EVENT ACTIONS PREDICT LOSS JUMPS AT SPIKE DELETION BOUNDARIES

Anonymous authors
Paper under double-blind review

ABSTRACT

Spiking neural networks combine continuous dynamics with discrete events triggered by threshold crossings. A finite parameter change can remove a spike, switching the network to a different event sequence and producing a loss change that is not represented by the gradient evaluated before the spike disappears.

We show that this loss jump can be predicted from the EventProp backward pass while the spike is still present. Deleting a spike removes both its membrane reset and its downstream synaptic transmission. We sum the corresponding backward terms and call the result the event action. Across 120 deletions of single spikes in a learned LIF network trained on the Spiking Heidelberg Digits (SHD) dataset, the event action computed before deletion closely matches the signed loss jump. The result replicates in an independently trained network and on examples from the speaker held out from training. Replacing the target spike with other spikes reduces prediction accuracy, and spike time, boundary location, spike position, and source neuron activity do not explain the relation.

The prediction magnitude also separates deletions with small and large effects: the highest predicted impact quintile has a median realized loss change about 803 times larger than the lowest. For 20 parameter directions that cross a deletion boundary, adding the predicted jump raises the rank correlation with the realized loss change from 0.281 to 0.997 and selects the candidate with the lowest loss. The backward pass contains information about the loss change caused by losing an event that the parameter gradient before deletion does not capture.

1 INTRODUCTION

Changing a synaptic weight generally shifts the timing of existing spikes. At certain parameter values, a spike can disappear altogether. The network then follows a different event sequence, and the loss can undergo a finite jump rather than continuing along the same branch. This matters during learning because parameter updates are finite: an update chosen from local gradient information can cross such a boundary and encounter a loss change that was not represented on the trajectory where that gradient was computed. Understanding these boundary effects is important for evaluating and eventually controlling parameter updates in spiking neural networks.

EventProp computes exact gradients in LIF networks simulated in continuous time by propagating adjoint variables backward through recorded spike events (Wunderlich & Pehle, 2021). These gradients are exact while the current event sequence remains unchanged. Once a spike appears or disappears, however, the trajectory switches to a different event sequence and the parameter gradient before deletion does not contain the resulting finite loss jump. Existing work addresses this difficulty by modifying either the loss (Nowotny et al., 2025) or the spiking dynamics (Klos & Memmesheimer, 2025). Here the original loss and dynamics are left unchanged, and the loss jump itself is estimated before deletion occurs.

A spike has two immediate effects on the forward trajectory: it resets the source membrane and transmits synaptic input downstream. Both effects disappear when the spike is deleted. We sum the corresponding EventProp backward terms and call the result the *event action*. The event action is defined for a specific spike rather than for a parameter direction.

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.

The EventProp parameter gradient and the deletion jump describe different changes. The backward pass is evaluated on the trajectory in which spike e still exists, whereas the loss jump is measured after that spike disappears and the network follows a different trajectory. Later spike times may also shift after deletion. Comparing a_e with the measured jump tests whether the backward information associated with the disappearing spike remains predictive after the event sequence changes.

Prior work has interpreted adjoint variables evaluated at spike times as measures of responsibility or "blame" assigned to spikes and synaptic weights (Mészáros et al., 2025). In contrast, our analysis uses the backward terms associated with an individual spike to predict the signed finite change in loss that occurs when that spike is deleted.

Experiments on learned LIF networks trained on the 20-class Spiking Heidelberg Digits (SHD) dataset show close agreement between the event action prediction and the measured loss jump for deletions induced by continuous weight scaling. The main analysis contains 120 cases in which exactly one spike disappears at the first boundary. The relation replicates in an independently trained network and on examples from the speaker held out from training. The prediction magnitude tracks the size of the measured jump. Replacing the target spike with other spikes reduces accuracy, while spike time, boundary location, source neuron activity, and position within the source spike train do not explain the relation. The predicted jump is also large enough to change the ordering of finite parameter updates.

The main contributions are:

- We introduce the event action, the sum of the EventProp backward terms associated with a spike's reset and downstream transmission, and show that its value before deletion predicts the signed loss jump produced when that spike disappears.
- We show the relation across 120 deletions and reproduce it in an independently trained network and on examples from the speaker held out from training. The target spike gives the most accurate prediction, and the relation remains accurate when later source neuron spike times shift.
- For finite parameter changes, the deletion jump often exceeds the loss change accumulated before deletion. Adding the predicted jump improves the ordering of candidate changes by their realized loss.

2 Related Work

Spike creation and deletion in SNN learning. Nowotny et al. (2025) address spike creation and deletion through loss shaping, reducing the disruption caused by event sequence changes during training. Klos & Memmesheimer (2025) instead modify the neuron dynamics so that spikes can appear and disappear continuously. Both approaches alter behavior near event sequence boundaries. Our experiments keep the original loss and dynamics and estimate the finite loss jump associated with a particular deletion.

Adjoint information at spike times. Mészáros et al. (2025) interpret EventProp adjoint variables at spike times as assigning responsibility or "blame" to spikes and the synaptic weights through which they act. Those adjoints are used to compute parameter gradients on the current event sequence. The event action studied here is the sum of the backward terms corresponding to a spike's reset and downstream transmission, and is compared with the finite loss jump measured after that spike disappears. Klawitter & Michmizos (2026) instead use spike removal for attribution, identifying presynaptic spikes necessary for a postsynaptic neuron to reach threshold. Their measure concerns causal influence on downstream firing rather than the change in task loss caused by a deletion induced by changing a parameter.

Estimating removal effects. Taylor pruning methods use derivatives to estimate the loss effect of removing parameters, filters, or neurons from otherwise smooth networks (Molchanov et al., 2019). A spike deletion differs because the spike is generated by the network dynamics itself. Removing it eliminates both the source neuron reset and downstream synaptic transmission and can also alter the timing of subsequent events.

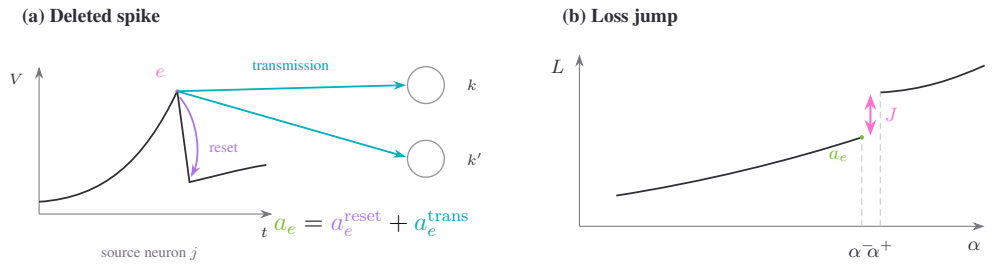


Figure 1: **Conceptual view of spike deletion and the associated loss jump. (a)** A hidden spike e from source neuron j resets the source membrane and transmits activity to downstream neurons k and k'. The backward terms form the event action a_e. **(b)** The event action is evaluated immediately before the spike disappears and is compared with the finite loss jump $J = L(\alpha^+) - L(\alpha^-)$ after the event sequence changes.

3 METHOD

Figure 1 summarizes the analysis. For each candidate hidden neuron, we scale its incoming weights along a path controlled by α, locate the first boundary at which a spike disappears, measure the resulting loss jump, and compare it with the event action prediction computed for that spike immediately before deletion.

3.1 MEASURED DELETION JUMP

For a hidden neuron j, the incoming weights are scaled as

$$W_{\cdot j}(\alpha) = (1 - \alpha) W_{\cdot j}(0). \tag{1}$$

Starting from $\alpha = 0$, we increase α until the hidden spike pattern first changes and refine that boundary by bisection.

Let α^- denote a point just before the boundary, where the target spike still exists, and α^+ a point just after it, where the spike has disappeared. The measured deletion jump is

$$J = L(\alpha^+) - L(\alpha^-). \tag{2}$$

The predictor is evaluated only at α^-. We retain a case when exactly one entry of the spike count matrix, indexed by example and hidden neuron, decreases by one at the first boundary and network activity continues afterward. Neither J nor the prediction error is used to select cases.

3.2 EVENT ACTION PREDICTOR

Deleting a spike removes two immediate effects from the forward trajectory: the source neuron reset and the downstream synaptic transmission. For a hidden spike e emitted by neuron j at time t_e, the corresponding EventProp backward terms are

$$a_e^{\text{reset}} = \tau_m(\vartheta - V_r)\lambda^+_{V,j}(t_e), \tag{3}$$

$$a_e^{\text{trans}} = -\sum_k w_{jk}\tau_{s,k}\lambda_{I,k}(t_e). \tag{4}$$

We use the term event action for their sum,

$$a_e = a_e^{\text{reset}} + a_e^{\text{trans}}. \tag{5}$$

Both terms are included because both forward effects disappear when the spike is deleted.

If spike e were suppressed while the state immediately before t_e remained unchanged, the missing state change would consist of the removed reset and transmission jumps. Under the EventProp sign conventions, contracting those changes with the adjoint gives the two terms in Eq. 5. The adjoint at t_e also reflects what happens later in the trajectory. As long as the remaining event sequence is

preserved, it propagates the effect of a state perturbation through later dynamics and shifts in event times. Changes in later spike times can be represented without changing the event sequence.

The actual deletion crosses to a different event sequence. We change the synaptic weights until the target spike no longer occurs, then compare the event action at α^- with the loss jump measured at α^+. EventProp exactness on the sequence before deletion does not determine this jump. Additional spike creation or deletion after the target event would introduce further discrete changes beyond the action of that event.

The implementation converts the recorded action to the same units of mean loss as J and defines the prediction $\widehat{J}$. The conversion is given in Appendix A.2; no coefficient is fitted to deletion outcomes.

To study the magnitude of the deletion effect, we use $|\widehat{J}|$. We rank the 120 deletions by $|\widehat{J}|$, divide them into five groups of 24, and report the realized $|J|$ values in each group.

3.3 Finite update decomposition

To measure how the deletion jump affects finite parameter changes, we split the realized loss change at the deletion boundary:

$$S = L(\alpha^-) - L(0), \tag{6}$$

$$J = L(\alpha^+) - L(\alpha^-), \tag{7}$$

$$T = L(\alpha^+) - L(0) = S + J. \tag{8}$$

Here S is the loss change accumulated before deletion, J is the jump at the boundary, and T is the final loss change after the spike has disappeared. We compare the ordering given by S with the ordering given by $S + \widehat{J}$ to measure how much the predicted jump changes the comparison between candidate parameter updates.

4 Experiments

The main analysis compares $\widehat{J}$ with J for 120 deletions. Two replications use 20 deletions each: one with an independently trained network and one with examples from the speaker held out from training. We then compare the target spike with other spikes, evaluate the reset and transmission terms separately, and test spike time, boundary location, spike position, and source neuron activity. The finite update experiment uses 20 directions.

4.1 Main deletion experiment

The experiments use the 20-class Spiking Heidelberg Digits (SHD) dataset (Cramer et al., 2022). The main network is a feedforward model with one hidden layer of 256 LIF neurons. It is trained for 100 epochs using the EventProp implementation and `sum_weigh_exp` objective described by Nowotny et al. (2025). The final epoch 99 checkpoint is analyzed. Training accuracy is 89.2%, and accuracy on the speaker held out from training is 63.0%. Appendix A.1 gives the remaining training settings.

The deletion search uses 64 examples spanning all 20 classes and tests 128 directions in the incoming weights. The directions are tested in the same sequence in every run, given in Appendix A.3. For each direction, we increase α until the hidden spike count pattern first changes and refine that boundary by bisection. We keep the direction if exactly one spike disappears at that boundary and network activity continues afterward. Of the 128 directions, 120 meet this criterion.

For each of these 120 deletions, $\widehat{J}$ is computed immediately before the boundary and J is measured across it. All 120 are included in the reported results.

4.2 Replications

The first replication changes the trained parameters while keeping the task and architecture unchanged. A second network with the same architecture and training procedure is trained independently. Its final checkpoint reaches 88.9% training accuracy and 66.4% accuracy on the held out

speaker. We use the same direction sequence and deletion criterion and evaluate the first 20 deletions that meet the criterion.

The second replication keeps the trained network unchanged and replaces the 64 analysis examples with examples from the speaker held out from training. We use the same class counts, direction sequence, boundary search, deletion criterion, and event action calculation, and evaluate the first 20 deletions that meet the criterion.

4.3 CONTROLS

All 120 deletions in the main network are used here. For each deletion, we compare the target event action with three replacements: the mean action of the other spikes from the same source neuron in the same example, the mean action of that neuron's spikes across the full 64 example batch with the target event excluded, and the action of the closest spike in time from another hidden neuron. We also evaluate the transmission and reset terms separately.

We compare $|J|$ with the target spike time, the boundary value α^-, the target spike's relative position within the source neuron's spike train, and the source neuron spike count.

The trajectory, derivative, and finite update analyses use the first 20 directions that meet the deletion criterion. We use the same 20 directions for all three analyses. Appendix A.3 gives the direction sequence and selection rule.

5 RESULTS

5.1 THE EVENT ACTION PREDICTS THE DELETION JUMP

Across the 120 deletions in the main network, $\widehat{J}$, computed immediately before deletion, closely matches the measured loss jump J (Fig. 2(a)). Pearson correlation is $r = 0.99995$, Spearman correlation is $\rho = 0.981$, median relative error is 3.92%, and the slope of a fit constrained through the origin is 0.9933, with bootstrap 95% interval $[0.9845, 0.9983]$. No coefficient is fitted to deletion outcomes.

The same cases show that prediction magnitude tracks deletion impact. Across all 120 deletions, $\rho(|\widehat{J}|, |J|) = 0.972$. Dividing cases into five equal quintiles by $|\widehat{J}|$ gives median realized magnitudes 4.47×10^{-8}, 1.05×10^{-7}, 1.10×10^{-6}, 5.53×10^{-6}, and 3.59×10^{-5} from the lowest to highest quintile. The highest median is about $803\times$ the lowest (Fig. 2(b)).

Sign agreement is 108/120. All 12 disagreements have $|J| < 10^{-7}$ and remain in the reported statistics. For cases with $|J| \geq 10^{-7}$, sign agreement is 85/85 and median relative error is 1.80%; for cases with $|J| \geq 10^{-6}$, sign agreement is 61/61 and median relative error is 1.43%.

5.2 THE RELATION REPLICATES ACROSS TRAINED PARAMETERS AND UNSEEN DATA

The same relation appears with different trained parameters and different examples. For the independently trained network, the first 20 deletions that meet the criterion give Pearson correlation $r = 0.99997$, Spearman correlation $\rho = 0.968$, median relative error 2.41%, and 19/20 correct signs. The single sign disagreement has $|J| < 10^{-7}$.

For the examples from the speaker held out from training, the first 20 deletions that meet the criterion give Pearson correlation $r = 0.99996$, Spearman correlation $\rho = 1.000$, median relative error 1.53%, and 20/20 correct signs. Table 1 summarizes the three evaluations.

5.3 THE TARGET SPIKE GIVES THE MOST ACCURATE PREDICTION

For the comparisons with other spikes, we repeat the boundary refinement for all 120 directions. The target event action gives 3.67% median relative error and 106/120 correct signs in this run. The small difference from the main results comes from repeating the boundary refinement; jumps extremely close to zero can change sign between the two runs (Appendix C.1).

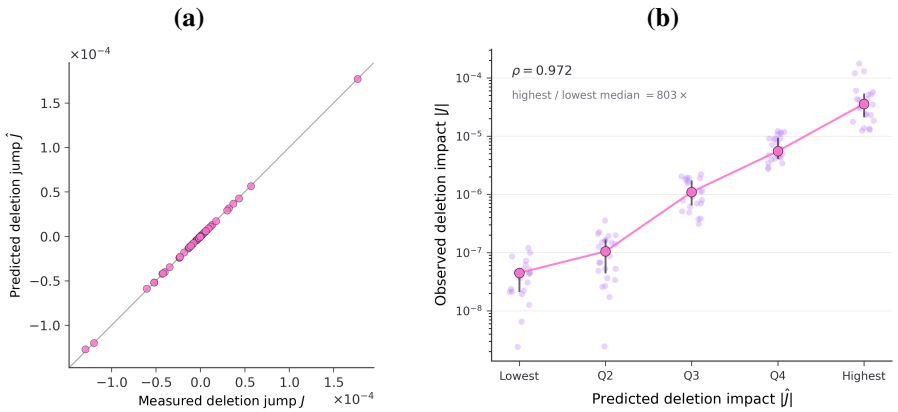


Figure 2: **The event action computed before deletion predicts both the signed loss jump and its magnitude.** **(a)** Measured J and $\widehat{J}$ for all 120 deletions in the main network. The diagonal is the identity relation; $r = 0.99995$ and the slope of a fit constrained through the origin is 0.9933. **(b)** The same 120 deletions divided into five equal groups by $|\widehat{J}|$. Small points are individual realized $|J|$ values, large markers are medians, and vertical segments show the interquartile range. The highest group has a median realized $|J|$ approximately $803\times$ larger than the lowest, and $\rho(|\widehat{J}|, |J|) = 0.972$.

Table 1: **Deletion jump prediction across the main evaluations.** Cases are selected by requiring exactly one spike to disappear at the first boundary.

Evaluation	N	Pearson r	Median rel. error	Sign
Main network	120	0.99995	3.92%	108/120
Independent network	20	0.99997	2.41%	19/20
Held out speaker	20	0.99996	1.53%	20/20

Replacing the target action with the mean over the other spikes from the same source neuron and example keeps a high correlation ($r = 0.994$), but median relative error rises to 11.91%. Both predictors have 106/120 correct signs because the disagreements occur for very small jumps. Averaging the same neuron's other events across the full 64 example batch gives $r = 0.129$ and 467.86% median relative error. The closest event in time from another hidden neuron gives $r = -0.126$ and 150.96% median relative error. Table 2 reports the full comparison.

Across all 120 deletions in the main network, spike time, boundary location, relative spike position, and source neuron spike count have absolute Pearson correlations of at most 0.141 with $|J|$. The event action magnitude gives $r = 0.99995$ (Table 3).

5.4 The transmission term dominates in this network

Across all 120 deletions, the full action gives $r = 0.99995$ and 3.92% median relative error. Transmission alone gives $r = 0.99999$ and 3.60% median relative error, whereas reset alone gives $r = -0.664$ and 100.85% median relative error (Table 4).

The event action includes both the reset and transmission terms because spike deletion removes both forward effects. In this feedforward network, transmission alone gives nearly the same predictive accuracy as the full action, while reset alone performs poorly.

5.5 Later source neuron spike times shift after some deletions

Using the recordings immediately before and after the boundary, we match the deleted spike to the originally identified event in 18 of 20 cases. Ten of those 18 show changes in later source neuron spike times, while none shows an additional change in the number of remaining source spikes. The largest timing shift is 7 ms.

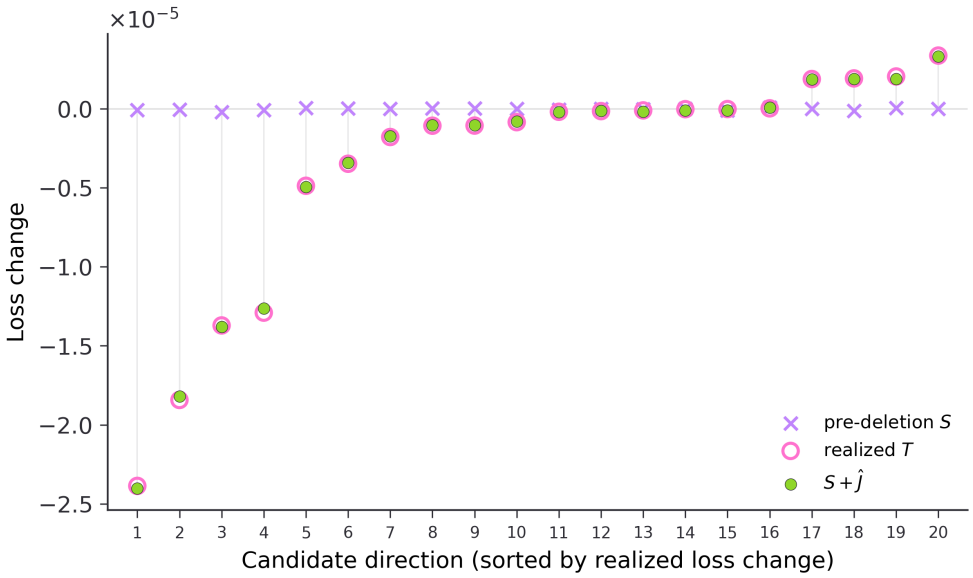


Figure 3: **The predicted deletion contribution changes the ordering of finite parameter updates.** Candidates are sorted by their realized total loss change T after crossing the deletion boundary. The measured change before deletion gives $\rho(S, T) = 0.281$, while adding the predicted jump gives $\rho(S + \widehat{J}, T) = 0.997$. The jump itself exceeds the change before deletion, $|J| > |S|$, in 18/20 cases.

The prediction remains accurate in the cases with later timing shifts. This analysis records the source neuron only, not the complete event sequence of the network after deletion.

5.6 The deletion jump changes finite update ordering

For each of the 20 directions, $L(0)$, $L(\alpha^-)$, and $L(\alpha^+)$ are measured in the same execution. The deletion jump is larger than the entire loss change before deletion in 18/20 cases, with median $|J|/|S| = 73.4$.

Ranking candidates by the measured change before deletion S alone gives $\rho(S, T) = 0.281$ with the realized total change T. Adding the predicted jump gives $\rho(S + \widehat{J}, T) = 0.997$. Minimizing S selects a candidate whose realized loss ranks third, whereas minimizing $S + \widehat{J}$ selects the candidate with the lowest loss among the 20 tested directions (Fig. 3). Because S is measured directly rather than approximated by a gradient, this comparison isolates the information added by the deletion prediction.

6 Discussion and Limitations

EventProp gives exact parameter gradients while the event sequence is unchanged. At a spike deletion boundary, the target event vanishes and the network follows a different trajectory. The event action prediction computed just before this boundary matches the finite loss jump measured just after it.

Actions from other spikes of the same source neuron and example remain strongly correlated with the target jump, but their mean has about three times the median relative error of the target action. Averaging across the full batch removes most of the relation, and a nearby spike from another neuron performs poorly. The target event action gives the lowest error.

The event action contains reset and transmission terms because both effects disappear when a spike is deleted. In this network, transmission alone nearly matches the full prediction, while the reset term alone performs poorly. Downstream transmission dominates the measured loss jump.

In 10 of the 18 cases where the deleted event can be matched directly, later spikes from the same source neuron shift in time. Accurate prediction does not require those later spike times to remain unchanged. EventProp's adjoint propagates state perturbations through subsequent dynamics and shifts in later event times. Additional spike creations or deletions would add discontinuities beyond the action of the deleted event.

For the 20 finite update directions, the measured change before deletion S has $\rho(S, T) = 0.281$ with the realized total change. The deletion jump exceeds $|S|$ in 18/20 cases. Adding the predicted jump raises the rank correlation to $\rho(S + \widehat{J}, T) = 0.997$ and selects the candidate with the lowest loss among the 20 directions. The deletion jump can dominate the comparison between finite updates.

These experiments cover deletions of single spikes induced by parameter changes in a shallow feed-forward LIF network. Multiple events can change at the same boundary, and in recurrent networks one deletion can trigger further changes. Spike creation is different because the new event is absent before the boundary, leaving no event action to evaluate for it.

7 Conclusion

We showed that the EventProp terms associated with a spike's reset and downstream transmission can be used before deletion to predict the loss jump that occurs when that spike disappears. Across 120 deletions, the prediction closely matches the measured jump. The same relation is reproduced in an independently trained network and on examples from the speaker held out from training.

The prediction is most accurate for the spike that actually disappears, and its magnitude tracks the size of the deletion effect. In this network, the transmission term accounts for most of the prediction accuracy. The prediction also remains accurate when later spikes from the source neuron shift in time. For the 20 finite update directions, adding the predicted deletion jump improves the ordering of the realized loss changes.

AI Use Statement

In this work, we used generative AI tools to assist with research ideation, literature search planning, code drafting and debugging, statistical analysis scripts, and manuscript drafting and editing. All scientific claims were checked against primary sources or experimental results, and all AI assisted code used for reported experiments was inspected and validated by the author. We take responsibility for the final content of this work, including text, claims, code, and artifacts produced with the aid of generative AI.

Reproducibility Statement

The appendix reports the training settings, data splits, direction sequence, deletion criterion, boundary search procedure, normalization, replication settings, comparisons with other spikes, reset and transmission results, trajectory analysis, derivative check, and finite update analysis. We plan to release the experiment code and result files after the review process.

References

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.

Jack Klawitter and Konstantinos P. Michmizos. Critical spike attribution for feature importance in spiking neural networks. In *2026 IEEE Neuro-Inspired Computational Elements (NICE)*. IEEE, 2026. doi: 10.1109/NICE69539.2026.11567461.

Christian Klos and Raoul-Martin Memmesheimer. Smooth exact gradient descent learning in spiking neural networks. *Physical Review Letters*, 134:027301, 2025. doi: 10.1103/PhysRevLett.134.027301.

Balázs Mészáros, James C. Knight, and Thomas Nowotny. Efficient event-based delay learning in spiking neural networks. *Nature Communications*, 16:10422, 2025. doi: 10.1038/s41467-025-65394-8.

Pavlo Molchanov, Arun Mallya, Stephen Tyree, Iuri Frosio, and Jan Kautz. Importance estimation for neural network pruning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 11264–11272, 2019.

Thomas Nowotny, James P. Turner, and James C. Knight. Loss shaping enhances exact gradient learning with eventprop in spiking neural networks. *Neuromorphic Computing and Engineering*, 5(1):014001, 2025. doi: 10.1088/2634-4386/ada852.

Timo C. Wunderlich and Christian Pehle. Event-based backpropagation can compute exact gradients for spiking neural networks. *Scientific Reports*, 11:12829, 2021. doi: 10.1038/s41598-021-91786-z.

A Experimental Details

A.1 Training and evaluation sets

Both trained networks use one hidden layer with 256 LIF neurons and no recurrence. The full speaker split contains 7,243 training examples and 913 examples from the speaker held out from training before balancing. Class reduction is disabled, class balancing is enabled for both training and evaluation, and spike number normalization is enabled.

Training uses the `sum_weigh_exp` loss, 100 epochs, learning rate 0.01, input dropout 0.2, and initialization standard deviation 0.37 for weights from the hidden layer to the output. We analyze the epoch 99 checkpoint specified by the training protocol.

For the main network, epoch 99 gives training accuracy 0.892, accuracy on the speaker held out from training 0.630, training loss 0.297, and loss on that speaker 1.944. For the independent network, the corresponding values are 0.889, 0.664, 0.304, and 1.624.

The main analysis uses 64 examples from the training split: four examples each for classes 0–3 and three examples each for classes 4–19, in split order. The speaker replication uses the same class counts and draws all 64 examples from the 913 examples belonging to the speaker held out from training. Dropout is disabled during all deletion analyses.

A.2 Event action normalization

In the EventProp backward equations,

$$a_e = \tau_m(\vartheta - V_r)\lambda^+_{V,j}(t_e) - \sum_k w_{jk}\tau_{s,k}\lambda_{I,k}(t_e). \tag{9}$$

The implementation stores the raw event action value r_e. For the 64 examples used in the deletion search, the reported prediction is

$$\nu_e = \tau_m B r_e, \qquad \widehat{J}_e = -\frac{\nu_e}{N}, \tag{10}$$

where B is the simulator batch size and N is the number of examples entering the reported mean loss. With $B = 32$, $N = 64$, and $\tau_m = 20$ ms,

$$\widehat{J}_e = -10 r_e. \tag{11}$$

No factor in this conversion is fitted to deletion outcomes.

The raw value contains both the membrane reset term and the transmission term from the hidden layer to the output. The measurement synapse is not used in the forward simulation; its weights are synchronized with the trained output weights so that the recorded action contains the actual downstream transmission contribution.

A.3 Deletion search and main analysis

We test hidden neurons one by one using the same sequence in every run:

$$n_i = (37i + 17) \bmod 255,$$

with neuron 255 excluded because it was used for preliminary verification. We evaluate the first 128 directions in this sequence. In 120 of them, the first boundary removes exactly one spike and later activity continues; in seven, the deleted spike is the final event; in one, more than one spike count entry changes at the first boundary.

For each direction, the incoming weights of the selected hidden neuron are scaled according to Eq. 1. We first locate the boundary coarsely and then refine it. We retain the boundary only when exactly one entry of the spike count matrix changes: the count for the selected example and hidden neuron decreases by one.

Across the 120 deletions, Pearson correlation is $r = 0.99995$, Spearman correlation is $\rho = 0.981$, sign agreement is 108/120, and median relative error is 3.92%. The median absolute error is 5.49×10^{-8}. The slope of a fit constrained through the origin is 0.9933, with bootstrap 95% interval $[0.9845, 0.9983]$.

All 12 sign disagreements have $|J| < 10^{-7}$. For cases with $|J| \geq 10^{-7}$ ($N = 85$), sign agreement is 85/85 and median relative error is 1.80%. For cases with $|J| \geq 10^{-6}$ ($N = 61$), sign agreement is 61/61 and median relative error is 1.43%. All 120 deletions are included in these statistics.

Prediction magnitude. Sorting the 120 deletions by $|\widehat{J}|$ and dividing them into five groups of 24 gives median realized $|J|$ values of 4.47×10^{-8}, 1.05×10^{-7}, 1.10×10^{-6}, 5.53×10^{-6}, and 3.59×10^{-5} from the lowest to highest group. The ratio of the highest to the lowest median is 802.7, and the Spearman correlation between $|\widehat{J}|$ and $|J|$ across all 120 cases is $\rho = 0.972$.

Analyses using 20 directions. The trajectory, derivative, and finite update results below use the first 20 directions that meet the deletion criterion. Their final loss jumps were measured after the direction sequence and selection rule were set. The comparisons with other spikes and the component results use all 120 deletions.

B Replication Analyses

B.1 Independent network

The independent network is trained with the same architecture and settings, and its final epoch 99 checkpoint is analyzed. We test hidden neurons in the same sequence as in the main analysis and use the first 20 directions whose first boundary removes exactly one spike and is followed by later activity.

Across these 20 cases, Pearson correlation between $\widehat{J}$ and J is $r = 0.99997$, Spearman correlation is $\rho = 0.968$, sign agreement is 19/20, and median relative error is 2.41%. The sole sign disagreement has $|J| < 10^{-7}$.

B.2 Speaker held out from training

The held out replication uses the main network but replaces the 64 training examples with a balanced set of 64 examples from the speaker held out from training. The class counts, hidden neuron sequence, deletion criterion, boundary search, event action calculation, and loss jump measurement procedure are unchanged.

Among the same 128 directions, 118 first boundaries remove exactly one spike and are followed by later activity, nine delete the final spike of the trajectory, and one changes more than one spike count entry. We use the first 20 directions that meet the criterion. All 20 final measurements confirm that exactly the intended spike disappears.

Table 2: **Comparisons with other spikes across 120 deletions.** The mean over other spikes from the same neuron and example remains highly correlated with the target jump but has larger error. Averaging across the full batch or using a spike from another neuron performs poorly.

Predictor	Pearson r	Spearman ρ	Sign	Median rel. error
Target event action	0.99995	0.985	106/120	3.67%
Other spikes, same neuron and example	0.994	0.984	106/120	11.91%
Same neuron, all 64 examples	0.129	0.014	63/120	467.86%
Nearest event, other neuron	-0.126	-0.143	45/120	150.96%

Table 3: **Spike, boundary, and activity variables across the 120 main network deletions.** Correlations are with deletion jump magnitude $|J|$.

Variable	Pearson r	Spearman ρ
$\lvert\widehat{J}\rvert$	0.99995	0.972
Spike time	-0.122	-0.138
Boundary location α^-	0.141	0.295
Relative spike position	-0.094	-0.173
Source neuron spike count	-0.141	-0.158

Across these 20 cases, Pearson correlation between $\widehat{J}$ and J is $r = 0.99996$, Spearman correlation is $\rho = 1.000$, sign agreement is 20/20, and median relative error is 1.53%. The slope of a fit constrained through the origin is 0.9887.

For cases with $|J| \geq 10^{-7}$ ($N = 19$), sign agreement is 19/19 and median relative error is 1.46%. For cases with $|J| \geq 10^{-6}$ ($N = 15$), sign agreement is 15/15 and median relative error is 1.12%.

C Controls

C.1 Comparisons with other spikes

Repeating the boundary refinement for all 120 directions gives 106/120 correct signs and 3.67% median relative error for the target event action, compared with 108/120 and 3.92% in the main analysis. Jumps near zero can change sign between the two refinements. The mean over other spikes from the same neuron and example also gives 106/120 correct signs, but its median relative error is 11.91%.

The mean over other spikes from the same neuron and example excludes the disappearing spike itself. The average across all 64 examples also excludes that spike. The final comparison uses the closest event in time from a different hidden neuron in the same example.

C.2 Spike, boundary, and activity variables

Across the 120 deletions in the main network, we compare $|J|$ with the target spike time, the boundary value α^-, the target spike's position in the source neuron spike train, and the source neuron spike count. Relative position is the spike's index in that train, starting from 1, divided by the total number of spikes from the source neuron. These variables do not have a sign that should match J, so we compare them with $|J|$.

The same four variables also show weak association with signed J; across them, $|r| \leq 0.217$ and $|\rho| \leq 0.219$. For comparison, $\widehat{J}$ versus signed J gives $r = 0.99995$ and $\rho = 0.981$.

C.3 Reset and transmission components

Using measurements from all 120 deletions, we evaluate the reset and transmission terms in Eq. 5 separately.

Table 4: **Event action components across 120 deletions.** The transmission term accounts for nearly all of the prediction accuracy.

Predictor	Pearson r	Spearman ρ	Sign	Median rel. error
Full event action	0.99995	0.981	108/120	3.92%
Transmission only	0.99999	0.981	108/120	3.60%
Reset only	-0.664	-0.969	12/120	100.85%

Table 5: **Finite update ranking.** The predicted jump restores the ordering of the realized total loss changes.

Score	Spearman with T	Selected rank	Regret
S	0.281	3	1.02×10^{-5}
$\widehat{J}$	0.988	1	0
$S+\widehat{J}$	0.997	1	0

The full event action contains both terms because deleting a spike removes both forward effects. Transmission alone gives nearly the same predictive accuracy as the full action.

C.4 Source neuron trajectory

For the first 20 directions that meet the deletion criterion, we repeat the boundary search. Later source neuron spike times change in 11 of 20 cases, and the number of remaining source spikes does not change further. We then record the source neuron spike sequence from the same α^- and α^+ simulations used to measure the loss jump. The deleted spike can be matched to the originally identified event in 18 of 20 cases using the 0.51 ms tolerance. Ten of those 18 show changes in later spike times, eight do not, and none shows another change in the number of remaining source spikes. The largest timing shift is 7 ms.

The two unmatched cases are omitted only from the comparison of individual spike times. They remain in the loss jump results.

C.5 Derivative with respect to outgoing weights

For hidden neurons 17, 54, 91, and 128, we scale all outgoing weights of one hidden neuron together and compare the EventProp directional derivative with finite differences at perturbation sizes 10^{-4}, 3×10^{-4}, and 10^{-3}. These perturbations preserve the event sequence. Before normalization, the finite difference to EventProp ratio is approximately $\tau_m = 20$, as required by the implementation convention. After dividing by τ_m, the 12 ratios have median 0.9993, mean 1.0023, and range $[0.9892, 1.0281]$. This verifies the derivative with respect to outgoing weights used by the implementation.

D Finite Update Analysis

For the same 20 directions, we record $L(0)$, $L(\alpha^-)$, and $L(\alpha^+)$ together. All 20 cases satisfy

$$S + J = T$$

to stored numerical precision.

Across these 20 directions, $\widehat{J}$ and J have Pearson correlation $r = 0.99990$, Spearman correlation $\rho = 0.991$, 20/20 sign agreement, and median relative error 2.69%. The median ratio $|J|/|S|$ is 73.4, and $|J| > |S|$ in 18/20 cases. Table 5 summarizes the ranking quantities.