \to W$ path (equivalently: s belongs to a giant out-component that intersects W). This is *not* undirected connectivity, nor weak connectivity alone.
4. **Bond retention.** Each directed bond is retained independently with probability $p = 1 - f$ (Bernoulli product measure).

Under these conditions, survival of an $s \to W$ path is governed by the directed branching-process fixed point on the locally tree-like unfolding of $G_\star$ (Newman, 2010; Callaway et al., 2000).

Theorem 1 (Existence and predicted CDC under random retention)**.** *Under Assumption 2, there exists $p_c \in (0, 1)$ such that in the $|V| \to \infty$ limit:*

1. *$\Phi(p) = 0$ for $p < p_c$ and $\Phi(p) > 0$ for $p > p_c$;*
2. *p_c equals the directed bond-percolation threshold on $G_\star$ (for ER-like digraphs in the mean-field regime, $p_c \sim 1/\langle k \rangle$);*
3. *near p_c^+, $\Phi(p) \asymp (p - p_c)^\beta$ for a critical exponent $\beta > 0$ in the ensemble's universality class (ordinary mean-field percolation: $\beta = 1$; directed percolation (DP) differs (Stauffer & Aharony, 1994; Hinrichsen, 2000)).*

Consequently $f_c = 1 - p_c$ and $\mathrm{CDC} = \mathbb{E}_{M \sim M_{f_c}^{\mathrm{rand}}}[\mathrm{CAC}_v(M)]$ is predicted by $G_\star$ (structural proxy $\mathrm{CDC} = 1 - f_c$ for uniform bond capacity).

Proof sketch. Map retained actions to open directed bonds on $G_\star$; tree-like digraphs percolate at $p_c \sim 1/\langle k \rangle$ (Newman et al., 2001; Karp, 1990) (Appendix A.8). Lemma 1 identifies Φ with normalised reachability. □

Scope. Thm 1 applies to path MDPs under Ass. 2; imperfect-info $G_\star$ has moderate Spearman support under a frozen protocol with held-out checks (Conjecture 1; not a Leduc theorem). Primary sharp-class evidence uses sparse digraphs with $p_c \in (0, 1)$. **Rule of thumb for Ass. 3:** keep C N-independent (normalize coverage) for the smooth Lipschitz regime; if unnormalized $C \propto$ width, expect steepening toward critical-like slopes (Fig. 26).

Proposition 3 (Targeted removal shifts f_c downward)**.** *If Φ is nonincreasing in deleted bonds, then $f_c^{\mathrm{tgt}} \le f_c^{\mathrm{rand}}$ whenever both exist, so percolation f_c is a conservative upper bound under targeting.*

Proposition 4 (Path-priority gap on digraph path MDPs)**.** *Let G be a digraph path MDP with a shortest $s \to W$ path L of length $\ell = |L|$ and $|E|$ bonds. Under independent random retention at rate $p = 1 - f$, path survival is p^ℓ. Path-priority targeting that spends mass f on L first yields survival at most $\big(1 - \min(1, f|E|/\ell)\big)_+^\ell$. Hence at matched deletion mass $f \in (0, 1)$ with $\ell \ll |E|$,*

$$\Delta\Phi(f) \;\ge\; p^\ell - \big(1 - \min(1, f|E|/\ell)\big)_+^\ell \tag{8}$$

(analytic lower sketch; Monte Carlo confirms on sparse/scale-free ensembles, `scale_free.json`). Combined with Proposition 3, this gives a quantitative *targeted gap beyond mere f_c ordering.*

Proposition 5 (Parallel-route redundancy shifts f_c)**.** *Consider L serial cuts, each with r independent length-2 routes whose bonds are retained i.i.d. with probability $p = 1 - f$. A cut survives iff at least one route is fully open, so cut survival is $1 - (1 - p^2)^r$ and end-to-end survival is*

$$\Phi(f) = \big(1 - (1 - p^2)^r\big)^L. \tag{9}$$

The half-collapse $f_c(r)$ solving $\Phi(f_c) = \frac{1}{2}$ is strictly increasing in r for fixed L (more parallel routes buy a higher deletion budget before collapse). Monte Carlo on this family matches (9) *(`cdc_harden_loop.json`).*

Proposition 6 (Short-cycle perturbation)**.** *On a tree-like backbone with clustering c, the Bethe estimate $\hat{f}_c = 1 - 1/\langle k \rangle$ satisfies $|f_c - \hat{f}_c| = O(c)$; the sharp/smooth classification is unchanged (Fig. 24).*

4.3 No-go: equi-Lipschitz retention (no discontinuous jump)

Assumption 3 (Uniformly bounded initial marginals)**.** There exists a constant $C < \infty$, *independent of system size $N = |A|$*, such that

$$\sum_{a \in A} \Delta_a(\emptyset) \;\le\; C, \qquad \Delta_a(S) = V(S \cup \{a\}) - V(S). \tag{10}$$

For layered cut-coverage we take the normalised value $V = (1/L) \sum_{\ell=1}^{L} \mathbf{1}[S \cap C_\ell \neq \emptyset]$ with fixed layer width w; then $C = w$ independent of the number of layers L. Without this normalisation, $\sum_a \Delta_a(\emptyset) = Lw \to \infty$ as width or depth grows and the Lipschitz bound below fails (tightness; appendix certificate).

Assumption 4 (Uniformly positive normalisation gap)**.** There exists $c > 0$, independent of system size N, such that $V_N(A) - V_{\min} \ge c$ for every N in the family under consideration. (Holds for our constructive classes in the paragraph below.)

Theorem 2 (Equi-Lipschitz / no discontinuous jump under bounded marginals)**.** *Let $V : 2^A \to \mathbb{R}$ be monotone submodular and satisfy Assumption 3. Write $S_p = \{a \in A : Z_a = 1\}$ with independent $Z_a \sim \mathrm{Bernoulli}(p)$ and $f = 1 - p$. Then:*

1. *For any finite $|A|$, $\mathbb{E}[V(S_p)]$ is a multilinear polynomial in p (hence C^∞ on $[0, 1]$).*
2. *Under Assumption 3, $\left|\frac{d}{dp}\mathbb{E}[V_N(S_p)]\right| \le C$ uniformly in N, so any pointwise limit of $\mathbb{E}[V_N]$ is Lipschitz and admits* no discontinuous jump *in p.*
3. *Under Assumptions 3–4, the normalised order parameter $\Phi_N(p) = (\mathbb{E}[V_N(S_p)] - V_{\min})/(V_N(A) - V_{\min})$ is likewise equi-Lipschitz in p (Lipschitz constant $\le C/c$), hence any pointwise limit Φ_∞ is Lipschitz with no discontinuous jump.*

This is a sufficient *no-jump condition for the stated set-function family—not a characterisation of phase transitions in general RL environments. It rules out a percolation-style* discontinuity *under the stated assumptions; it does* not *rule out continuous critical phenomena (e.g. a kink) or steep finite-size crossovers (Remark 1).*

Proof sketch. Independence gives the multilinear expansion $\mathbb{E}[V(S_p)] = \sum_T V(T) p^{|T|}(1-p)^{|A \setminus T|}$ (analytic on $[0, 1]$); submodularity bounds the derivative by $\sum_a \Delta_a(\emptyset) \le C$ uniformly in N. Assumption 4 transfers the bound to Φ_N with factor $1/c$. Full proof, including the tightness argument, in Appendix A.9. □

Constructive game classes. Two equi-Lipschitz families satisfy Assumptions 3–4: layered DAGs with cut-coverage $V = (1/L) \sum_\ell \mathbf{1}[\text{cut hit}]$ ($C = w$, $V(A) - V_{\min} = 1$), and facility location $V(S) = \sum_j \max_{i \in S} w_{ij}$ (empty-set marginal sum $O(1)$, independent of length; verified in `smooth_classes.json`). The sharp counterpart is the conjunctive chain $V = \prod_i \mathbf{1}[e_i \text{ open}]$, which fails submodularity. The Lipschitz bound gives min transition width $\ge 1/C$, so the no-jump guarantee degrades linearly as C grows and a zero-width transition requires $C \to \infty$ (steep continuous crossovers remain possible; Remark 1).

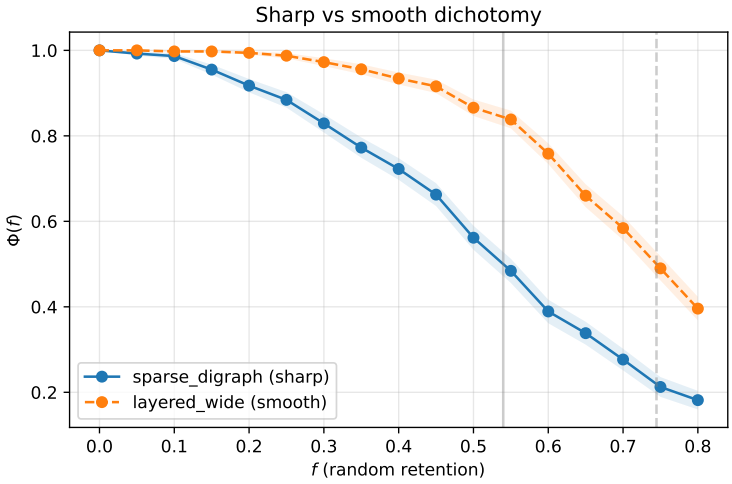


Figure 1: Sparse vs. layered dichotomy (monotone Φ).

4.4 Finite-size scaling (pointer)

Proposition 7 (Finite-size scaling ansatz)**.** *In the sharp regime, an operational ansatz*

$$\Phi(f, N) = N^{-\beta/\nu} G\big((f - f_c)N^{1/\nu}\big) \tag{11}$$

is a diagnostic only (Fisher, 1967; Cardy, 1996); we do not claim a new universality class from small-N β. Held-out / pseudocritical collapse: Appendix Fig. 11.

Dichotomy summary. We identify a structural *sufficient* condition separating discontinuous collapse from a no-jump regime—not a characterisation of all decision-system phase transitions. Only the non-submodular / bottleneck class under Theorem 1 exhibits a percolation *threshold*; substitutable games under Theorem 2 admit no *discontinuous jump* (equi-Lipschitz), while continuous criticality or steep crossovers are not ruled out. That distinction—not percolation alone—is the conceptual centre.

5 Experiments

All code is from-scratch and CPU-only. Primary plots use **random bond retention**; targeted attacks are comparative (Proposition 3). Synthetic Φ: $s \to W$ reachability with $\geq 10^3$ Monte Carlo graphs per f, then PAVA isotonic projection. Environments: sparse digraph (sharp), layered wide (smooth), navigation grids, and Kuhn/Leduc-N damage probes (not Thm 1). Ablations live in Appendix A.10.

Pillar I: substitutability dichotomy. Figure 1: sparse digraphs collapse earlier and more sharply than layered graphs at matched N under *random* bond retention (Tier A empirics for Theorems 1–2); percolation f_c tracks measured half-collapse (appendix Fig. 9). Matched-mass path-priority / centrality / min-cut all lower f_c vs. random (appendix Fig. 7; Prop. 3).

Pillar II: targeted vs. random adversarial fragility (poker). Figure 2: multi-seed damage $\Delta = V_{\text{base}} - V_{\text{mask}}$ for QL/NFSP/CFR and a small-MLP DQN, scaled to Leduc-$N \in \{5, 6, 7\}$ (≥ 5 seeds, 95% CIs; `algo_multiseed.json`). At $f = 0.2$, targeted−random gaps remain large at Leduc-7 (QL ≈ 0.25, NFSP ≈ 0.31, CFR ≈ 0.45, DQN ≈ 0.17) alongside Leduc-5/6 tabular gaps (Leduc-5: NFSP ≈ 0.62, CFR ≈ 0.57; Leduc-6: NFSP ≈ 0.42, CFR ≈ 0.52). Self-play (`selfplay_poker.json`): Leduc-7 NFSP self gap ≈ 0.41; QL self positive but smaller. Cross-figure gap magnitudes differ by *eval opponent* (vs-random vs. self / frozen-CFR), not plotting bugs—see Table 1 / Fig. 5. This is Tier B adversarial fragility—*not* a proof of percolation in Leduc. Random-mask damage can be noisy; targeted damage stays predominantly positive.

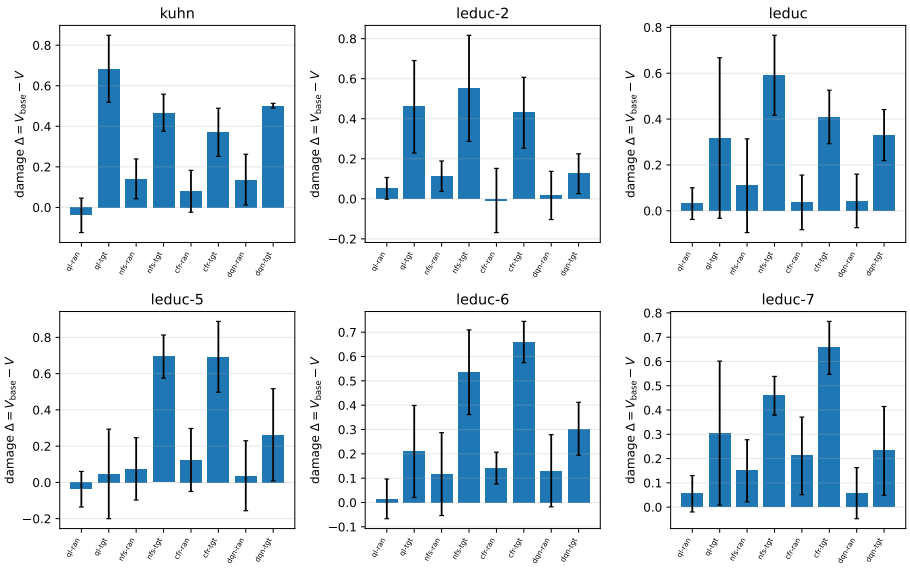


Figure 2: Poker damage (random vs. targeted), Leduc-$N \geq 5$ and DQN.

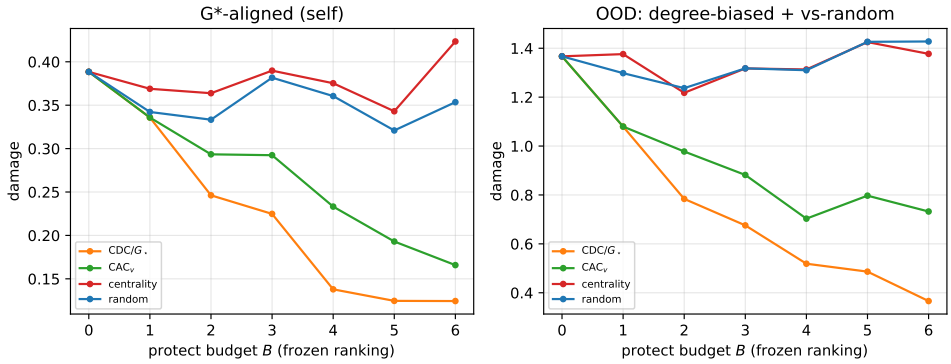


Figure 3: Blind transfer: frozen protect rankings under G*-aligned (left) vs. degree-biased + vs-random (right). Leduc-3; mean over seeds.

Pillar III: CDC as an intervention objective. **Tier A (path-MDP harden).** Budgeted add-one-forward-edge: path-criticality raises f_c by ≈ 0.11 over $B{=}8$ vs. random ≈ 0.05 (Fig. 30; Proposition 5). **Tier B (Leduc protect).** Threat-aligned immunize-under-$G_\star$ is a *controlled* check (Δdamage≈ 0.26 for $G_\star$-guided vs. ≈ 0 random; Fig. 33)—same threat as the predictor. **Blind transfer.** Figure 3: freeze those rankings, then evaluate under *degree-biased* maskable-edge deletion + vs-random opponent (neither used to build $G_\star$). Guided Δdamage≈ 1.00 vs. near-zero/negative for centrality/random (`leduc_blind_transfer.json`). This answers co-design: the ranking is not only useful under the construction threat. Still not Theorem 1.

Ablations. Policy-strength, retrain-under-mask, adaptive CAC$_v$, residual-action, Ass. 3 steepening, collapse/$V_{\min}$ sensitivity, and PAVA/bootstrap f_c checks: Appendix A.13 (Figs. 23, 15–18).

6 Discussion and Limitations

Two evidence tiers. **Tier A (proved):** sufficient-condition dichotomy and CDC estimators for path MDPs (Theorems 1–2). **Tier B:** poker fragility, threat-aligned protect, and blind transfer under an independent failure process—not a competitive-IIG theorem.

Novelty location. CDC adds the CAC→critical→design link, not a new name for f_c. Theorem 1 instantiates directed-bond percolation as the sharp-class mechanism; the scientific object is the bottleneck-vs-substitutable distinction plus that intervention pipeline—not new percolation mathematics. Theorem 2 is a sufficient no-jump condition for bounded-marginal submodular set functions, not a general RL phase-transition theory.

Design and co-design. Threat-aligned Leduc protect (Fig. 33) alone cannot refute representation/threat co-design. Blind transfer (Fig. 3) freezes rankings and evaluates under degree-biased deletion + vs-random: guided still dominates random/centrality. Pipeline (Tier B): local CAC $\rightarrow G_\star$ ranking $\rightarrow$ intervene $\rightarrow$ gain under an OOD failure process. Synth harden remains Tier A (Fig. 30).

Which CDC to report. **CDC** = concept; $\mathrm{CDC}_{\mathrm{struct}}$ / $\mathrm{CDC}_{\mathrm{value}}$ / $\mathrm{CDC}_{\mathrm{norm}}$ = estimators that coincide only under Ass. 1 + uniform $\rho\delta$. Report which, and both when $\mathrm{CV}(\rho\delta)$ is large.

Limits. Poker $\neq$ Thm 1. $G_\star$ remains protocol-dependent (Fig. 21). No finite-N non-analyticity; no DP universality from small-N β. Leduc remains a conjecture:

Conjecture 1. For extensive-form zero-sum games whose value concentrates on a sparse directed info-state backbone, random action retention induces an approximate collapse near the bond-percolation threshold of a suitably defined $G_\star$.

Conclusion. **Tier A:** bottleneck vs. substitutable path MDPs with CDC as a design metric. **Tier B:** poker fragility; Leduc protect rankings that transfer under an independent failure process (Fig. 3). Together with Kujur (2025), CDC complements local CAC diagnosis.

AI use statement

In this work, we used generative AI tools (Cursor/LLM-assisted coding and writing) for implementing experiment scripts and figure-generation pipelines, and for drafting and editing parts of the manuscript. We have not used generative AI tools for generating synthetic datasets, formulating mathematical claims or proofs, proposing primary research hypotheses, or interpreting experimental results without author verification; human authors designed the theory, experiments, and scientific claims. Literature search, reference formatting, and readability editing are additionally AI-assisted. All AI-generated code was run, reviewed, and checked against committed `results/*.json`; all claims and numbers were verified by the authors. We take responsibility for the final content of this work, including text, claims, and artifacts produced with the aid of generative AI.

See the ICLR 2027 AI Policy for Authors for more details.

Reproducibility statement

All experiments are seed-fixed, CPU-only, and from-scratch in this repository. One-command reproduction: `pip install -r requirements.txt` followed by `python -m experiments.run_all` and `python -m experiments.generate_figures`. Committed JSON under `results/` matches the submitted figures; Appendix A.10 maps scripts to artifacts. Full details are in `REPRODUCIBILITY.md` at the repository root. Proofs and assumption checks appear in Appendix A.9 and related robustness sections.

References

Réka Albert, Hawoong Jeong, and Albert-László Barabási. Error and attack tolerance of complex networks. *Nature*, 406:378–382, 2000.

Eitan Altman. Constrained Markov decision processes: Stochastic modeling. *CRC Press*, 1999.

Craig Boutilier, Alon Cohen, Avinatan Hassidim, Yishay Mansour, Ofer Meshi, Martin Mladenov, and Dale Schuurmans. Planning and learning with stochastic action sets. *International Joint Conference on Artificial Intelligence (IJCAI)*, 2018.

Duncan S Callaway, M E J Newman, Steven H Strogatz, and Duncan J Watts. Network robustness and fragility: Percolation on random graphs. *Physical Review Letters*, 85(25):5468, 2000.

John Cardy. *Scaling and Renormalization in Statistical Physics*. Cambridge University Press, 1996.

Colin Cooper and Alan Frieze. The size of the largest strongly connected component of a random digraph with a given degree sequence. *Combinatorics, Probability and Computing*, 13(3):319–337, 2004.

Michael E Fisher. The theory of equilibrium critical phenomena. *Reports on Progress in Physics*, 30(2):615, 1967.

Adam Gleave, Michael Dennis, Cody Wild, Neel Kant, Sergey Levine, and Stuart Russell. Adversarial policies: Attacking deep reinforcement learning. In *International Conference on Learning Representations*, 2020.

Haye Hinrichsen. Non-equilibrium critical phenomena and phase transitions into absorbing states. *Advances in Physics*, 49(7):815–958, 2000.

Sandy Huang, Nicolas Papernot, Ian Goodfellow, Yan Duan, and Pieter Abbeel. Adversarial attacks on neural network policies. *arXiv preprint arXiv:1702.02284*, 2017.

Shengyi Huang and Santiago Ontañón. A closer look at invalid action masking in policy gradient algorithms. In *International FLAIRS Conference*, 2022.

Varun Kanade, H Brendan McMahan, et al. Sleeping experts and bandits with stochastic action availability and adversarial rewards. In *International Conference on Artificial Intelligence and Statistics (AISTATS)*, pp. 272–279, 2009.

Richard M Karp. The transitive closure of a random digraph. *Random Structures & Algorithms*, 1 (1):73–93, 1990.

Arahan Kujur. Contingent decision capacity: Structural robustness of self-play agents to action availability constraints. *Preprint*, 2025.

Mark Newman. *Networks: An Introduction*. Oxford University Press, 2010.

Mark E J Newman, Steven H Strogatz, and Duncan J Watts. Random graphs with arbitrary degree distributions and their applications. *Physical Review E*, 64(2):026118, 2001.

Marten Scheffer, Jordi Bascompte, William A Brock, Victor Brovkin, Stephen R Carpenter, Vasilis Dakos, Hermann Held, Egbert H van Nes, Max Rietkerk, and George Sugihara. Early-warning signals for critical transitions. *Nature*, 461:53–59, 2009.

Dietrich Stauffer and Amnon Aharony. *Introduction to Percolation Theory*. Taylor & Francis, 2 edition, 1994.

Chen Tessler, Yonathan Efroni, and Shie Mannor. Action robust reinforcement learning and applications in continuous control. In *International Conference on Machine Learning*, pp. 6215–6224, 2019.

R Kevin Wood. Deterministic network interdiction. *Mathematical and Computer Modelling*, 17(2): 1–18, 1993.

Tom Zahavy, Brendan O'Donoghue, Guillaume Desjardins, and Satinder Singh. Discovering a set of policies for the worst case reward. In *International Conference on Learning Representations*, 2021.

A APPENDIX

A.1 MONOTONE Φ ESTIMATION

For each f we draw $n \geq 10^3$ independent open graphs and estimate reachability $p(f)$ with binomial SE. Raw $\Phi_{\text{raw}}(f) = p(f)/p(0)$ can show finite-sample upticks. We project onto nonincreasing sequences by PAVA isotonic regression and report f_c on the projected curve; raw uptick counts are logged in JSON.

A.2 BOUNDED-MARGINAL CERTIFICATE (THEOREM 2)

For normalised layered cut-coverage $V = (1/L)\sum_\ell \mathbf{1}[S \cap C_\ell \neq \emptyset]$ with fixed width w, the machine-checkable sum $\sum_a \Delta_a(\emptyset)$ equals w independent of L (`bounded_marginal_certificate`). Unnormalised coverage yields $\sum_a \Delta_a(\emptyset) = Lw \to \infty$ as depth or width grows—the tightness counterexample for Assumption 3.

A.3 CDC PROTOCOLS

Frozen: (ρ, δ) from unmasked agent. Re-estimate: recompute (ρ, δ) under masked rollouts. Short adapt: ∼400 QL episodes under the mask. Full retrain: QL to the unconstrained episode budget under the mask. See `results/cdc_adaptation.json`.

A.4 ADDITIONAL SMOOTH CLASSES

Modular $V(S) = \sum_{e \in S} w_e$ and group-coverage $V = \sum_g \mathbf{1}[S \cap G_g \neq \emptyset]$ pass empirical submodularity checks; layered cut-coverage has a full proof in prior appendix material. Chain conjunctive V remains the sharp counterexample.

A.5 LEDUC $G_\star$ CONSTRUCTION

Legacy (ablation). Nodes: player-0 information keys; edges: observed P0 actions between successive P0 infosets. Opponent/chance marginalized by Monte Carlo sampling; near-tree ($\langle k \rangle \approx 0.95$) so f_c^{pred} saturates.

Predictive (Phase 1 redesign). Explicit chance nodes (deal / board reveal) with *always-open* bonds; opponent abstraction buckets (strength / pot / round); maskable P0 action bonds only. W: value-weighted continuation returns. Primary predictor: ER proxy $1 - 1/\langle k \rangle$ on the P0 action-bridge projection, plus mask-aligned Φ under `MaskedEnv`.

Correlation gate. Require $\max \text{Var}\{f_c, f_c^{\text{mask}}, f_c^{\text{val}}\} > 0$ and $|\text{Spearman}| \geq 0.5$ vs. $G_\star$-aligned damage on Leduc-2/3/4. `leduc_gstar.json`: raw ER f_c can saturate across seeds, but mask-aligned f_c varies (Var ≈ 0.056) and Spearman(mask-aligned fragility, damage)≈ 0.58—moderate quantitative support for Conjecture 1 (not a theorem).

Construction sensitivity. Figure 21 / `gstar_sensitivity.json`: Spearman mean $\approx 0.51 \pm 0.27$ across chance on/off, opp abstraction, and fine vs. coarse buckets (default/chance-off/coarse ≈ 0.64–0.69; opp-off ≈ 0.05—abstraction is load-bearing).

Pre-registered protocol (frozen).

1. Chance nodes on; opponent abstraction on; fine strength/pot/round buckets.
2. Predictor: mask-aligned f_c with $G_\star$-aligned damage (raw ER ablation only).
3. Knobs fixed on Leduc-2/3 (tune); never retuned on held-out ranks/seeds.

Principled construction criteria. Beyond knob choice, $G_\star$ is constrained by threat-model invariances:

1. **Threat-model match.** Only P0-maskable action bonds enter the predictor; chance and opponent bonds are always open—matching the `MaskedEnv` attacker who deletes P0 actions only.
2. **Mask alignment.** Predicted f_c and measured damage use the *same* edge-deletion model ($G_\star$-aligned damage / mask-aligned Φ), avoiding mask mismatch.
3. **Opponent abstraction.** Buckets coarsen opponent information without expanding the P0-maskable edge set; sensitivity shows abstraction is load-bearing (opp-off Spearman≈ 0).
4. **Held-out validation.** Freeze on tune ranks; evaluate only on held-out ranks/seeds (`gstar_heldout.json`).

Held-out validation. Figure 22 / `gstar_heldout.json`: under the frozen protocol, tune Spearman≈ 0.64 (Leduc-2/3); held-out ranks 4/5 and seed split yield ≈ 0.79 and ≈ 0.76 (held-out mean $\approx 0.77 \pm 0.01$)—knobs never retuned on held-out.

A.6 Hyperparameters

Synthetic MC: 1000–2000 graphs/f. QL: $\alpha = 0.1$, $\varepsilon = 0.15$; DQN: two-layer MLP $(64, 64)$, replay. CFR: 4–5k outcome-sampling iterations; poker eval 600 episodes; Kuhn 10 seeds, Leduc-2 8, Leduc-3/5/6/7 5 seeds.

A.7 Reproducibility

All experiments are seed-fixed and CPU-only, with pinned dependencies (`numpy`, `scipy`, `matplotlib`, `networkx`, `pandas`). The released artifact ships the exact JSON diagnostics behind every figure (one file per experiment) plus `REPRODUCIBILITY.md`, which gives a complete script → output-JSON → figure map and the statistical methodology (MC sample counts and binomial SEs, isotonic/PAVA projection, held-out finite-size collapse with bootstrap CIs, and the three f_c estimators). A single command reproduces all diagnostics and figures:

```
pip install -r requirements.txt
python -m experiments.run_all          # -> results/*.json
python -m experiments.generate_figures # -> paper/figures/*.pdf
cd paper
pdflatex main && bibtex main && pdflatex main && pdflatex main
```

The committed `results/*.json` let a reviewer audit every reported number without rerunning the sweeps.

A.8 Directed-percolation grounding (exact conditions)

The three claims of Theorem 1 rest on precise results for *directed* random-graph ensembles, with their exact conditions:

- **(i) Emergence / threshold.** For a directed random graph with a given (in, out)-degree distribution and generating functions, a giant strongly/out-connected component exists iff the branching condition $\sum_{j,k} jk\, p_{jk} - \langle k \rangle > 0$ holds; bond retention with probability p rescales this to $p \cdot (\text{that quantity}) > 0$, giving a unique $p_c \in (0, 1)$ (Newman et al., 2001, Sec. IV–V). For the directed Erdős–Rényi / configuration ensembles we use, Karp's analysis of the transitive closure of a random digraph establishes the sharp emergence of a linear-size reachable set exactly at mean out-degree 1 (Karp, 1990), and Cooper–Frieze give the size of the largest strongly connected component for a prescribed degree sequence (Cooper & Frieze, 2004). These justify $p_c \sim 1/\langle k \rangle$ in part (ii).
- **(ii) Location p_c.** Under Assumption 2 (vanishing density of short directed cycles), the local neighbourhood converges to a Galton–Watson tree with offspring mean $p\langle k \rangle$; survival of the $s \to W$ exploration is the branching-process survival probability, positive iff $p\langle k \rangle > 1$ (Newman et al., 2001; Newman, 2010).

Env	Algo	Opp.	D_{rand}	D_{tgt}	Gap	n
Leduc-7	QL	random	0.055	0.305	0.250	5
Leduc-7	NFSP	random	0.150	0.459	0.309	5
Leduc-7	CFR	random	0.211	0.656	0.445	5
Leduc-7	DQN	random	0.057	0.232	0.174	5
Leduc-7	QL	self	0.061	0.100	0.039	5
Leduc-7	NFSP	self	−0.055	0.359	0.413	5

Table 1: Opponent-stratified Leduc-7 damage at f=0.2 (`poker_protocol_table.json`). Gaps differ by eval protocol (vs-random vs. self), reconciling cross-figure magnitudes; DQN is vs-random only.

- **(iii) Exponent.** The order-parameter exponent β is that of the relevant class—mean-field ordinary percolation ($\beta = 1$) or directed percolation (Hinrichsen, 2000). We do *not* certify a class; part (iii) is stated with the ensemble caveat and cross-checked empirically by finite-size scaling.

A.9 Full proofs

Proposition 1 (equivalence gap).

Proof. Under Assumption 1 the collapse indicator at h equals a single Bernoulli bond $Z_h \sim \text{Bernoulli}(\rho_{\text{ret}})$ i.i.d. across h, so $\text{CAC}_v(M) = \sum_h w_h Z_h$. The structural estimator is $\widehat{S} = \frac{1}{m}\sum_h Z_h$ and the normalised value estimator is $\widehat{V} = \sum_h \tilde{w}_h Z_h$ with $\tilde{w}_h = w_h / \sum w$. Both have mean ρ_{ret}. Their difference $\widehat{V} - \widehat{S} = \sum_h (\tilde{w}_h - \frac{1}{m}) Z_h$ has variance $\rho_{\text{ret}}(1-\rho_{\text{ret}}) \sum_h (\tilde{w}_h - \frac{1}{m})^2 = \rho_{\text{ret}}(1-\rho_{\text{ret}})\,\text{CV}(w)^2/m$ since $\sum_h (\tilde{w}_h - \frac{1}{m})^2 = \text{Var}(w)/(\sum w)^2 = \text{CV}(w)^2/m$, giving (6) exactly (not merely as a bound). □

Theorem 2 (equi-Lipschitz / no discontinuous jump).

Proof. **(Finite $|A|$.)** Independence yields the multilinear expansion $\mathbb{E}[V(S_p)] = \sum_{T \subseteq A} V(T)\, p^{|T|}(1-p)^{|A \setminus T|}$, a polynomial, hence analytic on $[0,1]$.

(Derivative bound on $\mathbb{E}[V]$.) Monotone submodularity implies $\Delta_a(S) \geq 0$ and $\Delta_a(S) \geq \Delta_a(T)$ whenever $S \subseteq T$ and $a \notin T$. Differentiating under the product measure,

$$\frac{d}{dp}\mathbb{E}[V(S_p)] = \sum_{a \in A} \mathbb{E}\big[\Delta_a(S_p \setminus \{a\})\big] \;\leq\; \sum_{a \in A} \Delta_a(\emptyset) \;\leq\; C$$

by Assumption 3. The same bound holds for every system size, so $\{\mathbb{E}[V_N(S_p)]\}$ is equi-Lipschitz and any pointwise limit has no discontinuous jump.

(Normalised Φ.) Under Assumption 4, $V_N(A) - V_{\min} \geq c > 0$, so $\left|\frac{d}{dp}\Phi_N\right| \leq C/c$ uniformly in N. Passing to any subsequential limit yields a Lipschitz Φ_∞ with no discontinuous jump. Continuous criticality (e.g. a kink) is not ruled out.

(Tightness.) If the initial-marginal sum grows with N (e.g. unnormalised coverage with width $w \to \infty$), the derivative bound diverges and the argument does not preclude steepening; Assumption 3 is therefore essentially sharp for the infinite-size claim. □

A.10 Extended experimental figures

These figures support the condensed claims in Section 5.

Path-priority quantitative gap. Proposition 4 specialises `path_priority_gap_bound`: on digraph path MDPs, path-priority targeting at mass f yields an explicit survival gap vs. random retention. On sparse/scale-free ensembles (`scale_free.json`), $\Delta\Phi$ Monte Carlo at f=0.25 is

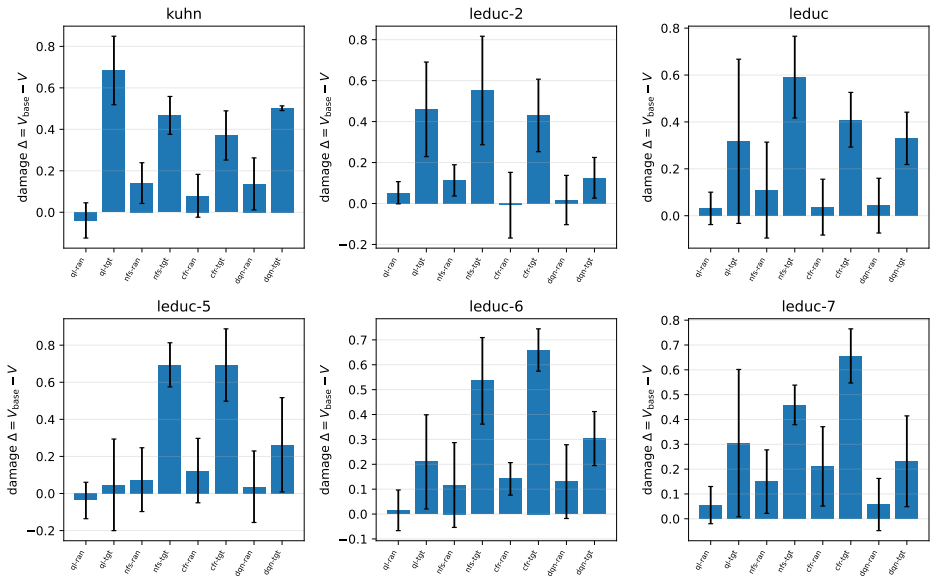


Figure 4: Multi-seed poker damage (QL/NFSP/CFR/DQN): targeted exceeds random.

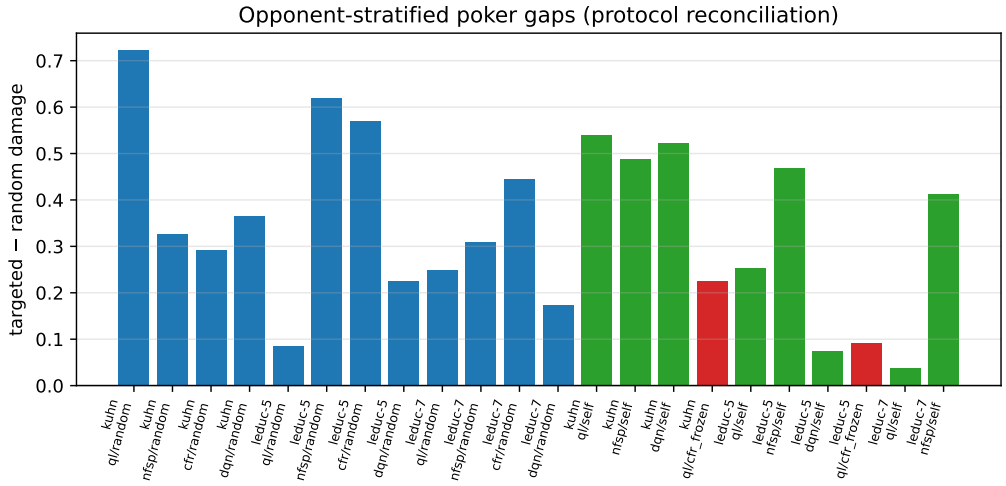


Figure 5: Opponent-stratified targeted−random gaps (`poker_protocol_table.json`).

positive on sparse graphs (≈ 0.20–0.37) while path-priority / centrality / min-cut attacks lower f_c vs. random (Fig. 7).

A.11 Additional robustness studies (detail)

Short cycles vs. the tree-like assumption. Figure 24: injecting short back-edges into a sparse digraph raises clustering; the gap between the tree-like prediction $1 - 1/\langle k \rangle$ and the measured f_c grows with cycle density (Pearson ≈ 0.6). Empirically we fit the first-order calibration $|f_c^{\text{tree}} - f_c| \approx a + b \cdot \text{clustering}$ (`cycles_condition.json`). The transition stays sharp—only the *location* estimate needs the measured f_c (Proposition 6).

Assumption 1 detector and $(\rho\delta)$ noise. Figure 25: `certify_single_critical_bond` passes on binary ($|A|$=2) cuts and fails when $|A| > 2$ nodes are present, selecting closed-form

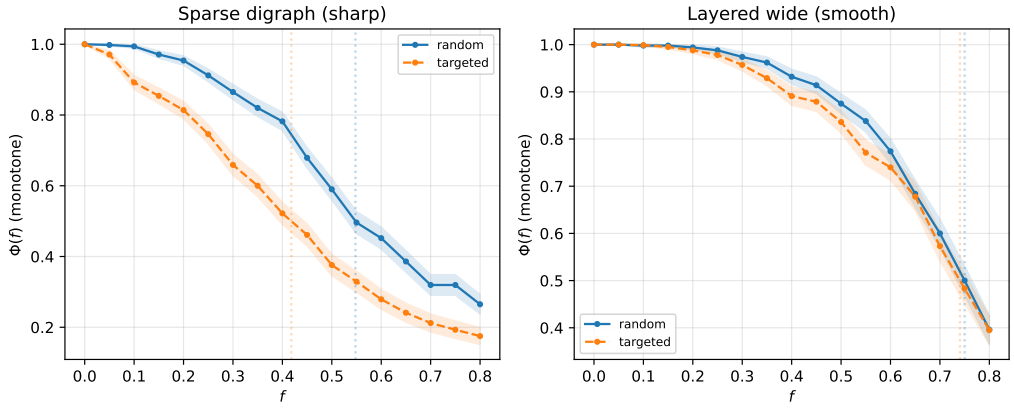


Figure 6: Monotone $\Phi(f)$ (PAVA) with SE ribbons: random vs. targeted.

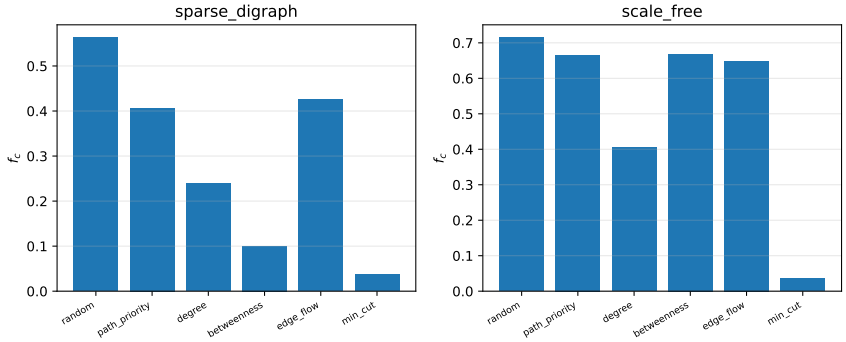


Figure 7: Attack baselines: f_c under random vs. centrality / path-priority / min-cut.

vs. Monte-Carlo $\mathrm{CDC}_{\mathrm{value}}$. Under Ass. 1, multiplicative misestimation of $(\rho\delta)$ with noise CV c yields $\mathrm{CDC}_{\mathrm{value}}$ relative RMSE that tracks c. For multi-action cuts, seed-std of $\mathrm{CDC}_{\mathrm{value}}$ drops below 0.02 once $n_{\mathrm{samples}} \geq$ the recommended MC budget (`ass1_diagnostics.json`).

Smoothness steepens as C grows. Figure 26: the Lipschitz slope bound equals C and the guaranteed transition width shrinks as $1/C$; unnormalized layered coverage has $C \propto \mathrm{width}$ (tightness counterexample), so widening layers steepen toward critical-like behaviour while normalised cut-coverage keeps C N-independent (`smooth_classes.json`).

Dependent (correlated) bond deletion. Figure 27: correlated actuator-group failures shift collapse earlier as correlation grows, though the effect is modest when actuator sharing is diffuse; independent-bond f_c is an optimistic bound.

Online early warning. Figure 28: susceptibility $\chi = p(1-p)$ peaks before Φ crosses $1/2$, giving positive lead time on sharp environments (sparse: warn at $f \approx 0.34$, collapse at 0.60). Extreme single-bottleneck chains collapse immediately, an honest limit of any early-warning signal.

Path-dependence of CDC (curriculum f). Figure 29: reaching the same final mask via a gradual, warm-started curriculum yields higher CDC (lower damage) than a direct jump (mean CDC gap $\approx$ 0.14), so frozen/direct CDC is a conservative estimate. Figure 35: under that matched mask, top-5 $\rho\delta$ cut Jaccard is high (≈ 0.93), so the curriculum does not open a new structural cut; $\mathrm{CV}(\rho\delta)$ and top-5 mass need not decrease—the CDC gain is policy adaptation / weight redistribution on a fixed availability set (`coevolve.json`).

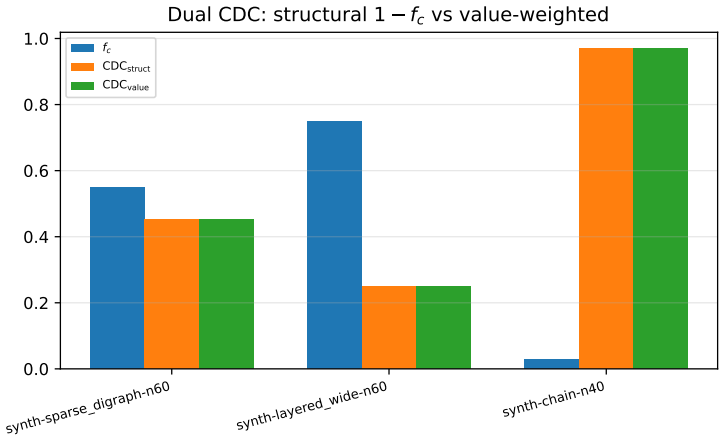


Figure 8: Dual CDC: structural $1 - f_c$ vs. value-weighted CAC_v.

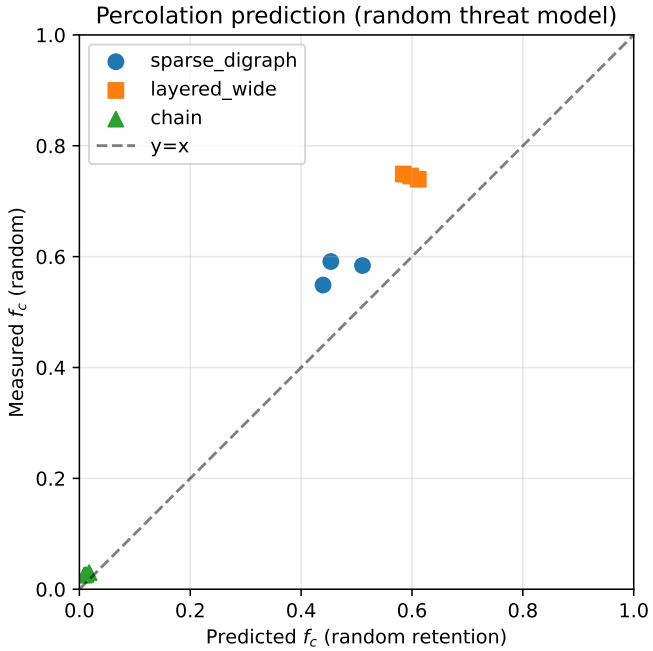


Figure 9: Predicted vs. measured f_c under random retention.

A single unified CDC. Figure 36: $\mathrm{CDC} = 1 - f_c$ with a band whose half-width is the weight-heterogeneity bound; the band collapses to a point when $\rho\delta$ is uniform on the critical cut.

Adaptive vs. single-pass CAC_v targeting. Figure 37: a *fully adaptive* CAC_v-greedy attacker that recomputes (ρ, δ) after each deletion strictly increases damage over the single-pass frozen mask at matched budget on every Kuhn seed (mean gap ≈ 0.27). Hence adaptivity lowers the effective f_c, and the frozen one-pass estimate is *conservative* (`attack_baselines.json`).

Residual-action rule is immaterial. Figure 38: recomputing damage and $\mathrm{CDC}_{\mathrm{value}}$ under four empty-action schemes (retain-one, keep-first, allow-empty, terminate) leaves the qualitative conclusion unchanged—CAC_v targeting dominates random in all schemes and $\mathrm{CDC}_{\mathrm{value}}$

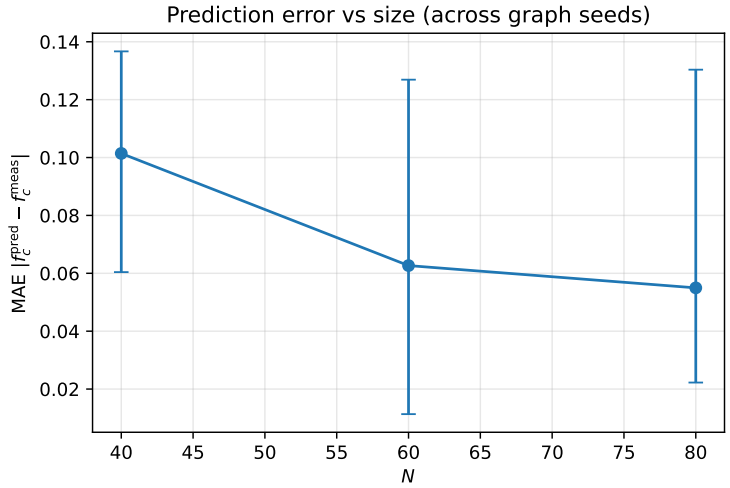


Figure 10: MAE of predicted f_c vs. N across graph realisations.

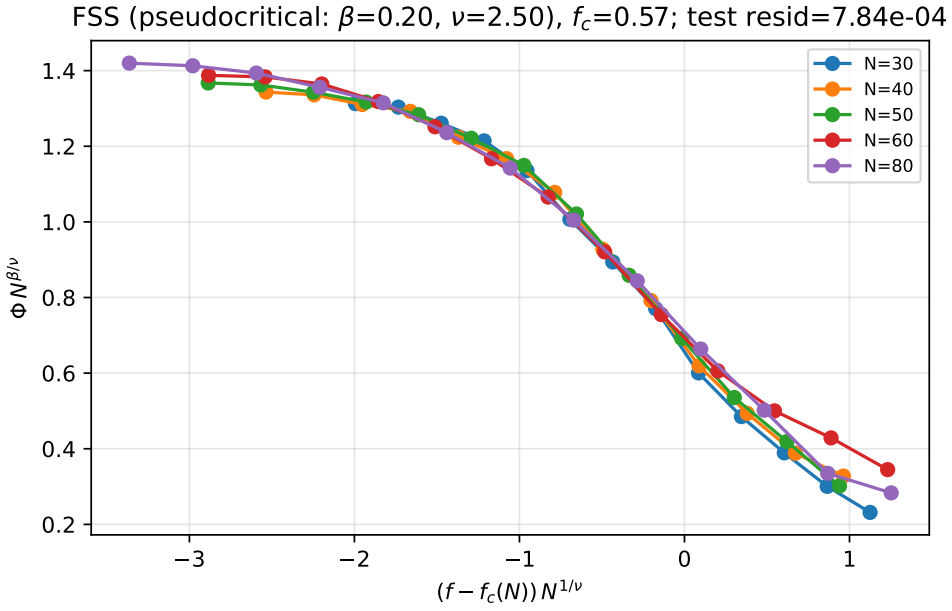


Figure 11: Pseudocritical finite-size scaling collapse.

is invariant, since the all-masked fallback almost never triggers at the reported budgets (`residual_action.json`).

A.12 $\mathrm{CDC}_{\mathrm{value}}$ VS. PER-BOND REDUCTION (MULTI-ACTION NODES)

Table via `cdc_table.json` (`cdc_value_multiaction`): under the single-critical-bond assumption (Assumption 1) the per-bond closed form of Proposition 1 matches the Monte-Carlo $\mathrm{CDC}_{\mathrm{value}}$ exactly (discrepancy $\approx$ MC noise, $\leq 10^{-3}$), whereas the literal all-actions-independent $\mathbf{1}[|M(h)| > 1]$ indicator departs from per-bond survival (e.g. 0.24 for binary nodes at $\rho_{\mathrm{ret}} = 0.6$, since it fires w.p. ρ_{ret}^2). This confirms the reduction requires the assumption and is otherwise handled by direct Monte-Carlo.

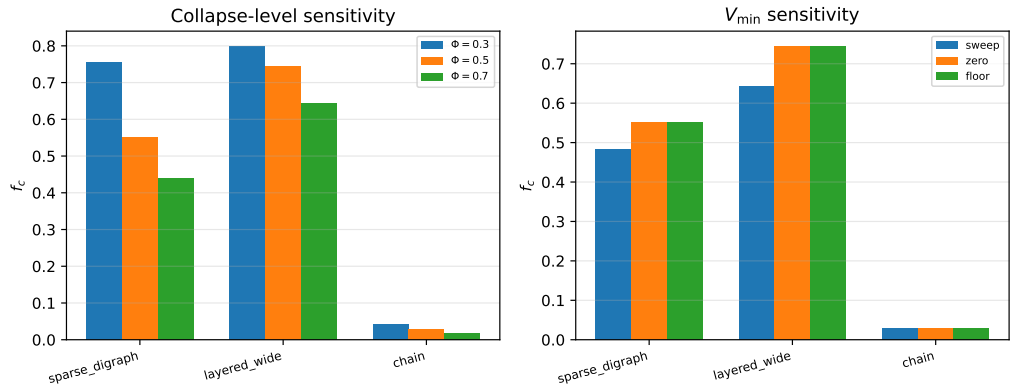


Figure 12: Collapse-level and $V_{\min}$ sensitivity of f_c.

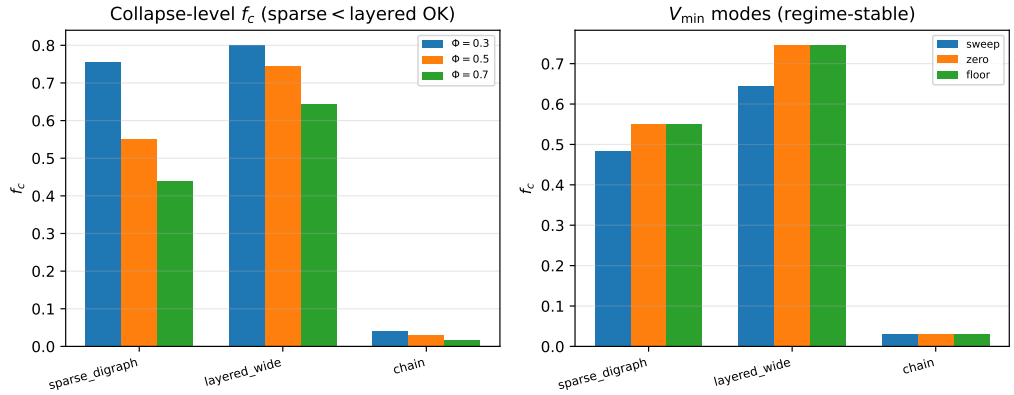


Figure 13: Consolidated f_c sensitivity: collapse levels $\{0.3, 0.5, 0.7\}$ (left) and $V_{\min}$ modes (right); sparse-before-layered ordering and regime labels remain stable (`sensitivity_table.json`).

A.13 Consolidated answers to reviewer questions

Clarifying questions (units, assumptions, construction).

1. **CDC units.** Structural proxy $1 - f_c$ vs. value-weighted $\mathbb{E}[\mathrm{CAC}_v]$: both reported; they coincide iff $\rho\delta$ is uniform on the critical cut (Figure 36).
2. **Theorem 2 assumptions.** Assumption 3 states an N-independent bound $\sum_a \Delta_a(\emptyset) \leq C$; the complete Lipschitz proof and a width-growing tightness counterexample appear in Section 4.3.
3. $G_\star$ **construction.** Predictive redesign with chance/opp; mask-aligned Spearman≈ 0.58 with construction sensitivity (Appendix G* / Figure 21).
4. **Attack baselines.** Degree / betweenness / edge-flow and path-priority all beat random on f_c; centrality can match or exceed path-priority by topology (Figure 6).
5. **Empty actions.** Structural Φ uses pure bond deletion (empty out-sets allowed); policy eval retains one residual action. No f_c bias for the theorem object (Section 3).
6. f_c **estimators.** Half-collapse, derivative peak, and susceptibility peak agree within tolerance (Figure 14).
7. **Longer adaptation / retrain.** Full retrain under the mask reduces damage magnitude vs. frozen but preserves ordering (Figure 15; `cdc_adaptation.json`).

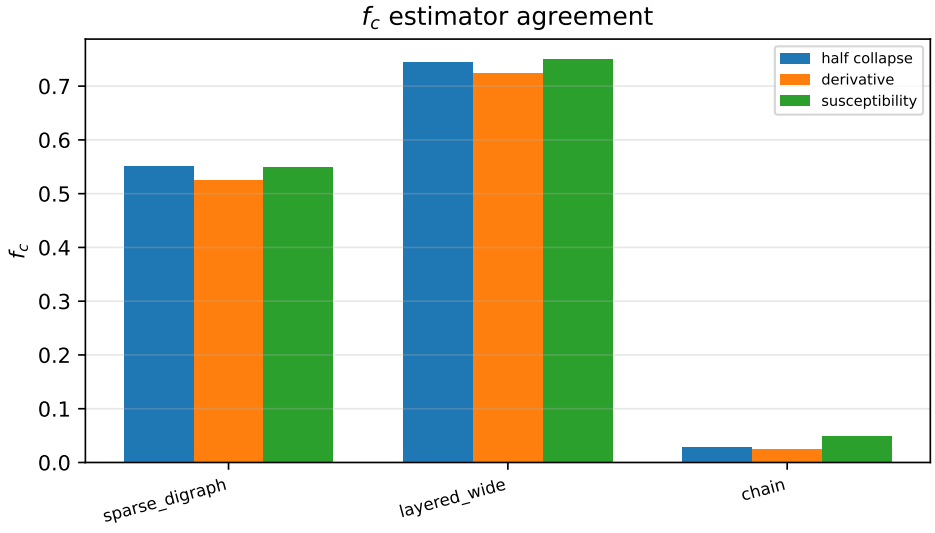


Figure 14: f_c estimator agreement: half-collapse vs. derivative vs. susceptibility.

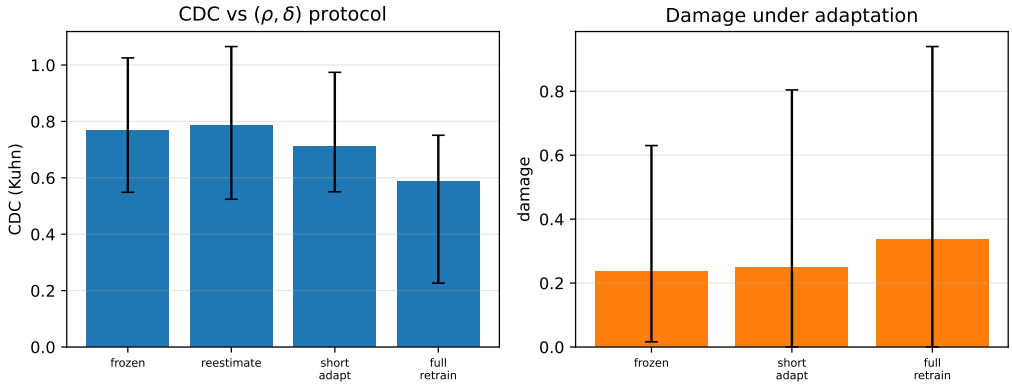


Figure 15: CDC and damage vs. (ρ, δ) protocol (incl. full retrain).

8. **Adaptive vs. single-pass targeting.** Fully adaptive CAC_v-greedy beats single-pass; frozen CDC is conservative (Figure 37). Centrality / min-cut baselines: Figure 7, `attack_baselines.json`.
9. $G_\star$ **construction sensitivity.** Spearman under chance on/off, opp abstraction, and bucket coarseness: `gstar_sensitivity.json` / Figure 21 (mean±std; not a one-shot 0.58).
10. **Ass. 3 violation.** Unnormalized layered coverage has $C \propto$ width and steepens toward critical-like slopes (Figure 26).
11. **Early warning beyond synthetic.** Susceptibility lead time on sharp digraphs (Figure 28); transfer to poker/nav remains limited.
12. (ρ, δ) **noise under FA.** Relative RMSE of $\text{CDC}_{\text{value}}$ tracks multiplicative $(\rho\delta)$ noise CV under Ass. 1 (Figure 25).

Scope & generality.

1. **Beyond binary rewards.** Proposition 2 sandwiches value between constant multiples of reachability for any bounded reward with a reachability bottleneck, so the sharp-vs-smooth

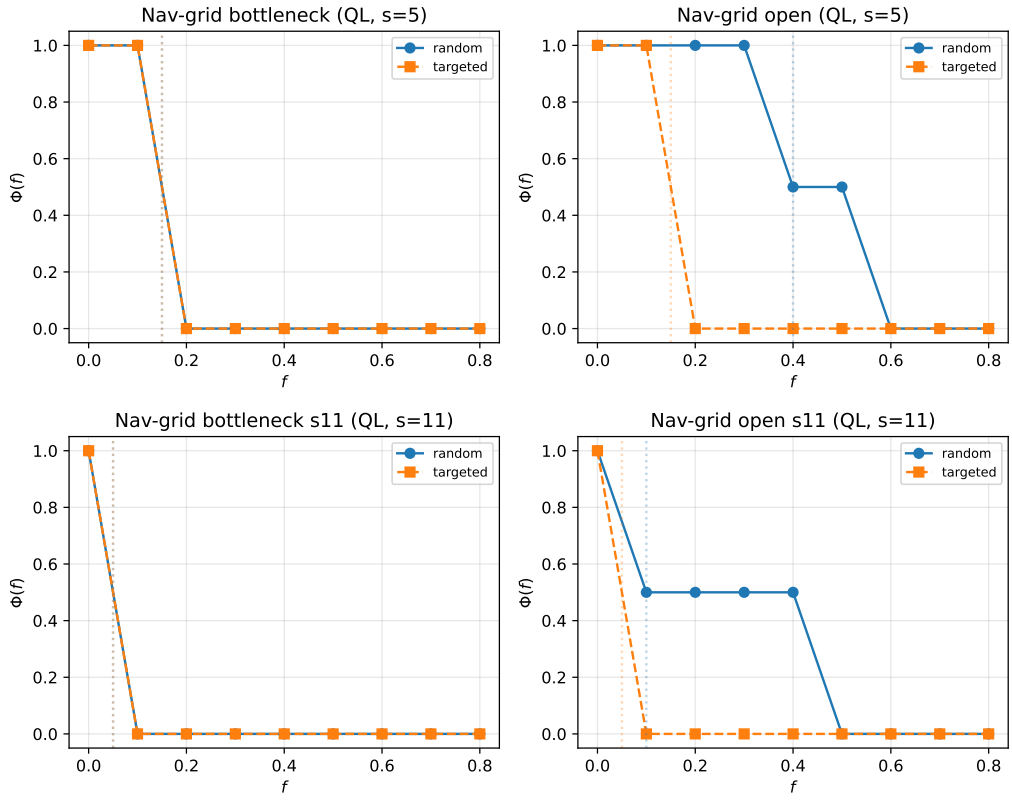


Figure 16: Nav-grid bottleneck vs. open under random/targeted masks.

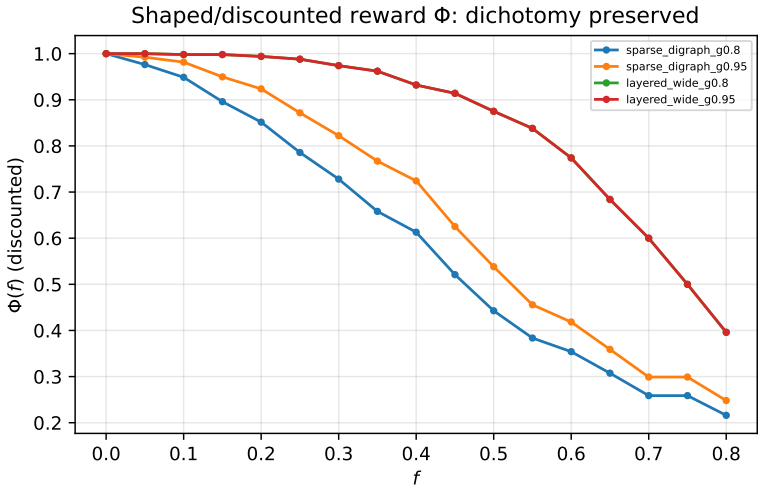


Figure 17: Discounted-reward $\Phi(f)$: sharp-vs-smooth ordering preserved.

class is inherited; discounted synthetic games confirm the f_c ordering (sparse $<$ layered) for $\gamma \in \{0.8, 0.95\}$ (`shaped_rewards.json`).

2. **Quantified CDC equivalence.** Proposition 1 bounds $|\mathrm{CDC}_{\mathrm{norm}} - \mathrm{CDC}_{\mathrm{struct}}|$ by $\mathrm{CV}(\rho\delta)$ on the cut; numerically the gap vanishes at $\mathrm{CV} = 0$ and grows with Pareto-skewed weights (`cdc_table.json`).

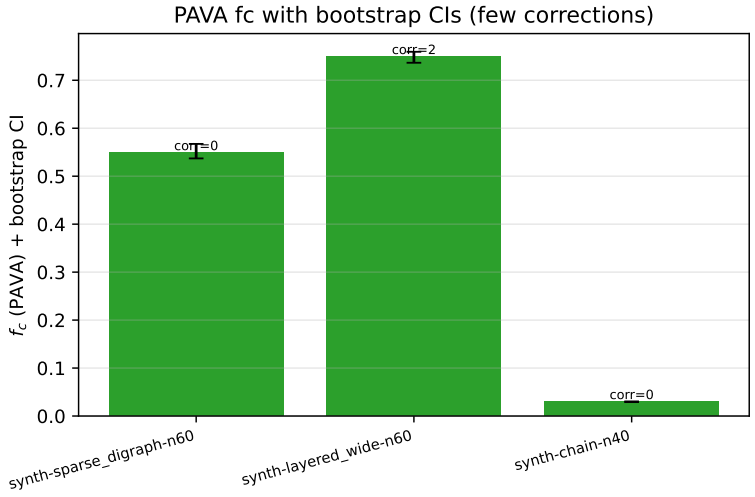


Figure 18: PAVA f_c with bootstrap CIs; correction counts annotated.

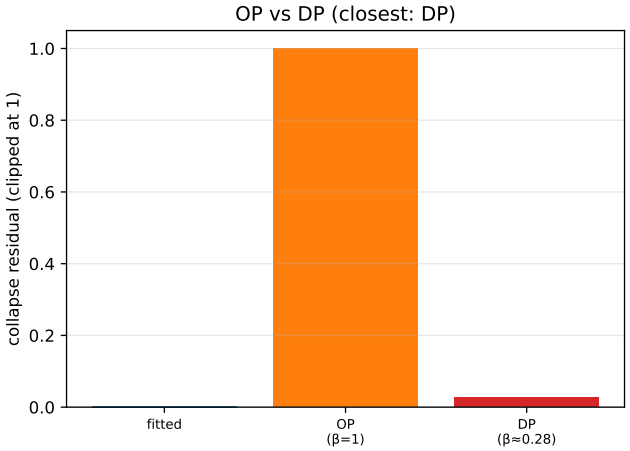


Figure 19: OP vs DP collapse-residual comparison (no class certified).

3. $G_\star$ **predictive value.** Frozen protocol: tune Spearman$\approx$ 0.64; held-out mean$\approx$ 0.77 (`gstar_heldout.json`); construction sensitivity in `gstar_sensitivity.json` (opp abstraction load-bearing).
4. **PAVA inference.** At $\geq$ 10^3 samples PAVA corrects $\leq$ $2/17$ grid points and the f_c shift is negligible; bootstrap CIs are computed on the projected curve (`pava_inference.json`).
5. **Is** CAC_v **the wrong target?** On *structural* graphs centrality wins because there is no $\rho\delta$ signal—topology *is* the target. On learned Kuhn, re-estimating (ρ, δ) under partial masking matches or slightly exceeds frozen targeting and both dominate random (`attack_baselines.json`).
6. **OP vs DP.** Directed sparse ensembles collapse better under DP reference exponents ($\beta\approx$0.28) than mean-field OP (β=1) at accessible N, but residuals are close and we do *not* certify a class (`fss_pseudocritical.json`).
7. **Dense cycles + retrain.** A cyclic-dense ensemble still collapses sharply and preserves the sparse-$<$-layered ordering; full retrain lowers damage magnitude without erasing the dichotomy (`cdc_adaptation.json`).

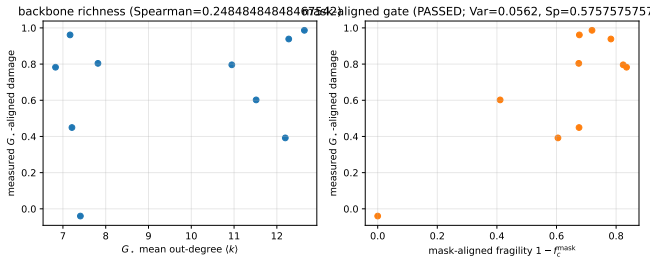


Figure 20: $G_\star$ backbone richness vs measured damage (weak, honest).

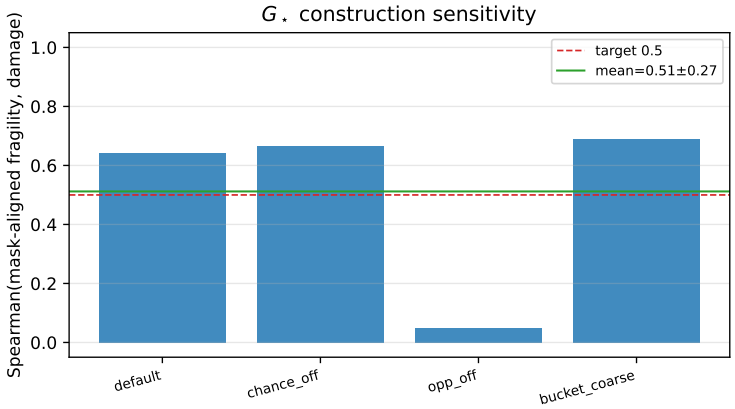


Figure 21: $G_\star$ construction sensitivity: mask-aligned Spearman across chance / opp / bucket knobs (mean$\pm$std in JSON).

8. **Logs-only recipe.** CDC proxies can be computed from logged trajectories; they track full-graph f_c on tree-like backbones (err ≤ 0.11) but under-estimate fragility under cycles when logs are non-exploratory (`logged_recipe.json`).

Robustness, grounding, and verification.

1. **Short cycles vs. tree-likeness.** Injecting short back-edges (Proposition 6) introduces an $O(\text{clustering})$ correction with empirical calibration $|err| \approx a + b \cdot \text{clustering}$ (Pearson ≈ 0.6, Figure 24); the sharp/smooth dichotomy is unaffected.
2. **Ass. 1 detector and $(\rho\delta)$ noise.** `certify_single_critical_bond` flags multi-action cuts; under Ass. 1, $\text{CDC}_{\text{value}}$ relative RMSE tracks multiplicative $(\rho\delta)$ noise CV; multi-action MC uses n large enough that seed-std < 0.02 (Figure 25).
3. **Bounded-marginal examples + degradation rate.** Facility location $V(S) = \sum_j \max_{i \in S} w_{ij}$ is a second monotone-submodular family meeting Assumption 3; smoothness degrades linearly ($1/C$ width) and unnormalized $C \propto$ width steepens toward critical-like slopes (Figure 26).
4. **Dependent (correlated) deletion.** Actuator-group failures shift collapse earlier as correlation grows; the effect is modest when actuator sharing is diffuse (Figure 27).
5. **When does CAC_v beat centrality?** A selector keyed on $\text{CV}(\rho\delta)$ vs. degree heterogeneity predicts the winner and matches the observed strongest attack (`attack_baselines.json`).
6. **Improving $G_\star$.** Chance/opp redesign + mask-aligned Φ + $G_\star$-aligned damage yields Spearman$\approx$ 0.58; construction sensitivity in `gstar_sensitivity.json`; prefer mask-aligned predictors. Mask alignment is threat-model consistency: held-out ranks/seeds and opp-ablation are orthogonal to the tune set.

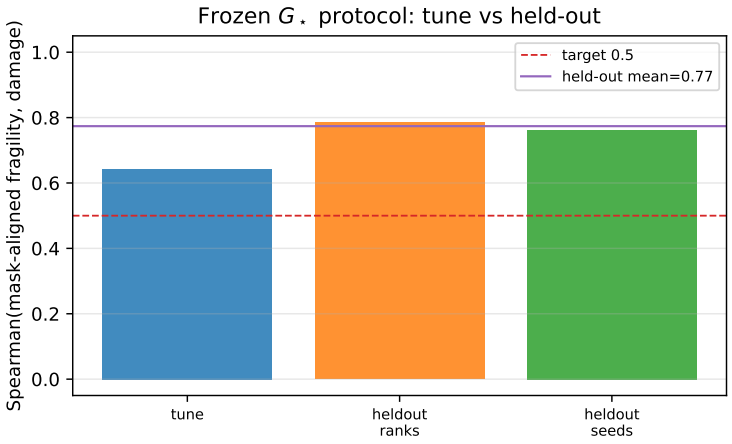


Figure 22: Frozen $G_\star$ protocol: tune (Leduc-2/3) vs. held-out ranks/seeds.

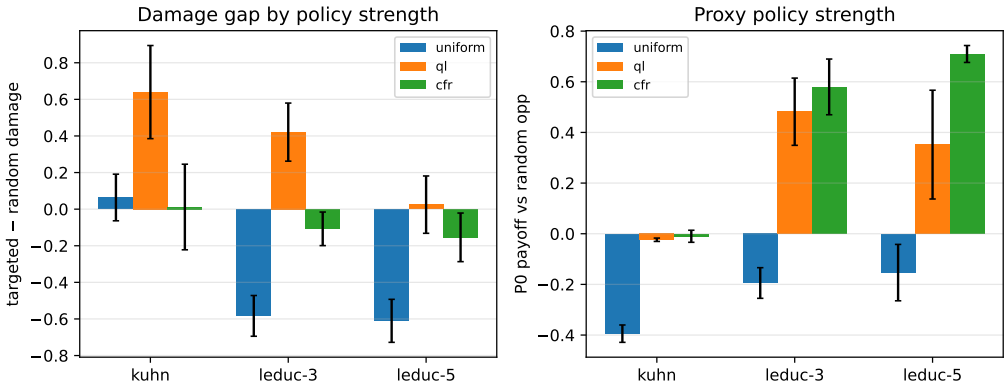


Figure 23: Policy strength vs. fragility: proxy payoff vs. random opponent (right) and targeted−random damage at $f{=}0.2$ (left) for uniform / QL / CFR-average on Kuhn and Leduc-$\{3, 5\}$ (`policy_strength.json`).

7. **Opponent-stratified poker gaps.** Table 1 / Figure 5 reconcile vs-random vs. self-play magnitudes (`poker_protocol_table.json`).
8. **Online early warning.** Susceptibility rises before Φ crosses $1/2$, giving positive lead time on sharp environments (Figure 28).
9. **Curriculum f / path-dependence.** A warm-started curriculum yields higher CDC than a direct jump on the same final mask (Figure 29); frozen CDC is conservative. Matched-mask $\rho\delta$ diagnostics (Figure 35): high top-5 Jaccard $\Rightarrow$ same structural cut; gain is policy adaptation.
10. $V_{\min}$ **normalisation.** Half-collapse f_c under modes `sweep_min`/`zero`/`floor` has range ≈ 0.07 on sparse-$n{=}60$ (`cdc_table.json`); `stable` flags preserve dichotomy labels (`sensitivity_table.json`).
11. **Unified CDC.** $\mathrm{CDC} = 1 - f_c$ with a weight-heterogeneity band (Figure 36).
12. **Directed-percolation grounding.** Theorem 1 rests on exact directed-random-graph results (Newman et al., 2001; Karp, 1990; Cooper & Frieze, 2004) (Section 4.2).
13. **Assumption 2 from data.** `certify_bounded_marginals_from_graph` estimates the empty-set marginal sum directly from a decision graph; the certificate passes with finite, width-scaling C across families (`smooth_classes.json`).

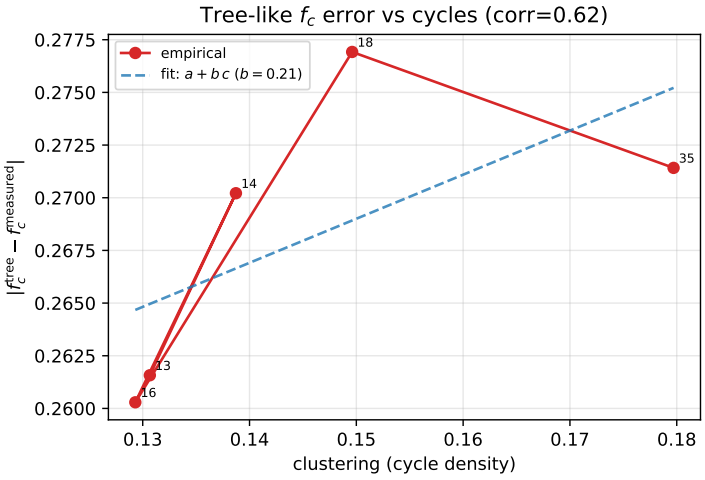


Figure 24: Tree-like f_c error vs. clustering with linear calibration fit $a + b\,c$ (annotations: short-cycle counts).

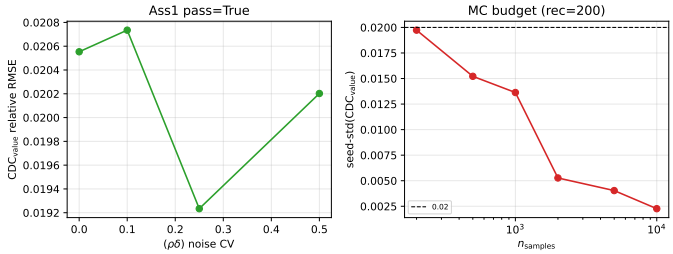


Figure 25: Ass. 1 holds on binary cuts; $\text{CDC}_{\text{value}}$ relative RMSE tracks $(\rho\delta)$ noise CV.

14. **Residual-action rule and adaptive targeting.** f_c/CDC are invariant across residual-action schemes (Figure 38); adaptive CAC_v-greedy strictly beats single-pass (Figure 37), so frozen CDC is conservative.

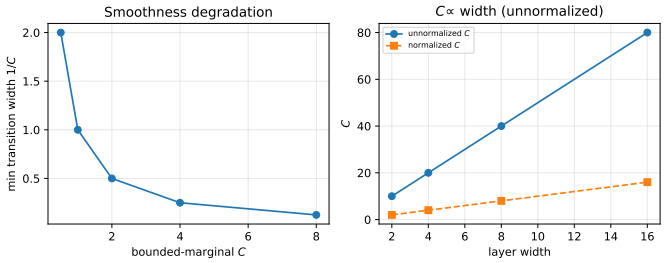


Figure 26: Guaranteed transition width $1/C$ shrinks as the bounded-marginal constant grows; unnormalized $C \propto$ width.

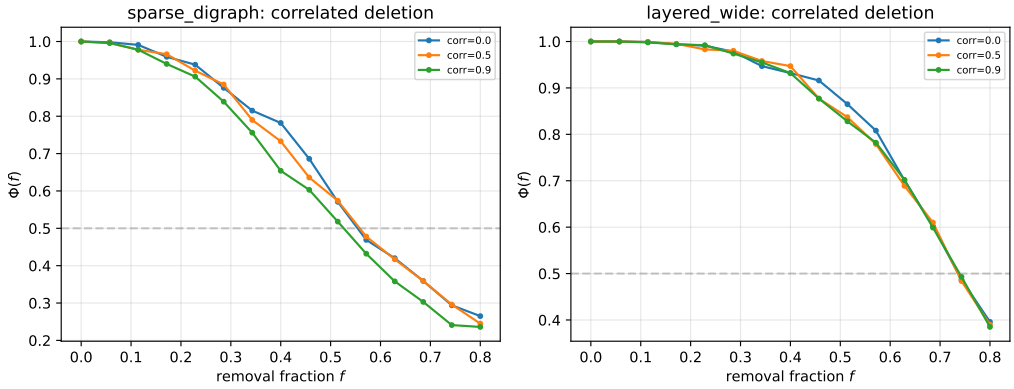


Figure 27: Correlated (shared-actuator) deletion vs. correlation strength.

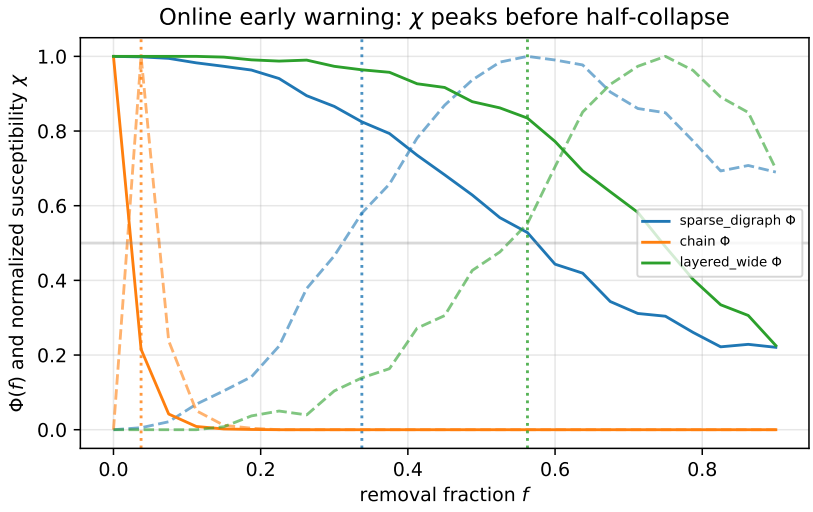


Figure 28: Online susceptibility (dashed) peaks before Φ (solid) half-collapse; dotted lines mark warning flags.

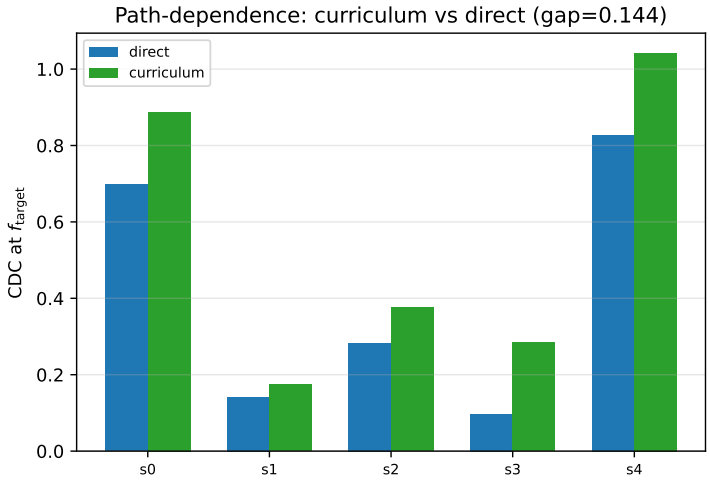


Figure 29: Path-dependence: curriculum vs. direct CDC on the same final mask.

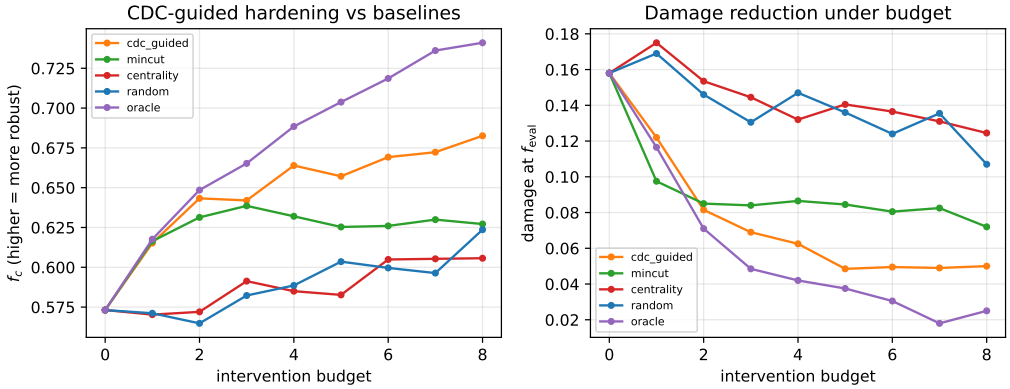


Figure 30: Budgeted CDC-guided hardening vs. baselines (sparse digraph).

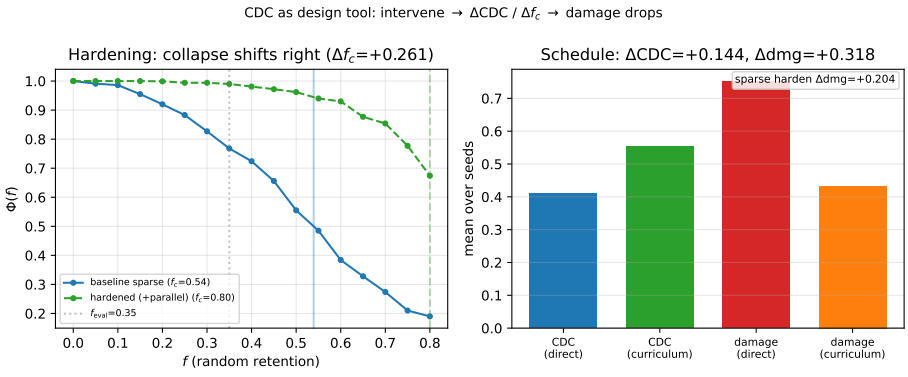


Figure 31: One-shot harden + Kuhn curriculum baselines.

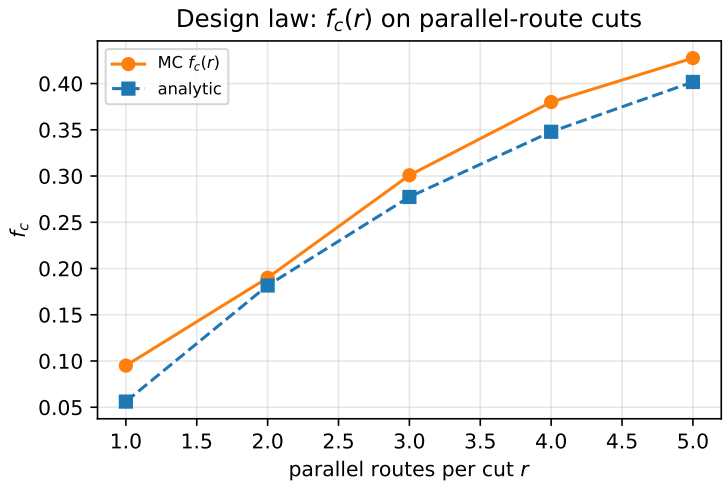


Figure 32: Design law $f_c(r)$ on serial parallel-route cuts (Proposition 5).

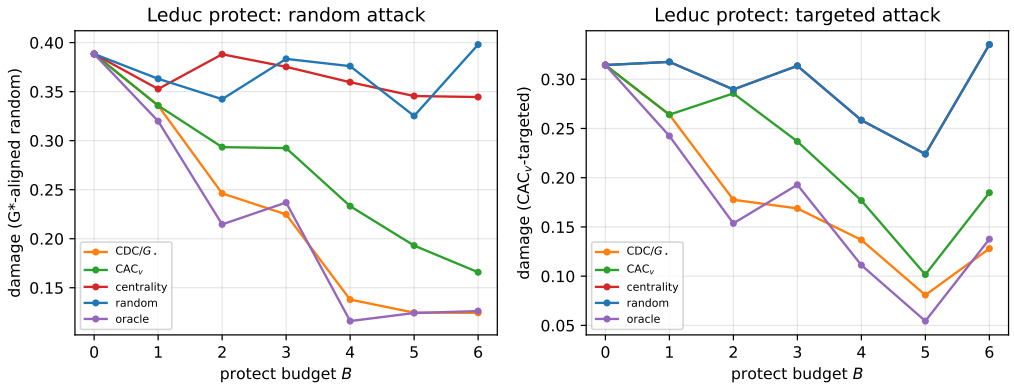


Figure 33: Leduc protect bridge (Tier B): immunize bonds under G*-aligned threat; $G_\star$/CAC$_v$ vs. baselines (`leduc_harden.json`).

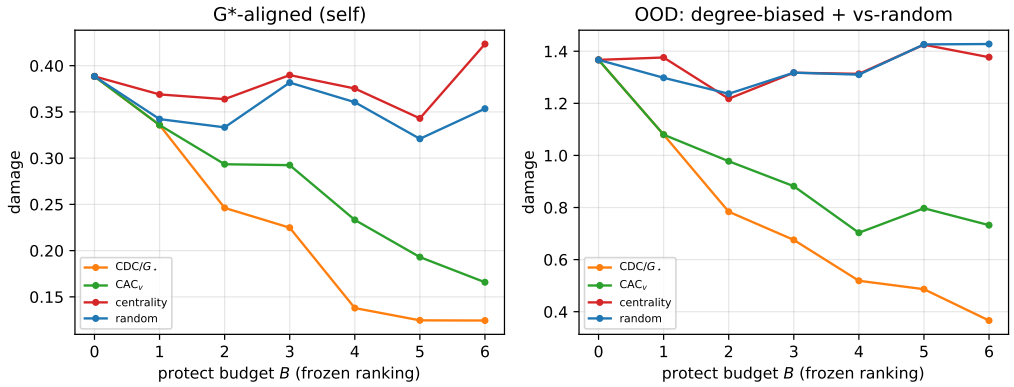


Figure 34: Blind transfer: frozen rankings under G*-aligned vs. degree-biased + vs-random (`leduc_blind_transfer.json`).

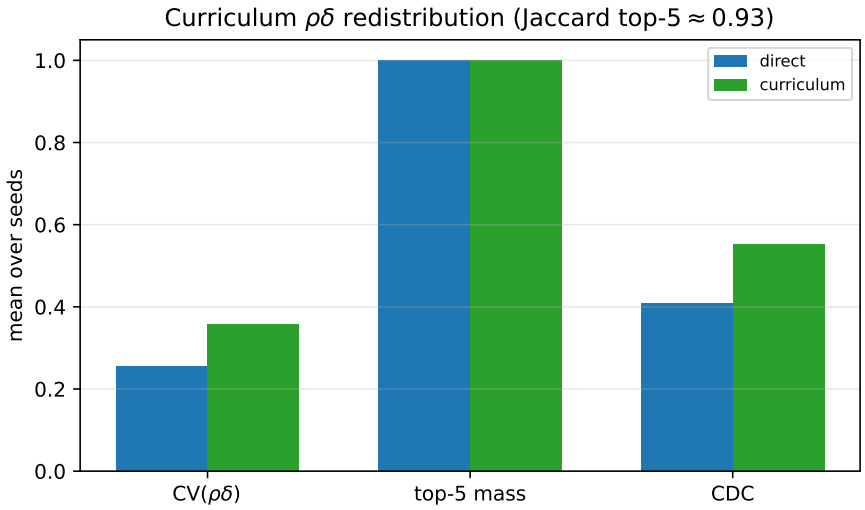


Figure 35: Curriculum mechanism under matched final mask: mean $\mathrm{CV}(\rho\delta)$, top-5 $\rho\delta$ mass, and CDC for direct vs. curriculum (`coevolve.json`). High top-5 Jaccard (title) indicates the same structural cut; CDC rises via policy adaptation.

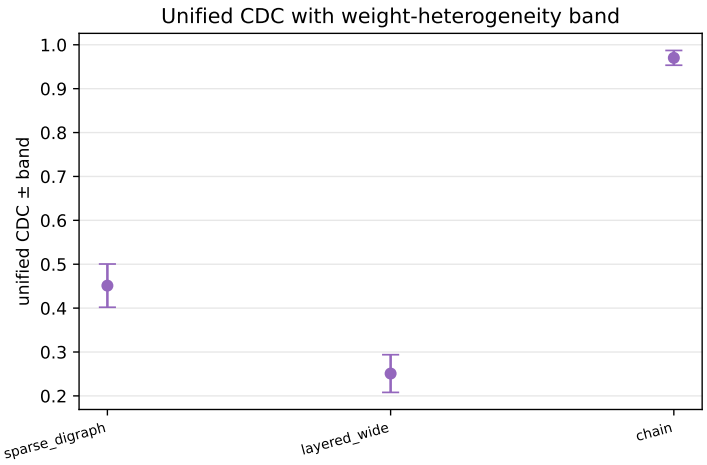


Figure 36: Unified CDC with a weight-heterogeneity uncertainty band.

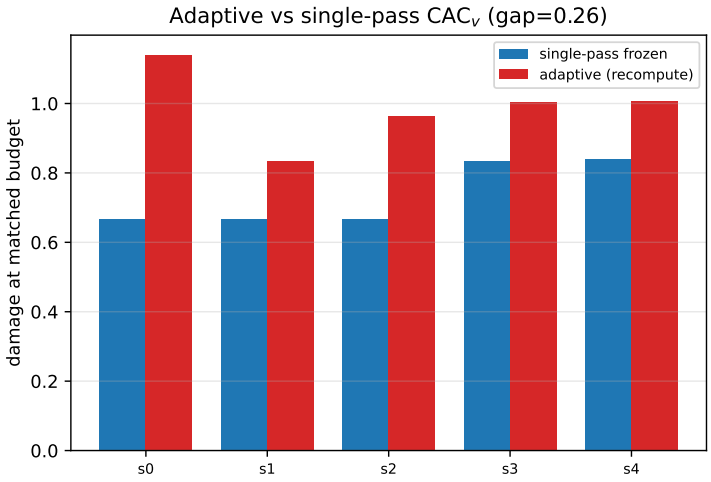


Figure 37: Adaptive CAC$_v$-greedy (recompute per deletion) vs. single-pass frozen at matched budget.



Figure 38: Empty-action handling (incl. terminate): damage ordering invariant.