STEALTH IS A RELATION, NOT A PROPERTY: HOW EVENT REPRESENTATIONS CREATE BLIND SPOTS FOR TIMING ATTACKS IN EVENT-BASED PERCEPTION

Anonymous authors
Paper under double-blind review

ABSTRACT

An event camera produces an asynchronous stream, but what is visible in that stream depends on how a downstream consumer, such as a model or detector, processes time. The same timestamp change may leave a coarse temporal representation unchanged while changing the response of a model that preserves finer timing. This creates blind spots that timing attacks can exploit. We characterize this dependence as observer-relative stealth. For recorded event streams, retiming an event within its protected accumulation window leaves the accumulated integer tensor exactly unchanged. Any consumer that receives only this tensor, therefore, sees the same input before and after the attack. We use this exact blind space to construct Null, a gradient-guided timestamp-retiming attack, and define $\mathrm{SC\text{-}ASR}_A(\tau)$ to measure attack success while bounding the change visible to observer A. Across five event-vision datasets, the exact blind space contains retimings that remain exactly hidden from the protected observer while causing victim failure. On DVS Gesture at a 10% event budget, Null reaches $81.56 \pm 5.81\%$ ASR on ConvSNN and $98.67 \pm 0.45\%$ on a GRU while preserving the protected tensor exactly. On DailyDVS-200, a protocol-scale Multi-View Fusion Network variant reaches $99.28 \pm 0.11\%$ exact-null ASR, compared with $9.70 \pm 1.06\%$ for its matched control. In a five-attack comparison, Null is the only method with nonzero attack success at exact observer equality, reaching 81.4% on DVS Gesture and 87.35% on DailyDVS-200. We also search the same exact blind space with an independently implemented constrained projected-gradient optimizer, C-PGD. At matched victim-gradient evaluations, C-PGD reaches $84.50 \pm 2.89\%$ ASR on DVS Gesture and $89.55 \pm 4.39\%$ on DailyDVS-200, again with exact protected equality. Both optimization procedures find retimings with high attack success within the exact same blind space. We then change the temporal observer and test the same attacks. Perturbations that are exactly hidden from the protected observer become visible under shifted, finer, overlapping, and randomized temporal views. Adding observer constraints reduces the real-valued blind-space fraction from 87.5% to 75.0% to 62.5%, while DVS ConvSNN ASR falls from 74.9% to 61.9% to 37.2%. These results show that stealth is not a property of the perturbation alone. It depends on the temporal information available to the observer.

1 INTRODUCTION

Event cameras produce asynchronous streams in which precise timing carries information (Gallego et al., 2022; Gehrig et al., 2019; Peng et al., 2023; Sabater et al., 2023; Zubic et al., 2024). The same stream can, however, be processed at different temporal resolutions. A temporal model may preserve fine event timing, while another component first accumulates events into coarser frames. The two consumers can therefore receive different information from the same recorded events.

We study the security consequence of this mismatch. A timestamp perturbation can change a fine-time victim while leaving a declared coarse temporal observer exactly unchanged. Prior work estab-

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.

lishes timestamps as an attack surface (Yu et al., 2026; Lee & Myung, 2022; Marchisio et al., 2021; Büchel et al., 2022; Bu et al., 2023; Yao et al., 2024; Lin et al., 2025). Count-preserving temporal transformations can also disappear after temporal aggregation in a poisoning setting (Riaño et al., 2026). We make the observer explicit and study whether its blind space contains test-time retimings that cause victim errors under finite event budgets.

Let A denote the temporal representation available to a downstream consumer. Exact stealth to A means

$$A(X') = A(X). \tag{1}$$

Any computation restricted to $A(X)$ receives exactly the same input before and after the attack. A different observer may still receive different information from the same pair of streams. Stealth therefore depends on the temporal information available to the observer.

Coarse temporal accumulation gives a direct discrete construction of this blind space. Moving an event to another fine bin within the same protected window leaves the accumulated integer count exactly unchanged. We search this feasible set with *Null*, a gradient-guided timestamp-retiming attack that keeps spatial location and polarity fixed. Across N-MNIST, DVS128 Gesture, CIFAR10-DVS, N-Caltech101, and DailyDVS-200, the resulting exact blind space contains retimings that cause victim errors. We observe the same behavior on Spiking Heidelberg Digits (SHD), an auditory event-stream benchmark.

DVS Gesture illustrates the effect. At a 10% event budget, exact-null attack success rate (ASR) is $81.56 \pm 5.81\%$ for a convolutional spiking neural network (ConvSNN), $34.06 \pm 3.69\%$ for a spike-element-wise residual network (SEW-ResNet) (Fang et al., 2021), $50.96 \pm 3.27\%$ for our temporal Transformer (Vaswani et al., 2017), and $98.67 \pm 0.45\%$ for a temporal gated recurrent unit (GRU) (Cho et al., 2014). The protected accumulated tensor remains exactly unchanged in every constrained row. When each moved event is restricted to one fine bin, ASR remains $34.7/17.3/58.0/98.7\%$ across the same victims. We report timestamp displacement separately because exact representation equality and physical-time displacement describe different properties of the perturbation.

We next test whether this vulnerability depends on Null's greedy search. Constrained projected gradient descent (C-PGD) provides an independent optimizer over the same unrestricted within-window exact-null feasible set. At matched victim-gradient evaluations on DVS Gesture, C-PGD reaches $84.50 \pm 2.89\%$ ASR and Null reaches $81.56 \pm 5.81\%$. Both satisfy $D_A = 0$, and additional C-PGD steps increase ASR further. A supporting DailyDVS-200 comparison shows the same qualitative result, with the low clean accuracy of that ConvSNN reported explicitly in Section 5.2. Both search procedures find retimings that cause victim errors inside the same exact blind space.

Raw ASR alone does not describe what the observer can see. We define the stealth-constrained attack success rate (SC-ASR)

$$\text{SC-ASR}_A(\tau) = \Pr[\text{victim failure} \wedge D_A(X, X') \leq \tau \mid \text{clean-correct}], \tag{2}$$

where $\tau = 0$ requires exact protected equality. We report the detector-specific undetected attack success rate (UASR) separately. These quantities distinguish victim failure, representation visibility, and detector evasion. In the five-attack DVS study, Yu and Free achieve high UASR despite nonzero deviation in the protected representation. Null satisfies $D_A = 0$ exactly while retaining substantial detector-specific UASR. This separates exact observer-level invisibility from empirical evasion of a particular detector.

An attack hidden from one temporal observer need not remain hidden when the temporal view changes. Canonical-null attacks have zero canonical deviation but become visible under shifted, half-width, overlapping, and randomized temporal views on DVS Gesture and DailyDVS-200. Requiring equality to additional fixed observers reduces the real-valued blind fraction from 87.5% to 75.0% to 62.5%. DVS ConvSNN ASR falls from 74.9% to 61.9% to 37.2%. Temporal information discarded by one observation can reappear under another.

Main contributions. (i) Observer-relative stealth. We formulate test-time stealth relative to a declared temporal observer and separate raw ASR, $\text{SC-ASR}_A(\tau)$, and detector-specific UASR. **(ii) Exact temporal blind space.** We construct exact integer timestamp retimings for coarse temporal accumulation and show that this blind space contains retimings that cause victim errors across datasets and victim families. **(iii) Independent search validation.** Null and C-PGD independently

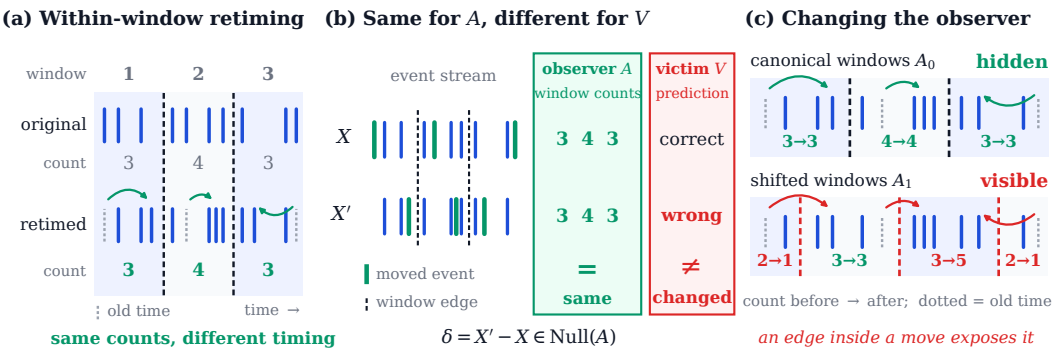


Figure 1: **Observer-relative stealth.** (a) Within-window retiming preserves window counts. (b) The canonical observer A sees no change, while a fine-time victim V can. (c) Shifted windows expose the same perturbation that the canonical windows $A_0 = A$ hide.

find retimings with high attack success within the same exact-null feasible set. **(iv) Observer-dependent visibility.** Changing and intersecting temporal observers changes which perturbations remain hidden, linking observer choice to attack visibility and the size of the common blind space.

2 Related Work and Novelty Boundary

Event-stream attacks. Event-stream attacks modify event values, spatial structure, or timestamps (Marchisio et al., 2021; Lee & Myung, 2022; Büchel et al., 2022; Bu et al., 2023; Yao et al., 2024; Lin et al., 2025). They have been studied with models ranging from asynchronous processing to token and transformer representations (Sekikawa et al., 2019; Sabater et al., 2022; Jiang et al., 2022; Peng et al., 2023; Sabater et al., 2023). Sparse Dynamic Attack (SDA), using potential-dependent surrogate gradients (PDSG), targets sparse perturbations of dynamic spiking neural network (SNN) inputs (Lun et al., 2025; Neftci et al., 2019; Zhou et al., 2023). Yu et al. (Yu et al., 2026) optimize spike retiming under tamper, displacement, and capacity constraints. Temporal Poisoning (Riaño et al., 2026) studies a clean-label training-time backdoor whose timestamp transformation preserves per-pixel and per-polarity counts and disappears after temporal aggregation. Our setting is test-time observer-constrained evasion. We define the observer A, measure exact and tolerance-based stealth, compare independent optimizers within the same exact-null feasible set, and evaluate the same attacks under different temporal observers.

Stealth and temporal representations. Adversarial ML distinguishes attack success from perceptual and detector-level stealth (Frederickson et al., 2018; Franzmeyer et al., 2024). Our $\mathrm{SC\text{-}ASR}_A(\tau)$ measures attack success subject to what observer A sees, while UASR_α measures evasion of a particular detector. These measures capture different properties. An attack can evade a detector while changing the protected representation. Exact A-stealth instead guarantees that the declared observer receives the same representation before and after the attack. This distinction matters because event representations retain different amounts of temporal information, from reconstructed intensity sequences and sparse asynchronous processing to task-specific encodings (Rebecq et al., 2019; Tulyakov et al., 2021; Zhang et al., 2020; Cadena et al., 2021). Driving and optical-flow pipelines can retain fine timing (Gehrig et al., 2021a;b), while converted neuromorphic datasets expose timestamped events directly (Orchard et al., 2015). We therefore define stealth relative to a named observer instead of a single generic event representation.

Representation invariance. Null spaces, Fourier zeros, and Poisson count laws describe information discarded by a temporal observer. For recorded streams, exact integer equality defines the feasible set used by Null and C-PGD and makes observer-level stealth directly measurable. Changing, refining, randomizing, or intersecting temporal observers can expose information hidden from the original observer. Appendix B.1 gives the expanded comparison.

3 OBSERVER-RELATIVE STEALTH AND BLIND SPACES

Let R be the representation available to an observer. We call X' *exactly stealthy to* R *at* X when $R(X') = R(X)$. For a linear observer A, perturbations in $\mathrm{Null}(A)$ form its blind space. Such perturbations can change the underlying event stream without changing the representation seen by the observer.

Proposition 1 (Observer-restricted detection)**.** *If $R(X') = R(X)$, any detector restricted to R receives the same input for the paired clean and attacked samples. A deterministic detector gives the same output. With independent detector randomness, the paired information alone cannot raise the true-positive rate above the false-positive rate.*

Proof. The claim follows directly from equality of the detector inputs. If $R(X') \neq R(X)$, attack-dependent information is present, although a particular detector may still fail to use it. The formal statement and proof are in Appendix A.

Exact invariance for recorded events. Let $X \in \mathbb{N}^{T\times C\times H\times W}$ be a fine-bin event tensor. Assume $T = KS$, giving K coarse windows, and let A sum blocks of S consecutive bins. Moving one event from bin s to s' within the same coarse window, while preserving location and polarity, gives $\delta = e_{s'} - e_s$. For several legal moves, write $\Delta = \sum_j \delta_j$.

Proposition 2 (Exact discrete blind space)**.** *Any finite collection of legal within-window retimings satisfies $A\Delta = 0$, so $A(X + \Delta) = AX$ exactly in integer arithmetic. With all aggregation rows present,* $\dim \mathrm{Null}(A) = KCHW(S - 1)$.

Proof. Each move removes and adds one count within the same aggregation row, so its row sum is zero. Linearity gives $A\Delta = 0$ for any collection of legal moves. Each length-S aggregation row contributes one constraint and $S - 1$ blind degrees of freedom, giving the stated nullity. This exact integer equality defines the feasible set used by Null and C-PGD. The complete proof is in Appendix A.8.

Changing the observer. For linear observers $A_1, \ldots, A_M$,

Proposition 3 (Intersection of blind spaces)**.**

$$\bigcap_i \mathrm{Null}(A_i) = \mathrm{Null}([A_1^\top \ \cdots \ A_M^\top]^\top). \tag{3}$$

Proof. A perturbation lies in the null space of every A_i exactly when their stacked operator annihilates it. Adding observer constraints can only preserve or shrink the common blind space. A retiming hidden from one observer can become visible when another observer retains different temporal information. Section 5.5 tests this prediction with shifted, finer, overlapping, randomized, and intersected observers. The full argument is given in Appendix A.

Continuous and stochastic views. The same observer-relative structure also appears beyond the discrete construction. For a linear temporal kernel, periodic rate perturbations vanish at zeros of its frequency response. Under an inhomogeneous Poisson model, preserving the integrated rate in each disjoint window preserves the joint frame-count distribution. Randomizing accumulation width removes fixed boxcar spectral nulls under the stated conditions. These results require additional assumptions and are separate from the exact recorded-event guarantee used by Null and C-PGD. Their statements, assumptions, and proofs are given in Appendix A.4 onward.

4 NULL-SPACE ATTACK AND THREE-LEVEL EVALUATION

4.1 ATTACK CONSTRUCTION

Legal retiming inside the blind space. Let an event occupy fine bin s in group $q = (k, c, y, x)$, where k indexes the protected coarse window. A legal destination s' remains in the same k and keeps spatial location and polarity fixed. Let n_{ev} denote the number of original events. With event

budget b, at most $\lfloor bn_{\text{ev}} \rfloor$ original events are moved, and each original event is used at most once. For victim loss $\mathcal{L}$ and $G = \nabla_X \mathcal{L}$, the first-order gain of a legal move is

$$\Gamma(q, s \to s') = G_{q,s'} - G_{q,s}.$$

We greedily choose high-gain legal moves and periodically refresh the gradient. Every move contributes $e_{s'} - e_s$ within one aggregation row, so the feasible set guarantees $A(X') = A(X)$ exactly. The optimization searches inside the protected blind space instead of balancing attack loss against a similarity penalty.

Independent exact-null optimizer. To separate the effect of the feasible set from Null's greedy search, we also use constrained projected gradient descent (C-PGD) (Madry et al., 2018). C-PGD maintains a continuous grouped event tensor. It removes the row-common gradient component, takes normalized projected-gradient steps, and projects each aggregation row back to its fixed event-count simplex. An integer transport projection then converts the relaxed solution back to a timestamp-only event stream under the requested moved-event budget. The resulting stream preserves location and polarity, uses the same unrestricted within-window legal moves as Null, and satisfies $A(X') = A(X)$ exactly. With one C-PGD step per progressive budget stage, the number of victim-gradient evaluations through 10% matches Null. Both use three evaluations on DVS Gesture and four on DailyDVS-200. Larger step counts are reported only as optimization-strength sensitivity. Appendix N gives pseudocode and the integer-projection details.

Controls and comparison attacks. We use budget-matched random retiming and, where feasible, an exact signed-displacement-matched control. The latter reproduces each optimized sample's signed shift multiset on randomly chosen legal events. This separates the choice of events from the size and direction of their shifts. The DVS comparison includes free gradient retiming, the official Yu PIL-L_0 implementation (Yu et al., 2026), and SDA/PDSG (Lun et al., 2025) and Yao/Gumbel attacks. Yu, Free, and Null share a common nominal retiming operating point. The native feasible sets and objectives of Yu, SDA, Yao, and Free differ from those of the exact-null attacks. We use these methods to study how fixed existing attacks behave under observer-relative evaluation, not as a same-feasible-set optimizer comparison. The direct optimizer comparison is Null versus C-PGD under the same attack population, legal moves, budget path, and matched gradient-evaluation counts. Appendix G.2 gives the comparison details.

4.2 Three-level evaluation

Raw ASR asks only whether the temporal victim changes from correct to incorrect. For protected observer A, we first measure representation visibility

$$D_A(X, X') = \frac{\|A(X') - A(X)\|_1}{\|A(X)\|_1}, \qquad D_\infty = \|A(X') - A(X)\|_\infty. \tag{4}$$

We then define stealth-constrained attack success

$$\text{SC-ASR}_A(\tau) = \Pr[\text{success} \wedge D_A \le \tau \mid \text{clean-correct}], \tag{5}$$

so $\tau = 0$ requires exact equality to the declared observer. For a fixed generated attack set, varying τ measures how much attack success remains as the observer tolerance is relaxed. The resulting curve is a post-hoc observer-tolerance curve. It is an attack-optimal Pareto frontier only if the attack is re-optimized separately at each τ. Where timestamp displacement is directly comparable, we also report mean absolute displacement d_t and $A(\epsilon) = \Pr[\text{success} \wedge d_t \le \epsilon \mid \text{clean-correct}]$. The two axes answer different questions. The τ axis limits what the declared observer sees. The ϵ axis limits how far events move in physical time. We use $A(\epsilon)$ as a complementary temporal-locality diagnostic, not as a substitute for representation stealth.

A practical defender introduces a third quantity. At clean false-positive rate α, we define the undetected attack success rate (UASR)

$$\text{UASR}_\alpha = \Pr[\text{success} \wedge \text{defender does not fire} \mid \text{clean-correct}]. \tag{6}$$

The detector is calibrated on held-out clean clips and frozen before attack scoring. Raw ASR measures victim failure. SC-ASR_A measures victim failure that also satisfies a declared observer constraint. UASR measures successful attacks that evade a particular detector. We evaluate these quantities separately because a change in representation does not determine whether a particular detector can exploit that change.

Changing the observer. Exact stealth is relative to the declared observer. We apply the same generated attacks to shifted, finer-width, overlapping, and randomized temporal observers and measure the resulting representation deviation separately from detector performance. We also impose equality to intersections of fixed observers and measure how the common blind space and attack success change as observer constraints are added. These experiments directly test the observer-relative prediction that information hidden from one temporal view can reappear under another. They measure recovered information and blind-space reduction without assuming that every detector can exploit the recovered signal. Randomized-observer detection is evaluated separately with schedule-aware and schedule-blind rules. Full observer definitions and defense diagnostics are given in Appendix M.

4.3 Evaluation protocol

Main Null results use three training seeds and several victim families. The unified five-attack studies evaluate Yu PIL-L_0, SDA, Yao/Gumbel, Free, and Null over three seeds using common attack populations within each dataset. ASR and SC-ASR are conditioned on clean-correct inputs. The Null–C-PGD comparison uses the same attack population, legal moves, progressive budget path, and matched victim-gradient evaluations at the primary comparison point. Across seeds we report mean$\pm$std descriptively. For repeated-clip comparisons, uncertainty uses a clip-clustered, seed-aware bootstrap that preserves attack pairing and each seed's clean-correct conditioning. Wilson intervals are retained only in per-population tables and are not treated as independent and identically distributed (IID) across repeated seed evaluations. Detector thresholds are calibrated on held-out clean clips and fixed before attack scoring. Appendix C gives preprocessing, hyperparameters, clean-pool construction, and statistical details.

5 Experiments

We first test whether the exact blind space contains retimings that cause victim errors across datasets and victim families. We then repeat the search with constrained projected gradient descent (C-PGD), evaluate existing attacks with raw ASR and $\mathrm{SC\text{-}ASR}_A(\tau)$, and separate representation visibility from detector-specific UASR. Finally, we change the temporal observer and test whether the same perturbations remain hidden. Unless noted, ASR is conditioned on clean-correct samples and headline values are mean$\pm$std over three seeds. Appendix C gives preprocessing, model definitions, budgets, per-seed tables, and confidence intervals.

5.1 Broad exact-null vulnerability

The protected operator is coarse temporal accumulation, and Null moves events only within their original protected windows. Every optimized row therefore satisfies $A(X') = A(X)$ exactly in integer arithmetic. Table 1 shows the common 10% operating point for four aligned event-vision datasets and four victim families. The ConvSNN is our compact convolutional SNN; SEW-ResNet follows spike-element-wise residual learning (Fang et al., 2021); the temporal GRU follows gated recurrent units (Cho et al., 2014); and our Event Transformer uses a standard Transformer encoder (Vaswani et al., 2017). Null exceeds the corresponding random control in every dataset–victim pair, although sensitivity inside the blind space varies substantially by architecture.

Table 1: **Exact-null attack across four aligned event-vision datasets.** Optimized ASR / random-control ASR (%). N-MNIST is reported separately below because its original and aligned protocols use different victim sets. DVS ConvSNN/SEW use displacement-conditioned random controls. DVS Transformer/GRU and the later dataset protocols use exact signed-displacement matching where reported. All optimized rows satisfy $D_A = D_\infty = 0$.

Dataset	ConvSNN	SEW	Transformer	GRU
DVS Gesture	81.6 / 0.6	34.1 / 0.8	51.0 / 1.1	98.7 / 2.2
CIFAR10-DVS	93.4 / 8.6	65.0 / 2.1	86.0 / 5.1	99.0 / 22.0
N-Caltech101	44.5 / 1.7	36.2 / 1.3	70.7 / 2.3	89.5 / 8.8
DailyDVS-200	84.3 / 9.8	34.2 / 9.3	98.3 / 13.6	98.9 / 16.2

On DVS Gesture, the protected CoarseFrameFormer reaches $93.81 \pm 0.22\%$ clean accuracy and its logits remain bitwise identical on exact-null rows (Appendix C.6). N-MNIST provides an additional aligned protocol check, DailyDVS-200 adds protocol-scale ACTION-Net (Wang et al., 2021), MVFNet (Wu et al., 2021), Swin-T (Liu et al., 2021), and TimeSformer (Bertasius et al., 2021)

variants, and SHD tests the same discrete mechanism outside vision. These supporting results are reported in Appendices I–K.

5.2 Null versus C-PGD in the same feasible set

Figure 2 compares Null with C-PGD inside the same exact-null feasible set. Null and C-PGD use the same ConvSNN victim, attack population, timestamp-only within-window moves, and requested 10% moved-event budget. At the primary comparison point, C-PGD uses one projected-gradient step per progressive budget stage, matching Null's victim-gradient evaluations through 10%.

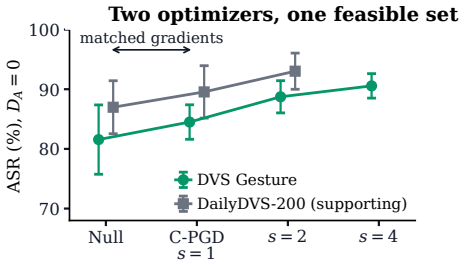


Figure 2: **Two optimizers, one exact-null feasible set.** ASR at $D_A = 0$ (mean±std, three seeds). C-PGD $s = 1$ matches Null's victim-gradient evaluations; $s = 2, 4$ test optimization sensitivity. DailyDVS-200 is supporting because its ConvSNN has low clean accuracy.

At matched victim-gradient evaluations, C-PGD reaches $84.50 \pm 2.89\%$ versus $81.56 \pm 5.81\%$ for Null on DVS, a paired difference of $+2.94$ points with 95% CI $[0.85, 5.12]$. DailyDVS gives $89.55 \pm 4.39\%$ versus $86.98 \pm 4.44\%$, with paired CI $[-0.29, 5.56]$. Additional C-PGD steps increase ASR while preserving $D_A = 0$. The DVS comparison is primary; the DailyDVS ConvSNN result is supporting because its clean accuracy is $13.55 \pm 0.60\%$. Appendix G.2 gives the full comparison, including the progressive-versus-single-pass DailyDVS reconciliation.

5.3 Existing attacks under observer-relative evaluation

The five-attack study keeps each method's native objective and scores its generated attacks with the same observer-relative metric. We compare Yu PIL-L_0 (Yu et al., 2026), Sparse Dynamic Attack (SDA) with potential-dependent surrogate gradients (Lun et al., 2025), Yao/Gumbel (Yao et al., 2024), unconstrained Free retiming, and Null. Because their feasible sets differ, τ is applied *post hoc*. Figure 3 is therefore a descriptive SC-ASR(τ) comparison, not an attack-optimal Pareto frontier.

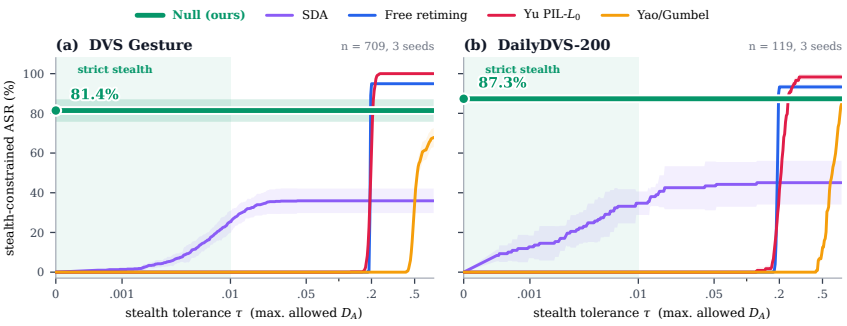


Figure 3: **Existing attacks under a post-hoc observer tolerance.** DVS Gesture (left) and DailyDVS-200 (right). τ filters fixed generated attacks by protected-representation deviation. Null is generated inside the exact-null feasible set and is the only method here with nonzero success at $\tau = 0$.

At $\tau = 0$, Null retains 81.4% attack success on DVS and 87.35% on DailyDVS, while the other four generated attack sets have zero success. Raw ASR is higher for several alternatives, showing that

Table 2: **Changing the observer changes visibility and the common blind space.** The top reports mean relative deviation of the same Null attacks under alternative observers over three seeds. The bottom reports DVS attacks re-optimized while enforcing equality to additional fixed phase observers.

Same attack, different observer. Relative representation deviation						
Dataset	Canonical	Shift-1	Shift-2	Shift-4	Half-width	Random mean
DVS Gesture	**0.000**	0.081	0.116	0.137	0.144	0.103
DailyDVS-200	**0.000**	0.078	0.111	0.131	0.143	0.099

DVS observer intersections. Blind-space size and attack success			
Observer constraints	Phase 0	Phases 0+2	Phases 0+2+4
Blind fraction (%)	87.5	75.0	62.5
ConvSNN ASR (%)	74.9	61.9	37.2
SEW ASR (%)	38.0	29.5	20.7

victim failure and observer-constrained success are different quantities. Temporal locality is also separate: in the unified seed-0 DVS run, $A(200\,\mathrm{ms})$ is 51.88% for Null, 7.11% for Yu, and 0% for Free. Full per-seed, footprint, temporal-radius, and statistical results are in Appendices G.1, G.2, and L.

5.4 Representation visibility versus operational detection

Representation visibility does not determine detector evasion. Figure 4 therefore reports undetected attack success rate (UASR) separately from raw ASR and exact $\mathrm{SC\text{-}ASR}_A(0)$. The same clean-only detector protocol is applied to all five DVS attacks.

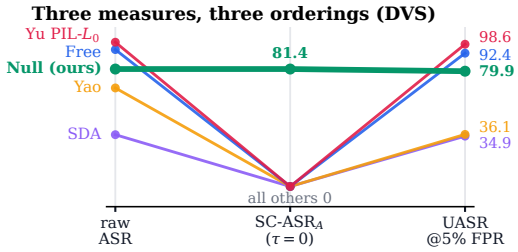


Figure 4: **Three measures, three orderings on DVS Gesture.** Raw ASR measures victim failure, $\mathrm{SC\text{-}ASR}_A(0)$ requires exact observer equality, and UASR is successful attack rate after a detector at 5% nominal false-positive rate. The same attack can rank very differently under the three questions.

Yu and Free retain high UASR despite changing the protected representation, whereas Yao is much easier to detect. Null satisfies $D_A = 0$ exactly while retaining 79.9% UASR. Yao's detector AUROC is 0.929; the remaining attacks are near chance in this DVS experiment. Complete AUROC, calibration, and DailyDVS diagnostic results are in Appendix L.

5.5 Cross-observer visibility

We next keep the generated Null attacks fixed and change only the temporal observer. Exact equality to the canonical observer need not survive a different temporal view. Table 2 shows zero canonical deviation but nonzero deviation under shifted, half-width, and randomized observers on both DVS and DailyDVS. Overlapping and multiscale diagnostics in Appendix M show the same qualitative effect.

The top panel supports the relational claim that canonical invisibility does not transfer automatically to another observer. The bottom panel tests the corresponding geometric prediction. Adding phase constraints shrinks the real-valued blind space and reduces ConvSNN and SEW attack success. GRU remains near saturation in this experiment (Appendix M), so the intersection experiment does not establish a universal defense. Randomized schedules expose the fixed canonical-null DVS attacks in about 97.5% of realized schedules. This is a representation-exposure result, not a complete defense benchmark. In the controlled randomized-accumulation study, a schedule-aware detector reaches AUROC about $0.97/0.999$, while a schedule-blind statistic remains near chance (Appendix H). The temporal observation rule therefore changes both the available evidence and the feasible blind space.

5.6 CONTROLS AND ABLATIONS

Targeted controls test simpler explanations. With a one-fine-bin cap, DVS ConvSNN/SEW/Transformer/GRU ASR remains 34.7/17.3/58.0/98.7%. A matched-effort CIFAR10-DVS comparison shows that only Null preserves the protected representation, and all 12 optimized-versus-control comparisons remain significant after Holm correction. In the controlled study, null-subspace projected gradient descent at $\epsilon = .25$ lowers LIF accuracy to 0.33% while the protected frame CNN remains at 97.80%, with no protected-label flips in $21{,}000$ attacks. Further ablations are in Appendices C.4, M, and G.4.

6 DISCUSSION AND CONCLUSION

Damaging directions inside an exact blind space. Coarse temporal accumulation discards within-window timing, yet changes along these directions can affect models that retain finer timing. Null leaves the protected integer tensor exactly unchanged while causing victim errors. On DVS Gesture at 10%, exact-null ASR reaches 81.56% for ConvSNN and 98.67% for GRU. C-PGD independently finds retimings with similar attack success inside the same exact-null feasible set. The vulnerability follows the blind space, not Null's particular search procedure. For the DVS temporal GRU and Event Transformer, matched controls preserve the optimized signed-shift multiset while assigning it to different legal event identities. All 6,336 sample-level matches complete exactly with zero fallback, and the controls remain far below the optimized attacks.

Attack success, observer stealth, and detection are distinct. Yu and Free achieve high raw ASR and UASR while changing the protected representation; Null retains attack success at exact equality, while Yao has larger representation deviation and is easier to detect on DVS. Raw ASR, $\mathrm{SC\text{-}ASR}_A(\tau)$, and UASR measure victim failure, observer-constrained success, and detector evasion, respectively. None should be used as a proxy for the others. Protected CoarseFrameFormer logits remain bitwise identical on the DVS exact-null evaluations, while the operational detector can use information outside that representation.

Changing the observer changes what remains hidden. Canonical-null attacks become visible under shifted, finer, overlapping, and randomized temporal views. Adding fixed observer constraints reduces the real-valued blind fraction from 87.5% to 75.0% to 62.5%, while DVS ConvSNN ASR falls from 74.9% to 61.9% to 37.2%. In the controlled randomized-schedule study, a schedule-aware detector reaches AUROC about 0.97/0.999 while a schedule-blind statistic remains near chance. Visibility therefore depends on the temporal information available to the observer, although these experiments do not establish observer diversification as a complete defense. GRU remains near saturation in the fixed-observer intersection experiment, so this is not a universal defense result.

Temporal locality is a separate constraint. The tolerance τ bounds protected-representation change, whereas $A(\epsilon)$ measures success within a timestamp-displacement radius. For the three retiming methods in the seed-0 DVS analysis, $A(200\,\mathrm{ms})$ is 51.88% for Null, 7.11% for Yu, and 0% for Free. Under a stricter one-fine-bin DVS cap, ASR remains 34.7/17.3/58.0/98.7% for ConvSNN/SEW/Transformer/GRU. Exact observer equality and temporal displacement therefore describe different properties of the perturbation. We therefore report representation deviation, event footprint, and timestamp displacement separately.

Statistical support and scope. Across three training seeds where applicable, the paired C-PGD$-$Null difference is $+2.94$ points on DVS (95% CI $[0.85, 5.12]$) and $+2.57$ on DailyDVS ($[-0.29, 5.56]$). SHD preserves exact protected equality with substantial ASR, extending the discrete mechanism beyond event vision; vulnerability remains architecture dependent. The breadth study retains low-sensitivity victims rather than selecting only architectures on which the attack is strong. Our threat model is digital post-sensor timestamp retiming, not an over-the-air physical attack. The appendix reports the matching, equality, population, and calibration checks.

Conclusion. Stealth depends jointly on the perturbation and the observer. A retimed stream can be unchanged to one temporal representation, cause errors in a finer-time model, and become visible under another view. $\mathrm{SC\text{-}ASR}_A(\tau)$ measures observer-constrained success, $A(\epsilon)$ measures temporal locality, and UASR measures detector evasion. Observer intersections show how additional temporal information changes the common blind space.

AI USE STATEMENT

We used generative AI tools for research brainstorming and conceptual framing, discussion of mathematical arguments and experimental design, code drafting and debugging, analysis scripts, result interpretation, and language editing. The authors made the final methodological decisions, ran all reported experiments, checked numerical claims against the saved result files, reviewed AI-assisted mathematical arguments, code, and prose, and take responsibility for the final manuscript, claims, citations, and artifacts.

ETHICS STATEMENT

This work studies adversarial manipulation of event-camera streams and has dual-use implications. The empirical study is limited to digital timestamp retiming of recorded events. It does not demonstrate an over-the-air physical attack. N-MNIST, DVS128 Gesture, CIFAR10-DVS, N-Caltech101, and DailyDVS-200 are public benchmark datasets, and we collected no new human-subject data. DVS128 Gesture contains recordings of participants performing gestures. We use only the released benchmark events and labels for recognition experiments.

REPRODUCIBILITY STATEMENT

Appendix A contains full proofs and assumptions. Appendix C documents preprocessing, model definitions, budgets, controls, and ablations. Appendix H gives the controlled spectral study. The supplementary submission will include the experiment code and saved per-seed result tables used for the reported values.

REFERENCES

Arnon Amir, Brian Taba, David Berg, Timothy Melano, Jeffrey McKinstry, Carmelo Di Nolfo, Tapan Nayak, Alexander Andreopoulos, Guillaume Garreau, Marcela Mendoza, Jeff Kusnitz, Michael Debole, Steve Esser, Tobi Delbrück, Myron Flickner, and Dharmendra Modha. A low power, fully event-based gesture recognition system. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 7243–7252, 2017.

Gedas Bertasius, Heng Wang, and Lorenzo Torresani. Is space-time attention all you need for video understanding? In *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pp. 813–824, 2021.

Tong Bu, Jianhao Ding, Zecheng Hao, and Zhaofei Yu. Rate gradient approximation attack threats deep spiking neural networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 7896–7906, 2023. doi: 10.1109/CVPR52729.2023.00763.

Julian Büchel, Gregor Lenz, Yalun Hu, Sadique Sheik, and Martino Sorbaro. Adversarial attacks on spiking convolutional neural networks for event-based vision. *Frontiers in Neuroscience*, 16: 1068193, 2022. doi: 10.3389/fnins.2022.1068193.

Pablo Rodrigo Gantier Cadena, Yeqiang Qian, Chunxiang Wang, and Ming Yang. Spade-e2vid: Spatially-adaptive denormalization for event-based video reconstruction. *IEEE Transactions on Image Processing*, 30:2488–2500, 2021. doi: 10.1109/TIP.2021.3052070.

Kyunghyun Cho, Bart van Merriënboer, Caglar Gulcehre, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, and Yoshua Bengio. Learning phrase representations using RNN encoder–decoder for statistical machine translation. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 1724–1734, 2014.

Benjamin Cramer, Yannik Stradmann, Johannes Schemmel, and Friedemann Zenke. The heidelberg spiking data sets for the systematic evaluation of spiking neural networks. *IEEE Transactions on Neural Networks and Learning Systems*, 33(7):2744–2757, 2022. doi: 10.1109/TNNLS.2020.3044364.

Xiao Du, Wanli Shi, Xiaohan Zhao, Yang Cao, Bin Gu, and Tieru Wu. Raw event-based adversarial attacks for spiking neural networks with configurable latencies. *Neural Networks*, 193:108026, 2026. doi: 10.1016/j.neunet.2025.108026.

Wei Fang, Zhaofei Yu, Yanqi Chen, Tiejun Huang, Timothée Masquelier, and Yonghong Tian. Deep residual learning in spiking neural networks. In *Advances in Neural Information Processing Systems*, volume 34, 2021.

Tim Franzmeyer, Stephen McAleer, Joao F. Henriques, Jakob Foerster, Philip H. S. Torr, Adel Bibi, and Christian Schroeder de Witt. Illusory attacks: Information-theoretic detectability matters in adversarial attacks. In *International Conference on Learning Representations (ICLR)*, 2024.

Christopher Frederickson, Michael Moore, Glenn Dawson, and Robi Polikar. Attack strength vs. detectability dilemma in adversarial machine learning. *arXiv preprint arXiv:1802.07295*, 2018.

Guillermo Gallego, Tobi Delbrück, Garrick Orchard, Chiara Bartolozzi, Brian Taba, Andrea Censi, Stefan Leutenegger, Andrew J. Davison, Jörg Conradt, Kostas Daniilidis, and Davide Scaramuzza. Event-based vision: A survey. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 44(1): 154–180, 2022.

Daniel Gehrig, Antonio Loquercio, Konstantinos G. Derpanis, and Davide Scaramuzza. End-to-end learning of representations for asynchronous event-based data. In *Proc. IEEE/CVF Int. Conf. Computer Vision (ICCV)*, 2019.

Mathias Gehrig, Willem Aarents, Daniel Gehrig, and Davide Scaramuzza. Dsec: A stereo event camera dataset for driving scenarios. *IEEE Robotics and Automation Letters*, 2021a. doi: 10.1109/LRA.2021.3068942.

Mathias Gehrig, Mario Millhäusler, Daniel Gehrig, and Davide Scaramuzza. E-raft: Dense optical flow from event cameras. In *International Conference on 3D Vision*, 2021b.

Wulfram Gerstner, Werner M. Kistler, Richard Naud, and Liam Paninski. *Neuronal Dynamics: From Single Neurons to Networks and Models of Cognition*. Cambridge University Press, 2014.

Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778, 2016.

Bin Jiang, Zhihao Li, M. Salman Asif, Xun Cao, and Zhan Ma. Event transformer. In *arXiv preprint arXiv:2204.05172*, 2022.

Xavier Lagorce, Garrick Orchard, Francesco Galluppi, Bertram E. Shi, and Ryad B. Benosman. HOTS: A hierarchy of event-based time-surfaces for pattern recognition. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 39(7):1346–1359, 2017.

Wooju Lee and Hyun Myung. Adversarial attack for asynchronous event-based data. In *Proc. AAAI Conf. Artificial Intelligence*, 2022.

Hongmin Li, Hanchao Liu, Xiangyang Ji, Guoqi Li, and Luping Shi. CIFAR10-DVS: An event-stream dataset for object classification. *Frontiers in Neuroscience*, 11:309, 2017. doi: 10.3389/fnins.2017.00309.

Guixu Lin, Muyao Niu, Qingtian Zhu, Zhengwei Yin, Zhuoxiao Li, Shengfeng He, and Yinqiang Zheng. Adversarial attacks on event-based pedestrian detectors: A physical approach. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pp. 5227–5235, 2025. doi: 10.1609/aaai.v39i5.32555.

Ze Liu, Yutong Lin, Yue Cao, Han Hu, Yixuan Wei, Zheng Zhang, Stephen Lin, and Baining Guo. Swin transformer: Hierarchical vision transformer using shifted windows. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 10012–10022, 2021.

Li Lun, Kunyu Feng, Qinglong Ni, Ling Liang, Yuan Wang, Ying Li, Dunshan Yu, and Xiaoxin Cui. Towards effective and sparse adversarial attack on spiking neural networks via breaking invisible surrogate gradients. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 3540–3551, 2025.

Aleksander Madry, Aleksandar Makelov, Ludwig Schmidt, Dimitris Tsipras, and Adrian Vladu. Towards deep learning models resistant to adversarial attacks. In *Proc. Int. Conf. Learning Representations (ICLR)*, 2018.

Alberto Marchisio, Giacomo Pira, Maurizio Martina, Guido Masera, and Muhammad Shafique. DVS-Attacks: Adversarial attacks on dynamic vision sensors for spiking neural networks. In *Proc. Int. Joint Conf. Neural Networks (IJCNN)*, 2021.

Emre O. Neftci, Hesham Mostafa, and Friedemann Zenke. Surrogate gradient learning in spiking neural networks. *IEEE Signal Processing Magazine*, 36(6):51–63, 2019.

Garrick Orchard, Ajinkya Jayawant, Gregory K. Cohen, and Nitish Thakor. Converting static image datasets to spiking neuromorphic datasets using saccades. *Frontiers in Neuroscience*, 9:437, 2015.

Yansong Peng, Yueyi Zhang, Zhiwei Xiong, Xiaoyan Sun, and Feng Wu. GET: Group event transformer for event-based vision. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 6038–6048, 2023.

Henri Rebecq, René Ranftl, Vladlen Koltun, and Davide Scaramuzza. High speed and high dynamic range video with an event camera. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2019.

Roberto Riaño, Gorka Abad, Stjepan Picek, and Aitor Urbieta. Temporal poisoning: Clean-label backdoors via event redistribution in SNNs. *arXiv preprint arXiv:2607.28075*, 2026.

Alberto Sabater, Luis Montesano, and Ana C. Murillo. Event transformer: A sparse-aware solution for efficient event data processing. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, pp. 2677–2686, 2022.

Alberto Sabater, Luis Montesano, and Ana C. Murillo. Event transformer$^+$: A multi-purpose solution for efficient event data processing. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(12):16013–16020, 2023. doi: 10.1109/TPAMI.2023.3311336.

Yusuke Sekikawa, Kosuke Hara, and Hideo Saito. Eventnet: Asynchronous recursive event processing. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 3887–3896, 2019.

Amos Sironi, Manuele Brambilla, Nicolas Bourdis, Xavier Lagorce, and Ryad Benosman. HATS: Histograms of averaged time surfaces for robust event-based object classification. In *Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition (CVPR)*, 2018.

Stepan Tulyakov, Daniel Gehrig, Stamatios Georgoulis, Julius Erbach, Mathias Gehrig, Yuanyou Li, and Davide Scaramuzza. Time lens: Event-based video frame interpolation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 16155–16164, 2021.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. Attention is all you need. In *Advances in Neural Information Processing Systems*, volume 30, pp. 5998–6008, 2017.

Qi Wang, Zhou Xu, Yuming Lin, Jingtao Ye, Hongsheng Li, Guangming Zhu, Syed Afaq Ali Shah, Mohammed Bennamoun, and Liang Zhang. Dailydvs-200: A comprehensive benchmark dataset for event-based action recognition. In *European Conference on Computer Vision (ECCV)*, 2024.

Zhengwei Wang, Qi She, and Aljosa Smolic. ACTION-Net: Multipath excitation for action recognition. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 13214–13223, 2021.

Wenhao Wu, Dongliang He, Tianwei Lin, Fu Li, Chuang Gan, and Errui Ding. MVFNet: Multi-view fusion network for efficient video recognition. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, pp. 2943–2951, 2021. doi: 10.1609/aaai.v35i4.16401.

Yanmeng Yao, Xiaohan Zhao, and Bin Gu. Exploring vulnerabilities in spiking neural networks: Direct adversarial attacks on raw event data. In *Proc. European Conf. Computer Vision (ECCV)*, 2024.

Yi Yu, Qixin Zhang, Shuhan Ye, Xun Lin, Qianshan Wei, Kun Wang, Wenhan Yang, Dacheng Tao, and Xudong Jiang. Time is all it takes: Spike-retiming attacks on event-driven spiking neural networks. In *Proc. Int. Conf. Learning Representations (ICLR)*, 2026.

Song Zhang, Yu Zhang, Zhe Jiang, Dongqing Zou, Jimmy Ren, and Bin Zhou. Learning to see in the dark with events. In *European Conference on Computer Vision*, pp. 666–682, 2020.

Zhaokun Zhou, Yuesheng Zhu, Chao He, Yaowei Wang, Shuicheng Yan, Yonghong Tian, and Li Yuan. Spikformer: When spiking neural network meets transformer. In *International Conference on Learning Representations (ICLR)*, 2023.

Nikola Zubic, Mathias Gehrig, and Davide Scaramuzza. State space models for event cameras. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5819–5828, 2024.

A PROOFS AND THEORETICAL DETAILS

A.1 CONTINUOUS AND STOCHASTIC STATEMENTS

This section gives the continuous-time and stochastic results summarized in Section 3.

Theorem 1 (Kernel-zero criterion). *For the linear response $y_k = \int g(t_k - t)\, dN(t)$ and a nonzero additive periodic rate perturbation $\delta\lambda(t) = \alpha \cos(2\pi f t + \varphi)$ with $\alpha \neq 0$, $\Delta\mathbb{E}[y_k] = 0$ for every k and phase φ iff $\widehat{g}(f) = 0$.*

Corollary 1 (Boxcar nulls). *For a boxcar of width Δ, nonzero zeros occur at $f = m/\Delta$. An idealized leaky integrate-and-fire (LIF) kernel has nonzero gain at every finite f.*

Theorem 2 (Poisson count indistinguishability). *For an inhomogeneous Poisson process, preserving integrated rate in every disjoint accumulation window preserves the joint distribution of the complete frame-count sequence.*

Theorem 3 (Randomized accumulation). *For positive random boxcar width W with characteristic function ϕ_W,*

$$\mathbb{E}_W |\widehat{g}_W(f)|^2 = \frac{1 - \operatorname{Re} \phi_W(2\pi f)}{2\pi^2 f^2}. \tag{7}$$

If W has a density, this energy is positive for every $f \neq 0$.

A.2 PROOF OF PROPOSITION 1

Let a detector restricted to R be written as $D(R(X), U)$, where U denotes any detector randomness. If $R(X') = R(X)$, then for every fixed realization u,

$$D(R(X'), u) = D(R(X), u).$$

A deterministic detector is the special case without U and returns the same output on the paired clean and attacked samples. If the clean and attacked evaluations use independent randomness with the same law, their output distributions are identical because their representation inputs are identical. The paired representation therefore contains no information that can increase true-positive rate above false-positive rate. This statement applies only to detectors that factor through R. A detector with access to the raw event stream or another representation is not covered.

A.3 BASELINE-RATE SCOPE

For a general multiplicative perturbation $\lambda(t) \mapsto \lambda(t)(1 + a(t))$, the attack-induced change is

$$\Delta\mathbb{E}[y_k] = \int g(t_k - t)\lambda(t)a(t)\, dt.$$

The transform condition applies to the product λa. Writing $\lambda(t) = \lambda_0$ over the support of g factors out the baseline and gives the pure-tone form used in Theorem 1. If the baseline changes inside a window, a sinusoidal a need not remain single-frequency after multiplication by λ. In the original controlled check, a linear-ramp baseline produced a residual about 7×10^{-3} of the frame mean and a full saccade profile produced about 9×10^{-2}. The recorded-event experiments bypass this approximation by enforcing exact per-window sums.

A.4 PROOF OF THEOREM 1

With $\delta\lambda(t) = \alpha \cos(2\pi f t + \varphi)$, Campbell's formula gives

$$\Delta\mathbb{E}[y_k] = \alpha \int g(t_k - t) \cos(2\pi f t + \varphi)\, dt.$$

Substitute $u = t_k - t$:

$$\begin{aligned}
\Delta\mathbb{E}[y_k] &= \alpha \int g(u) \cos(2\pi f(t_k - u) + \varphi)\ du \\
&= \alpha\ \operatorname{Re}\left\{ e^{j(2\pi f t_k + \varphi)} \int g(u) e^{-j2\pi f u}\, du \right\} \\
&= \alpha |\widehat{g}(f)| \cos(2\pi f t_k + \varphi + \angle \widehat{g}(f))\,.
\end{aligned}$$

If $\widehat{g}(f) = 0$, the response vanishes for every k and φ. Conversely, assume $\widehat{g}(f) \neq 0$. Set $\varphi = -2\pi f t_k - \angle\widehat{g}(f)$ for any chosen k. Then $\Delta\mathbb{E}[y_k] = \alpha|\widehat{g}(f)| \neq 0$. Linearity extends the claim to additive perturbations whose spectral support lies in the zero set.

A.5 Proof of Corollary 1

For $g = \mathbf{1}_{[0,\Delta)}$,

$$\widehat{g}(f) = \int_0^\Delta e^{-j2\pi f u}\, du = \Delta\,\mathrm{sinc}(f\Delta)e^{-j\pi f\Delta},$$

whose nonzero zeros satisfy $f\Delta \in \mathbb{Z} \setminus \{0\}$. For $h(t) = e^{-t/\tau}\mathbf{1}_{t\geq 0}/\tau$,

$$\widehat{h}(f) = \frac{1}{1 + j2\pi f\tau}, \qquad |\widehat{h}(f)| = \left(1 + (2\pi f\tau)^2\right)^{-1/2} > 0.$$

At the first boxcar null $f = 1/\Delta$, the LIF gain is $(1 + (2\pi\tau/\Delta)^2)^{-1/2}$, which equals about 0.62 when $\tau/\Delta = 0.2$.

A.6 Proof of Theorem 2

For an inhomogeneous Poisson process with intensity λ, the count in a Borel set B is Poisson with mean $\int_B \lambda$, and counts in disjoint sets are independent. Let $y_k = N([k\Delta, (k+1)\Delta))$. If the perturbation has zero integral in every coarse window, then each attacked window has the same Poisson parameter as its clean counterpart. Independence holds in both cases, so the product law of the full count sequence is identical. For any test ψ measurable with respect to $(y_k)_k$, $\mathbb{E}_{H_1}\psi = \mathbb{E}_{H_0}\psi$. Its power equals its size.

This conclusion is exact for Poisson window counts and for any statistic computed only from those counts. For a non-Poisson sensor process, equality of the window integrals guarantees equality of first moments but does not by itself guarantee equality of higher-order count statistics. Proposition 2 supplies the stronger realized-representation statement used in the recorded-event experiments.

A.7 Proof of Proposition 3

A perturbation δ is invisible to every monitored linear representation precisely when $A_i\delta = 0$ for all i. Stacking these equations gives

$$\begin{bmatrix} A_1 \\ \vdots \\ A_M \end{bmatrix} \delta = 0,$$

which proves Eq. (3). Appending rows to a linear operator cannot decrease its rank and cannot increase the dimension of its null space.

A.8 Proof of Proposition 2

For one coarse window and fixed (c, y, x), the aggregation row is $\mathbf{1}_S^\top$. Moving one event from s to s' produces $\delta = e_{s'} - e_s$, hence

$$\mathbf{1}_S^\top \delta = 1 - 1 = 0.$$

All other coarse rows are unchanged, so $A\delta = 0$. For a finite collection of legal moves, define the aggregate perturbation $\Delta = \sum_i \delta_i$. Linearity gives $A\Delta = 0$, and therefore $A(X + \Delta) = AX$ exactly in integer arithmetic. Any downstream function C of AX receives the identical argument.

A.9 Nullity calculation for Proposition 2

Each fixed tuple (k, c, y, x) contributes one row that sums S distinct fine-time coordinates. Rows for distinct tuples have disjoint support, so the $KCHW$ rows are linearly independent and $\mathrm{rank}(A) = KCHW$. The input dimension is $KSCHW$. Rank-nullity gives

$$\dim \mathrm{Null}(A) = KSCHW - KCHW = KCHW(S - 1).$$

This is the dimension over the reals. Legal timestamp moves occupy an integer subset further restricted by event occupancy and nonnegativity.

A.10 Derivation of the mismatch law, Eq. (8)

For fixed $m \in \mathbb{Z} \setminus \{0\}$ and small ϵ, let $f = (m+\epsilon)/\Delta$. Then

$$\operatorname{sinc}(m+\epsilon) = \frac{\sin(\pi(m+\epsilon))}{\pi(m+\epsilon)} = \frac{(-1)^m \sin(\pi\epsilon)}{\pi(m+\epsilon)}.$$

Hence

$$|\widehat{g}(f)|^2 = \Delta^2 \frac{\sin^2(\pi\epsilon)}{\pi^2(m+\epsilon)^2} = \Delta^2 \frac{\epsilon^2}{m^2}\left(1 + O(\epsilon)\right),$$

using $\sin(\pi\epsilon) = \pi\epsilon(1 + O(\epsilon^2))$ and $(m+\epsilon)^{-2} = m^{-2}(1 + O(\epsilon))$. Dividing $\Delta^2|\widehat{h}(f)|^2$ by this quantity gives

$$\rho(f) = \frac{m^2}{\epsilon^2}\left(1 + (2\pi f\tau)^2\right)^{-1}(1 + O(\epsilon)). \tag{8}$$

At $\epsilon = 0$, the denominator is zero and ρ diverges.

A.11 Proof of Theorem 3

For a boxcar of random width W,

$$\widehat{g}_W(f) = \int_0^W e^{-j2\pi f u}\, du = \frac{1 - e^{-j2\pi fW}}{j2\pi f}.$$

Thus

$$|\widehat{g}_W(f)|^2 = \frac{|1 - e^{-j2\pi fW}|^2}{4\pi^2 f^2} = \frac{1 - \cos(2\pi fW)}{2\pi^2 f^2}.$$

Taking the expectation and using $\mathbb{E}[\cos(2\pi fW)] = \operatorname{Re}\phi_W(2\pi f)$ gives Eq. (7). If this expectation is zero, then $\cos(2\pi fW) = 1$ almost surely, so W lies on the countable lattice $\{m/f : m \in \mathbb{Z}\}$ almost surely. A distribution with a density assigns zero mass to that set. The expectation is strictly positive for $f \neq 0$.

For $W \sim \mathrm{Unif}[\Delta(1-r), \Delta(1+r)]$,

$$\phi_W(\omega) = e^{j\omega\Delta}\operatorname{sinc}(\omega\Delta r/\pi),$$

so at $f = m/\Delta$,

$$\mathbb{E}|\widehat{g}_W(m/\Delta)|^2 = \frac{\Delta^2\big(1 - \operatorname{sinc}(2mr)\big)}{2\pi^2 m^2}.$$

For $m = 1$ this increases with r on $[0, 1/2]$. At $r = 0.25$ it equals $1.84 \times 10^{-2}\Delta^2$. A continuous width law is needed. A non-degenerate law supported on $\{1/f, 2/f\}$ still gives zero visibility at frequency f.

A.12 Schedule-aware deflection: full derivation

Let window k have width W_k and center t_k^c. Integrating the attacked rate over that window gives

$$\mu_k = \lambda_0 W_k + \lambda_0 \alpha\, G(W_k)\cos(2\pi f t_k^c + \varphi), \qquad G(W) = \frac{\sin(\pi fW)}{\pi f}. \tag{9}$$

At $f = m/\Delta$, write $W/\Delta = 1 + u$. If the law of W is symmetric about Δ, then u is symmetric and

$$\sin(\pi m(1+u)) = (-1)^m \sin(\pi m u),$$

an odd function of u. Hence $\mathbb{E}[G(W)] = 0$.

For a schedule-blind coherent statistic, use the nominal schedule rather than the realized width and set $\widetilde{z}_k = (y_k - \lambda_0\Delta)/\sqrt{\lambda_0\Delta}$. If the nominal phase is independent of W_k, the attack-induced mean factorizes and symmetry about Δ gives $\mathbb{E}[G(W_k)] = 0$ at the old null. This exact cancellation need not hold for arbitrary randomized schedules or for a statistic normalized by $1/\sqrt{W_k}$.

Now use the realized gain. Set

$$z_k = \frac{y_k - \lambda_0 W_k}{\sqrt{\lambda_0 W_k}}, \qquad w_k = \frac{G(W_k)e^{-j2\pi f t_k^c}}{\sqrt{W_k}}, \qquad S = \sum_k z_k w_k.$$

For conditionally independent Poisson counts in disjoint windows, under the null $\mathrm{Var}(S) = \sum_k G(W_k)^2/W_k$. From Eq. (9), the attacked mean of z_k is

$$\mathbb{E}[z_k] = \sqrt{\lambda_0}\alpha \frac{G(W_k)}{\sqrt{W_k}} \cos(2\pi f t_k^c + \varphi).$$

Multiplying by w_k and summing gives

$$\mathbb{E}[S] = \sqrt{\lambda_0}\alpha \sum_k \frac{G(W_k)^2}{W_k} \cos(2\pi f t_k^c + \varphi) e^{-j2\pi f t_k^c}.$$

Assume the widths are independent, or more generally mixing, and that the accumulated centers t_k^c yield sufficient phase dispersion at $2f$. Then the normalized double-frequency sum vanishes asymptotically. For finite K this is an approximation, leaving

$$\mathbb{E}[S] \approx \frac{\sqrt{\lambda_0}\alpha}{2} e^{j\varphi} \sum_k \frac{G(W_k)^2}{W_k}.$$

The squared deflection is

$$d^2 = \frac{|\mathbb{E}[S]|^2}{\mathrm{Var}(S)} = \frac{\alpha^2 \lambda_0}{4} \sum_k \frac{G(W_k)^2}{W_k} \approx \frac{K\alpha^2\lambda_0 \mathbb{E}|\widehat{g}_W(f)|^2}{4\Delta}. \tag{10}$$

For widths concentrated around Δ, the approximation replaces $1/W_k$ by $1/\Delta$ and the empirical mean of $G(W_k)^2$ by its expectation. Theorem 3 makes this expectation positive, so d^2 grows linearly with the number of observed windows.

A.13 Additional representation and temporal-budget details

Exact stealth corresponds to the zero-radius case of the representation-space ball $\mathcal{B}_\tau^A(X)$. The radius τ is a monitor tolerance rather than a second event-retiming budget. The same perturbed event stream can therefore have different stealth values for different representations A. For the block-sum operator, Fourier carriers occupy structured directions inside the larger aggregation null space. Legal timestamp moves occupy an integer subset further restricted by event occupancy and nonnegativity.

A B-bin voxel grid changes the temporal kernel. At the first null of the original frame, normalized visibility becomes $|\mathrm{sinc}(1/B)|$. It is 0 at $B = 1$, 0.64 at $B = 2$, and 0.98 at $B = 10$. Finer temporal summaries close one blind spot and create their own higher-frequency nulls. For the N-MNIST, DVS Gesture, CIFAR10-DVS, N-Caltech101, and DailyDVS-200 tensorizations used in Section 5, Proposition 2 gives 312,120, 1,146,880, 573,440, 573,440, and 573,440 real-valued blind directions, respectively.

The representation budget τ and temporal budget ϵ constrain different properties. The first limits what the protected observer sees. The second limits how far the attack moves events in time. A perturbation may lie in $\mathrm{Null}(A)$ and hence have $D_A = 0$ while still differing substantially in d_t. Reporting both $\text{SC-ASR}_A(\tau)$ and $A(\epsilon)$ separates observer-relative stealth from the amount of retiming required to achieve it.

B Supplementary comparisons and figures

This section collects two supporting views that are useful for interpreting the main result but are not needed for the core experimental sequence: a direct comparison with prior event-security work and a visualization of the continuous transfer-function mechanism. The first clarifies the novelty boundary; the second shows why the same temporal modulation can be hidden from one representation and visible to another.

Table 3: Comparison with event-security work. "Protected observation" denotes a guarantee on the realized representation available to a stated observer. Collapsed counts are one specific observation. "Exact chosen R" denotes the general invariance condition $R(X') = R(X)$, which reduces to count equality when R is frame accumulation.

Work	Perturbation	Count-preserving	Protected observation	Visibility / defense analysis
DVS-Attacks (Marchisio et al., 2021)	event-sequence perturbation	no constraint	–	sensor-filter study
Lee–Myung (Lee & Myung, 2022)	time shifts + added events	no	–	–
Yao et al. (Yao et al., 2024)	direct raw-event attack	no constraint	–	–
Du et al. (Du et al., 2026)	raw-event, configurable-latency attack	no exact count constraint	–	latency-robust optimization
Yu et al. (Yu et al., 2026)	timing-only retiming	yes	–	–
Temporal Poisoning (Riaño et al., 2026)	timestamp redistribution	yes	collapsed counts	detector study
This work	exact-null retiming	yes	**exact chosen R**	**observer-relative metric + cross-observer tests**

B.1 Event-security comparison

Table 3 separates three axes that are easy to conflate. Yu et al. study timing-only, count-preserving retiming, while Temporal Poisoning uses a count-preserving timestamp transformation in a training-time backdoor setting. Our test-time setting instead declares the protected representation R, measures success subject to $R(X') = R(X)$ or a tolerance around it, and then changes the observer to test whether the same perturbation remains hidden.

B.2 Transfer-function visualization

Figure 5 visualizes the continuous counterpart of the discrete blind-space argument. At boxcar harmonics, the accumulated-frame response is zero while the idealized LIF response remains nonzero. Near a boxcar null, a small frequency mismatch restores frame visibility according to Eq. (8). This visualization is explanatory; the recorded-event attacks use the exact integer guarantee of Proposition 2.

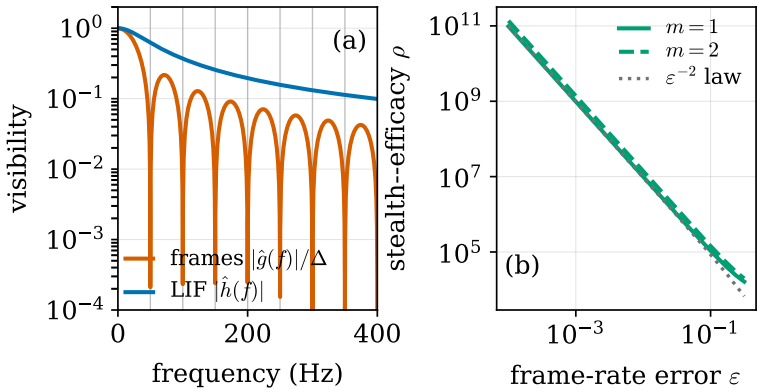


Figure 5: Visibility of one temporal modulation to two consumers. Boxcar accumulation has zeros at frame-rate harmonics. The LIF membrane has nonzero gain. The right panel shows the ϵ^{-2} growth of the stealth–efficacy ratio near a boxcar null.

C Experimental details

C.1 N-MNIST protocol

We parse the original binary N-MNIST files directly. Events are binned into 150 steps of 2 ms over a 300 ms recording. Two polarities are retained separately on the 34×34 sensor grid. The coarse

representation sums every 10 fine bins, producing 15 windows. We train three seeds for a study-specific frame CNN, a leaky integrate-and-fire (LIF) MLP (Gerstner et al., 2014), and a temporal GRU (Cho et al., 2014). The headline attack uses 1,000 balanced test samples (100 per digit) and unique-event budgets $\{5, 10, 20, 30\}\%$.

The LIF model uses an exponential decay corresponding to $\tau = 4$ ms. The strict unique-event attack maintains an availability tensor containing only original events that have never been moved. Four gradient refresh rounds are used. The random control uses the same unique-event budget and samples a legal destination uniformly inside the original coarse window.

Table 4: N-MNIST ASR (%) over three seeds.

Victim	Attack	5%	10%	20%	30%
LIF	gradient retiming	8.66±0.15	23.12±0.98	54.81±1.77	75.92±1.71
LIF	random retiming	0.20±0.10	0.17±0.12	0.34±0.21	0.55±0.26
GRU	gradient retiming	66.08±2.45	90.70±1.64	98.44±1.05	99.45±0.26
GRU	random retiming	0.17±0.12	0.24±0.15	0.37±0.06	0.41±0.20

The mean adversarial absolute displacement decreases mildly with budget because later moves use lower-gain destinations. For LIF it is 11.71, 11.16, 10.54, and 10.10 ms across the four budgets. For the GRU it is 10.54, 9.97, 9.50, and 9.23 ms. The random baseline has mean displacement about 7.33 ms. This N-MNIST baseline matches the event budget but not the displacement distribution. We therefore do not use it for the event-identity isolation claim. On DVS Gesture, exact signed-displacement matching is used for the GRU and Transformer event-identity isolation claim. ConvSNN and SEW use displacement-conditioned controls. CIFAR10-DVS, N-Caltech101, and DailyDVS-200 use the exact-control protocol reported in their respective sections.

The raw-stream export routine materializes count-space retiming back into timestamp-edited event lists while preserving event count, x, y, and polarity. Three saved examples pass exact reconstruction checks against the attacked count tensor.

C.2 DVS Gesture protocol

Each DVS128 Gesture segment is stored as an $N \times 4$ event array (x, y, p, t), and train and test subject IDs do not overlap. Sensor coordinates remain 128×128 in the source and are deterministically pooled by 2×2 blocks to 64×64 for model input. Each gesture duration D_i is normalized to 160 temporal bins, so the physical width of a fine bin is $D_i/160$. Eight fine bins form one coarse frame.

Three seeds are trained for a frame ResNet-18-style model (He et al., 2016), our compact ConvSNN, a SEW-ResNet-18-style SNN (Fang et al., 2021), a temporal GRU (Cho et al., 2014), our Event Transformer using a standard Transformer encoder (Vaswani et al., 2017), and our CoarseFrameFormer. Main attack budgets are $\{2, 5, 10, 20\}\%$. The full test attack set contains 264 samples. The SEW implementation is a compact PyTorch reimplementation of the element-wise residual principle, not the official reference code. The temporal GRU, Event Transformer, and CoarseFrameFormer use the same cached event tensors and subject-disjoint split.

Table 5: DVS Gesture ASR (%) over three seeds. "Disp.-cond. random" samples attack-derived absolute shift magnitudes, then chooses random legal event identities/directions.

Victim	Attack	2%	5%	10%	20%
ConvSNN	gradient retiming	24.98±3.43	56.45±5.26	81.56±5.81	90.98±1.27
ConvSNN	disp.-cond. random	0.14±0.25	0.14±0.24	0.56±0.49	1.70±0.75
SEW-ResNet	gradient retiming	10.44±3.75	23.22±3.74	34.06±3.69	48.17±1.40
SEW-ResNet	disp.-cond. random	0.27±0.23	0.55±0.23	0.82±0.70	1.77±0.92

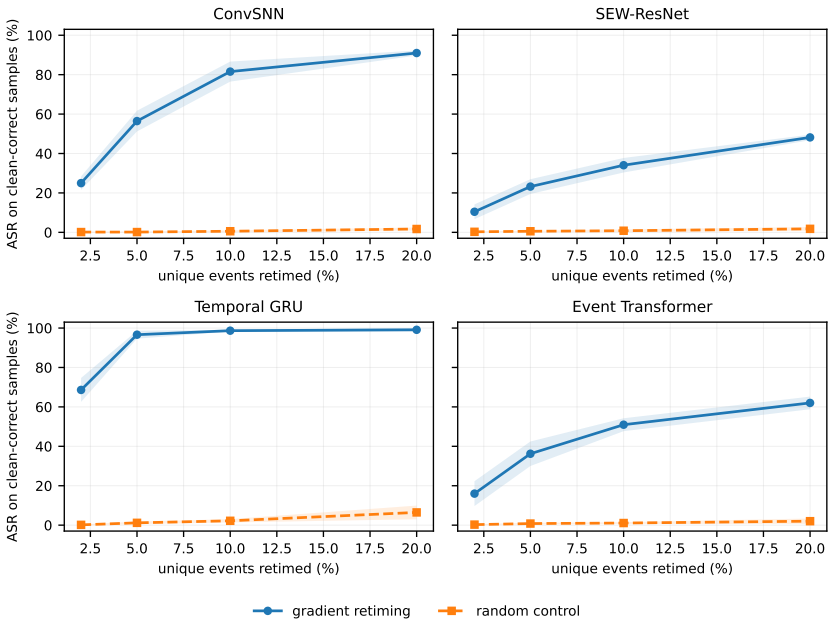


Figure 6: DVS Gesture ASR versus the fraction of unique original events retimed. Solid curves show optimized null-space retiming and dashed curves show random controls. GRU and Transformer controls reproduce each sample's signed-shift multiset exactly. ConvSNN and SEW use displacement-conditioned controls. Lines are means over three seeds and bands show ± 1 standard deviation.

C.3 Exact displacement-matched random control

For the temporal GRU and Event Transformer, we use an exact signed-displacement matcher. Let the optimized attack on a sample produce signed fine-bin shifts $\{d_i\}_{i=1}^{M}$. The control randomly assigns these same shifts to distinct original events for which $s_i + d_i$ remains inside the protected coarse window. The random and adversarial perturbations therefore have the same event budget, signed displacement histogram, mean absolute displacement, maximum displacement, and coarse-window support. Only the selected event identities differ.

Verification is exhaustive over the saved evaluation. For each of the temporal GRU and Event Transformer victims, all 264 DVS Gesture test samples are checked at four budgets and three seeds. This gives $2 \times 264 \times 4 \times 3 = 6{,}336$ sample-level matched controls. Every signed histogram matches exactly and the fallback count is zero. Table 6 reports the resulting ASR.

Table 6: DVS Gesture ASR (%, mean±std over 3 seeds) for the recurrent and attention-based victims. "Exact random" reproduces the adversarial signed-shift multiset sample by sample.

Victim	Attack	2%	5%	10%	20%
Event Transformer	gradient	16.00±6.30	36.23±6.22	50.96±3.27	62.01±3.21
Event Transformer	exact random	0.27±0.23	0.81±0.40	1.08±0.45	2.01±1.37
Temporal GRU	gradient	68.61±6.07	96.63±2.01	98.67±0.45	99.11±0.90
Temporal GRU	exact random	0.15±0.26	1.18±0.66	2.21±0.84	6.46±3.47

Stronger coarse consumer. CoarseFrameFormer operates only on the 20 aggregated DVS Gesture frames and obtains clean accuracies 93.56%, 93.94%, and 93.94% over seeds 0–2 ($93.81 \pm 0.22\%$). During the GRU and Transformer attack runs, both the frame ResNet and CoarseFrameFormer are evaluated on every constrained adversarial tensor. For every seed and budget, their logits are bitwise identical to clean logits and their prediction flip rates are zero.

C.4 MAXIMUM-SHIFT ABLATION

The maximum-shift ablation uses seed 0, the full DVS Gesture test set, a 10% unique-event budget, and the same progressive $2 \rightarrow 5 \rightarrow 10\%$ optimization path as the main attack. Figure 7 reports the attack under restricted temporal movement. Every condition retains exact coarse-frame equality. The unrestricted condition agrees exactly with the main seed-0 result for ConvSNN, SEW, and GRU. The Transformer differs by 1.22 percentage points. We therefore flag that consistency check and do not use the ablation to validate its main-table value.

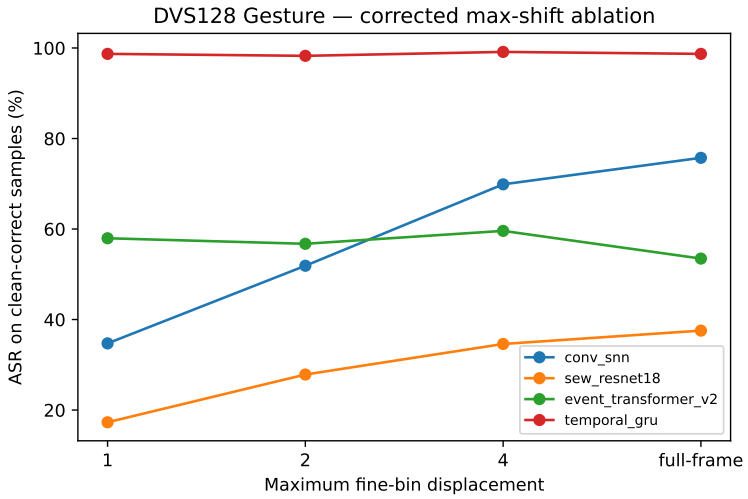


Figure 7: DVS Gesture maximum-shift ablation at a 10% unique-event budget (seed 0) using the same progressive path as the main attack. One fine bin corresponds to a sample-dependent physical duration. The mean absolute displacement in the 1-bin condition is 42.46 ms.

Table 7: DVS Gesture shift ablation, seed 0. ASR (%) under the main progressive path.

Victim	1 bin	2 bins	4 bins	full window
ConvSNN	34.73	51.88	69.87	75.73
SEW-ResNet	17.30	27.85	34.60	37.55
Event Transformer	57.96	56.73	59.59	53.47
Temporal GRU	98.70	98.27	99.13	98.70

C.5 COST-OF-STEALTH DIAGNOSTICS

The cost-of-stealth studies are smaller single-seed experiments. The N-MNIST comparison uses four gradient rounds for both arms on the same 100 samples at a 20% event budget. For the LIF victim, exact-null and free retiming reach 50% and 100% ASR. For the temporal GRU both reach 100% ASR. The exact-null arm has zero coarse-frame difference in both cases, while free retiming reaches maximum integer differences 14 and 15.

The CIFAR10-DVS comparison uses the same progressive $1 \rightarrow 2 \rightarrow 5 \rightarrow 10 \rightarrow 20\%$ path for both arms on 250 samples. At 10%, null/free ASR is $96.6/98.9\%$ for ConvSNN, $71.8/85.3\%$ for SEW, $88.1/79.7\%$ for Event Transformer, and $100.0/95.7\%$ for the GRU. Every null-space row has zero coarse difference and bitwise-identical CoarseFrameFormer logits. Free retiming changes the protected prediction on 28.0–44.4% of these samples and has maximum integer coarse-frame differences 161–192. Removing the null constraint does not uniformly improve attack efficacy across architectures. In every case, however, it removes the exact protected-representation guarantee in this diagnostic.

C.6 ATTENTION-BASED AND RECURRENT VICTIMS

The Event Transformer, named `event_transformer_v2` in the released code, combines a residual GroupNorm spatial encoder, token projection to 384 dimensions, a depthwise/pointwise temporal convolutional mixing stage, and an eight-layer, eight-head Transformer encoder. Its clean test

accuracies are 92.80%, 94.70%, and 93.56% over the three seeds ($93.69 \pm 0.95\%$). The attention-based victim is evaluated at $93.69 \pm 0.95\%$ clean accuracy, close to the strongest event models in the study.

The DVS temporal GRU provides a non-spiking sequence baseline and a cross-dataset counterpart to the N-MNIST GRU. Its clean accuracies are 87.50%, 84.47%, and 84.47% ($85.48 \pm 1.75\%$). Its attack results are reported with the same strict unique-event budget and exact signed-displacement-matched control as the Transformer.

D CIFAR10-DVS PROTOCOL

We use the original CIFAR10-DVS recordings (Li et al., 2017) after removing AEDAT2 trigger and external-event records before DVS decoding. Deterministic 2×2 spatial pooling maps the 128×128 sensor to 64×64. Each recording is represented by $T = 80$ fine bins with $S = 8$ bins per protected window and ten coarse frames. The split uses 800 train and 200 test recordings per class with split seed 2027. The balanced attack subset contains 100 test recordings per class.

The protected model is CoarseFrameFormer. Temporal victims are ConvSNN, SEW-ResNet-18 style, Event Transformer, and Temporal GRU. Budgets are $\{1, 2, 5, 10, 20\}\%$ unique original events. Legal moves remain inside the original coarse window. Controls include uniform random retiming and exact signed-displacement-matched retiming. The protocol also includes a 10% maximum-shift sweep over $\{1, 2, 4, \text{full}\}$ fine bins and a free-retiming cost-of-stealth diagnostic.

Table 8: CIFAR10-DVS clean accuracy and ASR (%, mean±std over three seeds). Random columns use the exact signed-displacement-matched control.

Victim	Clean acc.	5% ASR	5% random	10% ASR	10% random
ConvSNN	69.20±0.39	82.50±3.36	4.32±0.92	93.42±2.07	8.59±1.66
SEW-ResNet	72.02±0.33	46.23±1.73	1.70±0.10	64.98±3.60	2.11±0.11
Temporal GRU	55.70±0.35	97.72±0.47	11.67±0.51	98.98±0.22	22.04±1.46
Event Transformer	69.83±1.26	78.64±0.28	3.71±0.65	86.00±0.39	5.10±1.19

All constrained attack and exact-control rows have zero integer coarse-tensor difference, bitwise-identical CoarseFrameFormer logits, and zero protected-consumer prediction flips.

E N-CALTECH101 PROTOCOL AND RESULTS

N-Caltech101 (Orchard et al., 2015) is represented with $T = 80$ fine bins, eight bins per protected window, and a fixed aspect-preserving letterbox from 240×180 to 64×64. We use all 101 classes, an $80/20$ split with seed 2027, three training seeds, and a deterministic proportionally stratified attack subset of $1{,}000$ test streams. The victim registry and budgets are the same as CIFAR10-DVS. Table 9 reports the complete three-seed result.

Table 9: N-Caltech101 ASR (%, mean±std over three seeds). "Exact" is the signed-displacement-matched random control.

Victim	Attack	1%	2%	5%	10%	20%
ConvSNN	optimized	12.08±0.63	19.81±1.69	30.89±2.09	44.48±1.57	60.01±1.28
ConvSNN	exact	0.42±0.09	0.73±0.24	1.25±0.32	1.67±0.60	1.57±0.16
SEW-ResNet	optimized	7.45±1.85	14.58±1.03	26.89±1.59	36.23±3.39	46.71±3.56
SEW-ResNet	exact	0.71±0.62	0.52±0.54	0.80±0.54	1.27±0.71	2.46±1.08
Transformer	optimized	26.89±2.26	45.96±3.36	60.10±2.73	70.66±2.83	80.01±2.18
Transformer	exact	0.77±0.08	0.88±0.09	1.70±0.37	2.25±0.35	5.33±0.50
GRU	optimized	35.52±1.92	59.05±1.63	79.96±3.57	89.48±3.34	93.77±2.27
GRU	exact	2.15±0.32	2.82±0.20	6.44±0.55	8.78±1.35	14.68±1.01

F DAILYDVS-200 PROTOCOL AND RESULTS

DailyDVS-200 (Wang et al., 2024) is decoded from the released AEDAT4 recordings and letterboxed from 320×240 to 64×64. We use $T = 80$, eight fine bins per protected window, all 200 classes, three seeds, and a proportionally stratified $1{,}000$-stream attack subset. The split contains 31 training, 9 test, and 7 validation subjects with no subject overlap after the released validation/test overlap is

removed from validation. The train/test recording overlap is zero. Table 10 reports the complete three-seed attack result. ConvSNN clean accuracy is only $13.55 \pm 0.60\%$, so its ASR is supporting rather than headline evidence. The remaining fine-time victims reach 27.31–34.25% clean accuracy on the 200-way task.

Table 10: DailyDVS-200 ASR (%, mean±std over three seeds). "Exact" is the signed-displacement-matched random control.

Victim	Attack	1%	2%	5%	10%	20%
ConvSNN	optimized	34.39±5.21	50.30±8.38	70.37±5.20	84.29±6.23	94.89±2.43
ConvSNN	exact	4.75±1.60	5.75±0.50	7.15±1.31	9.81±2.01	15.81±5.14
SEW-ResNet	optimized	11.35±0.99	17.46±1.32	27.11±2.50	34.23±3.26	46.14±1.83
SEW-ResNet	exact	3.64±1.84	5.19±1.34	5.62±0.92	9.30±1.81	11.27±0.77
Transformer	optimized	65.24±1.18	85.50±1.07	94.34±1.13	98.27±0.96	99.38±0.84
Transformer	exact	3.00±1.04	5.12±1.52	8.20±1.35	13.65±3.18	24.75±1.96
GRU	optimized	66.33±1.27	87.24±1.24	95.25±1.11	98.88±0.36	99.25±0.39
GRU	exact	2.38±0.81	3.75±0.71	7.25±0.83	16.15±1.42	37.43±1.10

All optimized and exact-control DailyDVS-200 rows preserve the coarse tensor exactly. CoarseFrameFormer logits are bitwise identical and its prediction flip rate is zero in every saved constrained row.

G DIRECT RETIMING COMPARISON AND REPRESENTATION-SPACE METRIC

The direct comparison uses the official Yu et al. (2026) PIL-L_0, free gradient retiming, and our null-space retiming on the DVS128 Gesture ConvSNN. The full run is complete for three seeds (239, 235, and 235 clean-correct samples). For the nonnegative count tensor used here, $\|A(X)\|_1$ equals the total event-unit count, so D_A in Eq. (4) is computed exactly from the saved coarse-frame L_1 difference divided by the saved event-unit count. Across seeds, raw ASR is $100.0 \pm 0.0\%$ for PIL-L_0, $95.06 \pm 0.23\%$ for free retiming, and $81.56 \pm 5.81\%$ for null-space retiming. Exact-stealth SC-ASR is 0, 0, and $81.56 \pm 5.81\%$, respectively. For interpretation, define $\mathrm{SPR}_A(\tau) = \Pr[D_A \le \tau \mid \text{clean-correct}]$ and $\mathrm{ASR}_{\mathrm{pass}}(\tau) = \Pr[\text{success} \mid D_A \le \tau, \text{clean-correct}]$. Whenever the latter is defined, $\text{SC-ASR}_A(\tau) = \mathrm{SPR}_A(\tau)\,\mathrm{ASR}_{\mathrm{pass}}(\tau)$. We also report stealth retention $\text{SC-ASR}_A(\tau)/\mathrm{ASR}_{\mathrm{raw}}$ when raw ASR is nonzero. At $\tau = 0$, null-space retiming has 100% pass rate and retention, while PIL-L_0 and free retiming have zero pass rate and retention in the complete seed.

Table 11: Three-seed DVS direct comparison: SC-ASR (%, mean±std). The raw row is unconstrained ASR.

τ	Yu PIL-L_0	Free	Null
0	0.0±0.0	0.0±0.0	81.56±5.81
0.18	5.65±1.79	0.0±0.0	81.56±5.81
0.19	23.42±1.12	0.99±0.49	81.56±5.81
0.195	41.20±2.06	59.93±2.03	81.56±5.81
0.20	62.34±1.49	95.06±0.23	81.56±5.81
0.205	80.11±4.08	95.06±0.23	81.56±5.81
Raw	100.0±0.0	95.06±0.23	81.56±5.81

Table 12: $\text{SC-ASR}_A(\tau)$ (%) over the full measured post-hoc representation-budget grid for the complete direct-comparison seed. Values in brackets are pointwise 95% Wilson intervals over 239 clean-correct samples.

τ	Yu PIL-L_0	Free retiming	Null-space retiming
0	0.00 [0.00,1.58]	0.00 [0.00,1.58]	74.90 [69.03,79.97]
0.01	0.00 [0.00,1.58]	0.00 [0.00,1.58]	74.90 [69.03,79.97]
0.05	0.00 [0.00,1.58]	0.00 [0.00,1.58]	74.90 [69.03,79.97]
0.1	0.00 [0.00,1.58]	0.00 [0.00,1.58]	74.90 [69.03,79.97]
0.15	0.00 [0.00,1.58]	0.00 [0.00,1.58]	74.90 [69.03,79.97]
0.17	0.84 [0.23,3.00]	0.00 [0.00,1.58]	74.90 [69.03,79.97]
0.18	5.02 [2.90,8.57]	0.00 [0.00,1.58]	74.90 [69.03,79.97]
0.19	23.01 [18.13,28.75]	1.26 [0.43,3.62]	74.90 [69.03,79.97]
0.195	39.33 [33.35,45.65]	61.92 [55.63,67.85]	74.90 [69.03,79.97]
0.2	62.34 [56.05,68.24]	94.98 [91.43,97.10]	74.90 [69.03,79.97]
0.205	80.75 [75.28,85.25]	94.98 [91.43,97.10]	74.90 [69.03,79.97]
0.21	92.47 [88.41,95.18]	94.98 [91.43,97.10]	74.90 [69.03,79.97]
0.22	96.65 [93.54,98.29]	94.98 [91.43,97.10]	74.90 [69.03,79.97]
0.23	99.58 [97.67,99.93]	94.98 [91.43,97.10]	74.90 [69.03,79.97]
0.24	100.00 [98.42,100.00]	94.98 [91.43,97.10]	74.90 [69.03,79.97]

The seed-0 table above gives Wilson intervals. Table 11 is the completed three-seed headline comparison. The strict point $\tau = 0$ is one operating point of the monitor. The complete grid shows how much realized attack efficacy remains as the allowed representation distortion is relaxed. The three-seed raw ASRs are 100.0%, $95.06 \pm 0.23\%$, and $81.56 \pm 5.81\%$ for PIL-L_0, free retiming, and null-space retiming. Their exact-stealth SC-ASRs are 0, 0, and $81.56 \pm 5.81\%$.

G.1 Temporal locality of the retiming attacks

The representation budget τ constrains what is visible to the protected observer but does not constrain how far events move in time. We also report

$$A(\epsilon) = \Pr[\text{attack succeeds and } d_t(X, X') \leq \epsilon \mid \text{clean-correct}],$$

where d_t is mean absolute timestamp displacement. The unified runner contains five attacks, but the temporal-radius analysis covers Yu PIL-L_0, free retiming, and Null. At the seed-0 10% DVS operating point, mean displacement is 276.9 ms for Yu, 2.565 s for free retiming, and 179.5 ms for Null. Table 13 reports selected radii. Adapted SDA/PDSG and Yao/Gumbel remain in the five-method $\mathrm{SC\text{-}ASR}_A(\tau)$ evaluation but are not included in this completed temporal-radius diagnostic. These values supersede displacement values from an earlier direct-comparison implementation. The representation-space τ results are unaffected because they are recomputed from the protected tensors.

Table 13: Temporal-radius success for the three timestamp-retiming methods in the unified DVS seed-0 run at the 10% operating point. Values are $A(\epsilon)$ in percent.

ϵ (ms)	Yu PIL-L_0	Free retiming	Null-space retiming
100	0.42	0.00	1.67
200	7.11	0.00	51.88
300	68.20	0.00	74.48
404	96.65	0.00	74.48
600	99.58	0.00	74.90
1000	100.00	0.42	74.90
3000	100.00	74.48	74.90

The τ and ϵ analyses answer different questions. The τ axis measures observer-relative stealth, while ϵ measures temporal movement. In the unified run, Null remains exactly invariant and requires less temporal movement than free retiming. Yu achieves high temporal-radius success but does not satisfy exact protected equality.

G.2 Null–C-PGD comparison details

The direct Null–C-PGD comparison uses the same victim family and exact-null feasible set. Table 14 separates this comparison from the heterogeneous five-attack evaluation. Null and C-PGD

Table 14: Threat-model comparison. "Same" means identical between Null and C-PGD in the direct exact-null experiment. Prior attacks retain their native constraints and are therefore evaluated post hoc rather than treated as a common feasible-set contest.

Method	timestamp-only	$x/y/p$ fixed	exact A	event budget	displacement rule	role
Null	yes	yes	yes	10% upper	within window	exact-null optimizer
C-PGD	yes	yes	yes	10% upper	within window	exact-null optimizer
Free	yes	yes	no	$\approx$10%	native/free	post-hoc evaluation
Yu PIL-L_0	yes	yes	no chosen-A constraint	$\approx$10% realized	native PIL constraints	post-hoc evaluation
SDA / Yao	native attack	native attack	no chosen-A constraint	native realized	native	post-hoc evaluation

use identical attack populations and clean-correct masks within each dataset and seed. DVS uses the 264-clip released test population with 239/235/235 clean-correct clips across seeds. DailyDVS uses the frozen 1,000-stream attack subset with 116/134/131 clean-correct clips. The matched comparison uses three victim-gradient evaluations on DVS and four on DailyDVS. C-PGD uses continuous projected-gradient updates followed by row-simplex projection and integer transport; Null uses discrete greedy move ranking.

The two DailyDVS ConvSNN Null values reported in the paper use the same frozen 1,000-clip attack subset, the same checkpoints, and identical clean-correct sample UIDs: 116/134/131 clips across the three seeds. The broad experiment reports $84.29 \pm 6.23\%$ at 10% from a single-pass budget path. The direct Null–C-PGD experiment reports $86.98 \pm 4.44\%$ after the progressive $1\% \rightarrow 2\% \rightarrow$

$5\% \rightarrow 10\%$ path with four gradient refreshes, with the defense pipeline enabled. The 2.69-point difference therefore comes from the attack schedule rather than a different evaluation population. The saved manifest is `attack_subset_manifest.csv` (SHA-256 prefix `7bf9a4c4`).

G.3 Statistical tests

For repeated-clip, repeated-seed results, we use a 20,000-replicate clip-clustered, seed-aware bootstrap. Each replicate resamples original evaluation clip IDs, computes ASR separately within each seed using that seed's clean-correct mask, and then averages the seed rates. Paired method differences use the same resampled clips. On DVS Gesture, Null has 81.56% exact-null ASR with 95% CI $[77.24, 85.63]\%$, while matched-gradient C-PGD has 84.50% with $[80.33, 88.45]\%$. The paired C-PGD$-$Null difference is 2.94 points with $[0.85, 5.12]$. On DailyDVS-200, the corresponding values are 86.98% $[82.32, 91.38]$, 89.55% $[85.14, 93.60]$, and a 2.57-point difference with $[-0.29, 5.56]$. Mean$\pm$std across the three training seeds is retained as a descriptive summary. Pooled Wilson intervals remain in saved single-population tables but are not used as primary across-seed uncertainty.

Paired McNemar tests on the complete original direct-comparison seed confirm the raw-ASR differences of free retiming and PIL-L_0 versus null-space retiming ($p = 5.6 \times 10^{-10}$ and 1.7×10^{-18}). Across CIFAR10-DVS, N-Caltech101, and DailyDVS-200, all 12 optimized-versus-exact-control comparisons remain significant after Holm correction. The largest adjusted value is $p = 0.0119$.

G.4 Attack-generation cost

On an NVIDIA H100 80GB, the full DailyDVS-200 progressive $1 \rightarrow 2 \rightarrow 5 \rightarrow 10 \rightarrow 20\%$ attack path takes 467 ± 1 s for ConvSNN, 1490 ± 41 s for SEW-ResNet, 272 ± 9 s for the Event Transformer, and 315 ± 10 s for the GRU across three seeds. These times cover attack generation on the same 1,000-stream subset and exclude model training and preprocessing. The current implementation uses dense fine-time tensors, so its cost increases with temporal and spatial resolution.

For recorded DVS streams, randomized schedules are evaluated on fixed canonical-null attacks. About 97.5% of realized schedules expose a perturbation that is invisible to the canonical accumulator. Because the victim is not retrained under randomized accumulation, this experiment measures observer visibility rather than complete randomized-consumer robustness. Appendix H separately evaluates a schedule-aware detector.

H Controlled spectral validation

The recorded-event experiments establish exact integer invariance on event streams. A controlled study separately isolates the spectral mechanism and randomized accumulation under the assumptions of Section 3. We generate saccade-derived activity over 60 substeps of 2 ms, aggregate every ten substeps ($\Delta = 20$ ms), and compare a frame CNN with an LIF network at $\tau = 4$ ms. The activity tensor in this controlled study is continuous-valued. We use this continuous-valued study only to validate the mechanism. The headline N-MNIST, DVS Gesture, CIFAR10-DVS, N-Caltech101, and DailyDVS-200 experiments use discrete recorded events.

H.1 Invisible-subspace attack and the price of the constraint

For the boxcar observer, the discrete invisible subspace contains all perturbations whose fine-time entries sum to zero inside each coarse window. A sinusoidal null carrier is a structured Fourier direction in this space. A white-box null-PGD attack instead optimizes the victim loss and projects each iterate back into the zero-sum subspace. Table 15 gives the complete controlled attack comparison.

Across the 21 invisible-subspace configurations (7 settings $\times$ 3 seeds, 1,000 streams each), none of the 21,000 attacked streams changes a frame-CNN label. The largest relative logit change is 2×10^{-7}. At $\epsilon = .25$, the constrained and free attacks both saturate the fine-time victim, but the free attack has a frame footprint nearly six orders of magnitude larger. At $\epsilon = .10$, the constraint has a measurable efficacy cost, consistent with the smaller real-event diagnostics in Section 5.

Table 15: Controlled attack study (mean±std over 3 seeds, 1,000 streams/seed). δ_F is the relative change of the accumulated frame representation. "Flips" is the percentage of frame-CNN labels that change. Rows labeled null lie in the protected aggregation null space.

Attack	LIF acc. (%)	frame-CNN acc. (%)	δ_F	flips
none	96.90 ± 0.08	97.80 ± 0.43	–	–
null tone, $a = 2$	96.50 ± 0.29	97.80 ± 0.43	2×10^{-7}	0
null carrier, $a = 2$	96.57 ± 0.33	97.80 ± 0.43	2×10^{-7}	0
null-PGD, $\epsilon = .10$	50.73 ± 0.53	97.80 ± 0.43	3×10^{-7}	0
null-PGD, $\epsilon = .25$	0.33 ± 0.12	97.80 ± 0.43	3×10^{-7}	0
null-PGD, $\epsilon = .50$	0.00 ± 0.00	97.80 ± 0.43	3×10^{-7}	0
free PGD, $\epsilon = .10$	5.43 ± 0.57	95.80 ± 0.22	0.08	2.1%
free PGD, $\epsilon = .25$	0.00 ± 0.00	90.97 ± 0.52	0.19	7.0%

H.2 Transfer-function mechanism and synchronization error

At $\tau/\Delta = 0.2$, Corollary 1 predicts LIF gain $|\widehat{h}(1/\Delta)| \approx 0.62$ at a frequency the frame accumulator rejects exactly. Sweeping $\tau/\Delta \in \{0.1, 0.2, 0.35, 0.5, 0.75, 1\}$ at sub-saturating budgets yields Pearson $r = 0.99$ between predicted gain and measured accuracy drop. This dependence vanishes only after the attack saturates the classifier, which is why the mechanism sweep uses smaller budgets.

A Poisson testbed with $\lambda_0 = 2000\,\mathrm{s}^{-1}$ and $\Delta = 10$ ms directly checks Theorem 2. With a null-frequency modulation of amplitude 0.5, a two-sample Kolmogorov–Smirnov test over 2×10^4 frames gives $p = 0.57$ and variance ratio 1.01. At a non-null frequency the same-amplitude modulation gives $p < 10^{-3}$ and variance ratio 1.83. The integrated rate per null window is unchanged to 2×10^{-11}. Eq. (8) matches the exact stealth–efficacy ratio to 2.0% at relative frequency error 10^{-2} and 0.2% at 10^{-3}. The corresponding ratios are approximately 3.9×10^3 and 3.9×10^5.

H.3 Randomized accumulation and schedule-aware detection

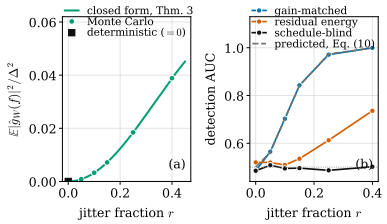


Figure 8: Randomized accumulation: expected response energy (left) and schedule-blind, residual-energy, and schedule-aware detection (right).

Figure 8 validates Theorem 3 and Eq. (10). Monte Carlo estimates of $\mathbb{E}|\widehat{g}_W(f)|^2$ agree with the closed form to within 0.25%. Over 400-frame streams at attack amplitude $\alpha = 0.4$, a schedule-blind detector remains at chance (0.49–0.50 AUROC) across the jitter sweep. The residual-energy detector reaches AUROC 0.61 and 0.74 at jitter $r = .25$ and .40, while the schedule-aware gain-matched detector reaches AUROC 0.97 and 0.999. The deflection-based prediction matches the gain-matched AUROC across all 18 evaluated settings with mean absolute error 0.011 (worst case 0.03).

H.4 Representation refinements and scope

The location of the null depends on the consumer's temporal representation. At the first null of a Δ-wide boxcar, splitting the interval into B equal sub-bins changes normalized carrier visibility to $|\mathrm{sinc}(1/B)|$. It is 0 for $B = 1$, about 0.64 for $B = 2$, and 0.98 for $B = 10$. Finer voxelization closes this blind spot but introduces higher-frequency nulls, where the LIF response is more attenuated. Eq. (8) quantifies the trade. An unsigned activity statistic $\int g\, d|N|$ need not share the same polarity-summed null. For randomized widths, the positivity claim requires a distribution that breaks null-preserving lattices, and detection requires access to the realized schedule.

I ALIGNED N-MNIST FIVE-MODEL PROTOCOL

To separate the null-space property from the older N-MNIST model choices, we reran N-MNIST with the same five-model family used in the larger event benchmarks, $T = 150$ fine bins and $S = 10$ bins per protected window. Clean test accuracy is approximately 99% for CoarseFrameFormer, ConvSNN, SEW-ResNet18, and Event Transformer and 98.5% for the GRU. Every optimized attack row has $D_A = D_\infty = 0$.

Table 16: Aligned N-MNIST null-space ASR (%, mean±std over three seeds).

Victim	10%	20%	30%
ConvSNN	4.11±0.32	10.00±0.51	18.71±1.67
SEW-ResNet18	1.45±0.51	3.50±0.24	4.81±0.32
Event Transformer	2.89±0.31	8.68±0.98	19.61±4.42
Temporal GRU	61.23±4.16	90.59±1.74	97.87±1.19

A max-shift ablation on seed 0 shows that the GRU already reaches 48.1% ASR when every moved event is restricted to one fine bin (2 ms), compared with 56.4% for unrestricted within-window moves at the same 10% budget. This supports a local-timing mechanism rather than a dependence on large timestamp displacements.

J DAILYDVS-200 ARCHITECTURE STRESS TEST

We also evaluate protocol-scale variants of ACTION-Net (Wang et al., 2021), MVFNet (Wu et al., 2021), Swin-T (Liu et al., 2021), and TimeSformer (Bertasius et al., 2021). These models are trained from scratch under our common DailyDVS preprocessing. They are *not* released benchmark checkpoints, and published benchmark scores are used only as architectural context. Table 17 reports the measured clean accuracy and clean-correct denominators needed to interpret conditional ASR. We use this result as a model-strength and architecture stress test under a controlled common representation. We do not claim that our training reproduces the published DailyDVS leaderboard. We report the low Swin-T clean accuracy explicitly. ASR remains conditioned on clean-correct clips for every model.

Table 17: DailyDVS protocol-scale architecture results at the 10% exact-null operating point. These are models trained under our common preprocessing, not official benchmark checkpoints. Clean Top-1 is measured on our exact evaluation split. n_{cc} lists clean-correct attacked clips for seeds 0/1/2. All optimized and matched-control rows preserve the protected tensor exactly and give bit-identical protected-consumer logits.

Victim	Clean Top-1	n_{cc} (s0/s1/s2)	Null ASR	matched control
ActionNet	42.18%	412/421/434	97.55±0.28	12.27±2.26
MVFNet	50.21%	506/522/507	99.28±0.11	9.70±1.06
Swin-T	22.74%	211/247/194	99.66±0.60	20.01±2.94
TimeSformer	37.63%	367/330/361	97.94±0.61	9.74±2.61

K SHD CROSS-MODALITY EXTENSION

SHD (Cramer et al., 2022) is a neuromorphic auditory benchmark and is used only to test whether the discrete null-space mechanism extends beyond event vision. We use the same $T = 80, S = 8$ construction, and all constrained rows satisfy exact protected equality.

The SHD result shows the same exact observer invisibility outside event vision, while vulnerability within the blind space remains architecture dependent.

Table 18: **Cross-modality extension to SHD.** Three-seed ASR (%, mean±std) under exact protected equality.

Victim	10% budget	20% budget
ConvSNN	53.9 ± 3.4	74.2 ± 3.7
SEW	12.9 ± 1.1	23.9 ± 1.2
Transformer	49.5 ± 4.3	67.1 ± 2.6
GRU	59.0 ± 2.6	73.3 ± 1.6

L UNIFIED FIVE-ATTACK AND OPERATIONAL-DEFENSE DETAILS

The unified DVS run evaluates Yu PIL-L_0, Free, Null, adapted SDA/PDSG, and adapted Yao/Gumbel on the same ConvSNN/CoarseFrameFormer population. Seed-0 headline values are reproduced below. The clean-only detector is fit on held-out clean clips and evaluated on the attack

Table 19: Unified DVS seed-0 attack/defense result.

Method	ASR	mean D_A	protected flip	AUROC	UASR@5
Null	74.90	0.000	0.0	0.493	73.64
SDA	40.59	0.010	0.0	0.494	40.59
Free	94.56	0.194	3.77	0.509	92.05
Yu PIL-L_0	100.00	0.198	2.51	0.432	98.74
Yao	64.02	0.523	34.73	0.932	35.56

population. The realized clean FPR for the combined detector is 2.51% for the main DVS rows at the nominal 5% operating point (Yao: 3.35% under its fold-level calibration). We report the realized FPR together with nominal operating points in the saved tables.

Across the completed three-seed DVS evaluation, Yu/SDA/Yao/Free/Null have AUROC $0.431/0.495/0.929/0.510/0.492$ and UASR@5 $98.6/34.9/36.1/92.4/79.9\%$, respectively. These aggregate values are the operational numbers reported in Section 5.4. The table above is retained as the seed-0 view.

A controlled Free sweep holds the attack family and approximate 10% edit footprint fixed while changing only the maximum allowed shift. Mean D_A progresses $0.0418 \to 0.0772 \to 0.1422 \to 0.1759 \to 0.1889$ for max-shift $2, 4, 8, 16, 32$ fine bins. ASR progresses $37.7 \to 54.4 \to 64.9 \to 72.8 \to 74.1\%$. Detector true-positive rate (TPR) at fixed false-positive rate (FPR) does not increase monotonically. Representation distance therefore does not provide a universal measure of operational detectability in this sweep.

The available DailyDVS unified operational diagnostic contains 119 pooled clean-correct evaluations across three seeds, substantially fewer than the broad 1,000-stream protocol. It shows the same non-monotonic pattern (Null $D_A = 0$, Yu/Free around 0.2, Yao around 0.63). All but Yao remain near chance on DVS, and all methods are near chance on this smaller Daily reporting population. We treat this as an underpowered diagnostic and do not use it for a main cross-dataset detection claim.

M OBSERVER-FAMILY ABLATIONS

Table 20 summarizes the DVS observer-family experiments. The fixed-attack rows apply the same Null streams to different temporal views. The intersection rows re-optimize the attack while enforcing equality to additional phase observers.

Exponential moving-average (EMA) summaries are also non-identical under the fixed attacks, although their normalized L_1 changes are smaller. The GRU remains near saturation in the intersection experiment, so adding observers does not give a universal robustness guarantee. Randomized schedules expose about 97.5% of the saved canonical-null DVS attacks, while the schedule-blind visibility statistic remains zero; Appendix H gives the corresponding detector experiment.

We also verified zero fallback in the signed-displacement matcher, exact integer equality of the protected tensor, bitwise equality of protected-consumer logits, fixed evaluation populations, and clean-only detector calibration. These checks ensure that a change in the protected representation or evaluation population is not counted as stealth.

Table 20: DVS observer-family ablations. Top: mean relative representation change for the same canonical-null ConvSNN attacks ($n = 239$). Bottom: blind-space size and ASR when the attack is constrained to multiple phase observers.

Fixed attack, changed observer			
Observer	Mean relative change		
Canonical width-8, phase 0	0.0000		
Width-8, phase 4	0.1357		
Width-8, stride-4 overlap	0.1291		
4/8/16-bin multiscale	0.1435		
Attack constrained to observer intersections			
Observers	Blind dim. (fraction)	ConvSNN ASR	SEW ASR
Phase 0	1,146,880 (87.5%)	74.9%	38.0%
Phases 0+2	983,040 (75.0%)	61.9%	29.5%
Phases 0+2+4	819,200 (62.5%)	37.2%	20.7%

N CONSTRAINED PROJECTED GRADIENT DESCENT (C-PGD)

C-PGD searches the same unrestricted within-window exact-null feasible set as Null. Let X_0 be the clean grouped count tensor, with one row for each protected (window, p, y, x) group and one column for each fine temporal bin. The primary comparison uses one projected-gradient step per progressive budget stage ($s = 1$); $s = 2$ and $s = 4$ are optimization-strength ablations.

Algorithm 1 C-PGD inside the exact-null feasible set

Require: Clean grouped counts X_0, label y, budget stages $\mathcal{B}$, steps per stage s, nominal step $\eta = 1$

$Z \leftarrow X_0$
for $b \in \mathcal{B}$ **do**
  for $j = 1, \ldots, s$ **do**
    $g \leftarrow \nabla_Z \mathcal{L}(Z, y)$
    for each protected row r **do**
      $g_r \leftarrow g_r - \text{mean}(g_r)$; normalize nonzero g_r by $\|g_r\|_\infty$
    end for
    $\eta_{\text{eff}} \leftarrow \eta$; increase it as needed, for at most 20 expansions
    $Z \leftarrow \Pi_{\text{row}}(Z + \eta_{\text{eff}} g)$, projecting each row onto $\{z \geq 0 : \sum_j z_j = \sum_j X_{0j}\}$
    if $\frac{1}{2}\|Z - X_0\|_1 > b$ **then**
      Contract Z toward X_0 to the budget boundary
    end if
  end for
end for
Integerize row deficits and surpluses by deterministic largest-remainder transport
Materialize legal within-window source-to-destination timestamp shifts
return integer event stream X' with $A(X') = A(X_0)$

The row-wise mean subtraction removes the component that would change a protected row sum. Simplex projection preserves each row's total mass, and the final transport step converts the relaxed tensor to integer event counts while keeping the requested moved-event budget. Ties use stable index order. The implementation checks nonnegativity, total event count, moved-event budget, and exact coarse equality after integerization.

At $s = 1$, Null and C-PGD are matched by victim-gradient evaluations: three through 10% on DVS Gesture and four on DailyDVS-200. Wall-clock time is not matched because the optimizers perform different non-gradient operations. The comparison tests whether a second optimizer can find damaging retimings in the same exact blind space.