Poisson SPKD Sigmoidal Model

Data

Victor–Purpura Distance (SPKD)

The **Victor–Purpura distance (SPKD)** quantifies the dissimilarity between two spike trains by measuring the minimum cost required to transform one train into the other. This transformation involves three operations:

1. **Delete a spike** (Cost $= 1$): remove a spike from one train.
2. **Add a spike** (Cost $= 1$): insert a spike to align with another.
3. **Shift a spike** (Cost $= q \cdot \Delta t$): move a spike in time, where cost scales with the time difference and a user-defined cost parameter q.

Example

Given:

- Spike Train 1 ($t_{\ell i}$): `[1, 2.5, 3.5, 6, 9]`
- Spike Train 2 ($t_{\ell j}$): `[1.5, 2, 3.7, 4, 8, 10]`
- Cost parameter: $q = 0.5$

Operations:

- Shift 1 to 1.5: $0.5 \times |1 - 1.5| = 0.25$
- Shift 2.5 to 2: $0.5 \times |2.5 - 2| = 0.25$
- Shift 3.5 to 3.7: $0.5 \times |3.5 - 3.7| = 0.1$
- Delete spike 6: cost $= 1$
- Shift 9 to 8: $0.5 \times |9 - 8| = 0.5$
- Add spike 10: cost $= 1$

Total Cost: $0.25 + 0.25 + 0.1 + 0.5 + 1 + 1 = 3.1$

Generating the Data

To generate the SPKD data across cost values and firing rates, we used the following procedure:

1. For each firing rate λ in a predefined range (1 to 100 spikes/s), we generated N spike trains of fixed duration $T = 1$ s by sampling inter-spike intervals from a homogeneous Poisson process. Spike trains were created by binning time into small intervals Δt and drawing spike counts from a Bernoulli approximation via a Poisson sampler.

2. From each set of N spike trains, we computed the Victor–Purpura distance across all unique train pairs (i, j) using a fixed cost parameter q.

3. For each pair, we calculated the per-spike distance by dividing the total distance by the expected number of spikes in a train (i.e., λT).

4. We repeated this process for a range of cost values q (log-spaced between 10^{-3} and 10^{6}) and recorded the mean per-spike SPKD over all train pairs.

This procedure yielded a surface of SPKD vs. cost for various firing rates.

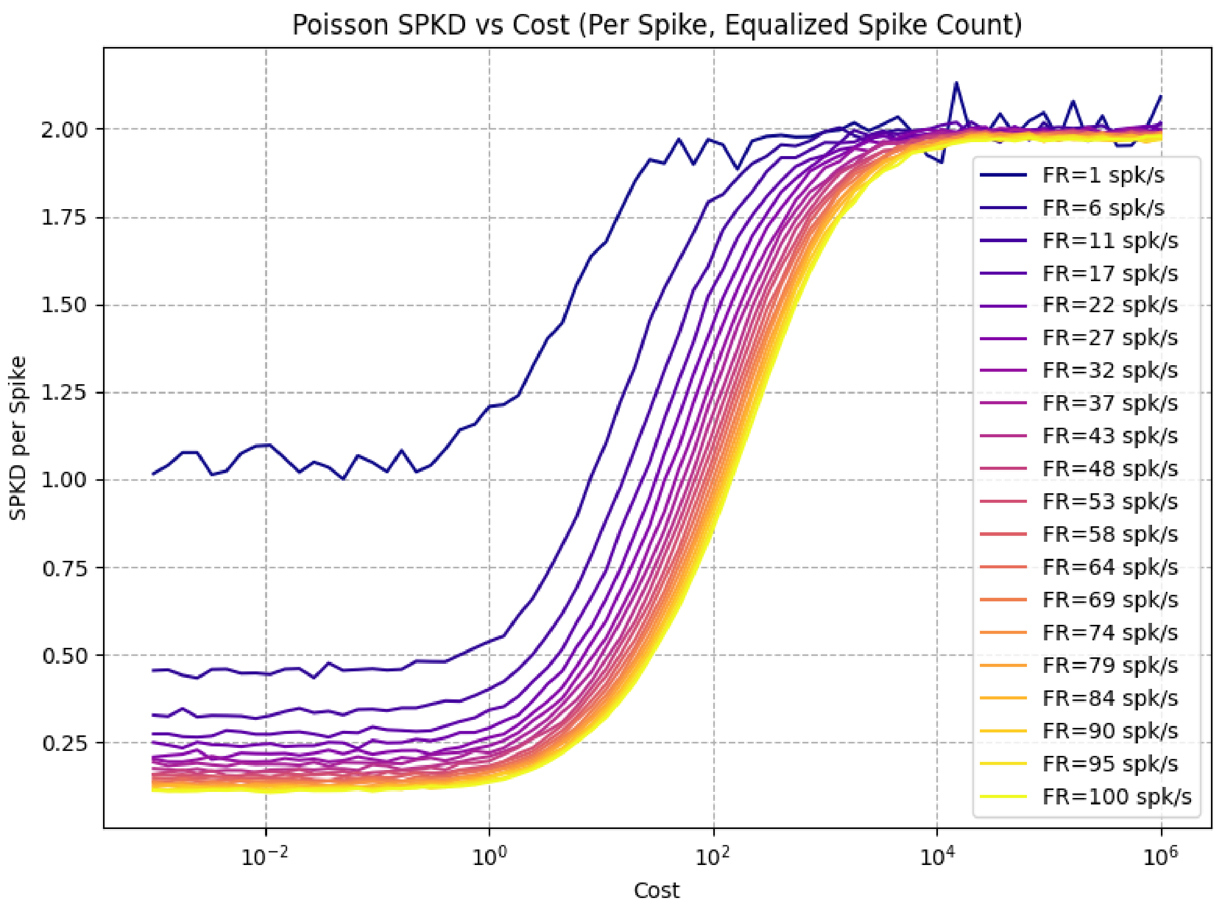

Figure 1: Surface of SPKD per spike as a function of cost and firing rate. Values were computed from simulated Poisson spike trains using the Victor–Purpura distance.

Model

We wanted a simple and flexible model that captures how SPKD per spike changes with cost and firing rate. When we looked at the data (Figure 1), the curves had a common shape: they start flat, rise quickly, and then level off. This kind of shape is well described by a sigmoid function, so we used one with four parameters.

Final Model Structure

Our final model uses analytic expressions for the lower and upper asymptotes, combined with a logistic-shaped transition between them. For the firing rate λ (in spikes / s) and the cost q, the model is:

$$\mathsf{SPKD}(q;\lambda) = \alpha(\lambda) + \frac{\beta(\lambda)}{1 + \exp\left[-\gamma(\lambda)(\log_{10} q - \delta(\lambda))\right]}$$

Here:

- $\alpha(\lambda) = \sqrt{\frac{4}{\pi\lambda}}$
- $\beta(\lambda) = 2 - \alpha(\lambda)$
- $\gamma(\lambda) = \frac{6.074}{\lambda+7.299} + 1.870$
- $\delta(\lambda) = 0.396 \cdot \log(\lambda + 1.506) + 0.367$

How We Fit It

For each fixed firing rate λ, we first fit a 4-parameter sigmoid function to the empirical SPKD per spike value over a range of cost values q. This gave us an estimate of $\gamma(\lambda)$ and $\delta(\lambda)$ from the lower and upper asymptotes of the sigmoid.

To model how γ and δ change with λ, we tested different smooth functional forms. For $\gamma(\lambda)$ (which decays), we tried:

- Exponential: $ae^{-b\lambda} + c$
- Power-law: $a\lambda^{-b} + c$
- Rational decay: $\frac{a}{\lambda+b} + c$

For $\delta(\lambda)$ (which grows and saturates), we tried:

- Logarithmic: $a\log(\lambda + b) + c$
- Inverse: $a - \frac{b}{\lambda+c}$
- Logistic: $\frac{L}{1+e^{-k(\lambda-x_0)}}$

We used SciPy's `curve_fit` to optimize each candidate and selected the best using R^2 as the evaluation metric. Rational decay was the best fit for $\gamma(\lambda)$, and the logarithmic function worked best for $\delta(\lambda)$.

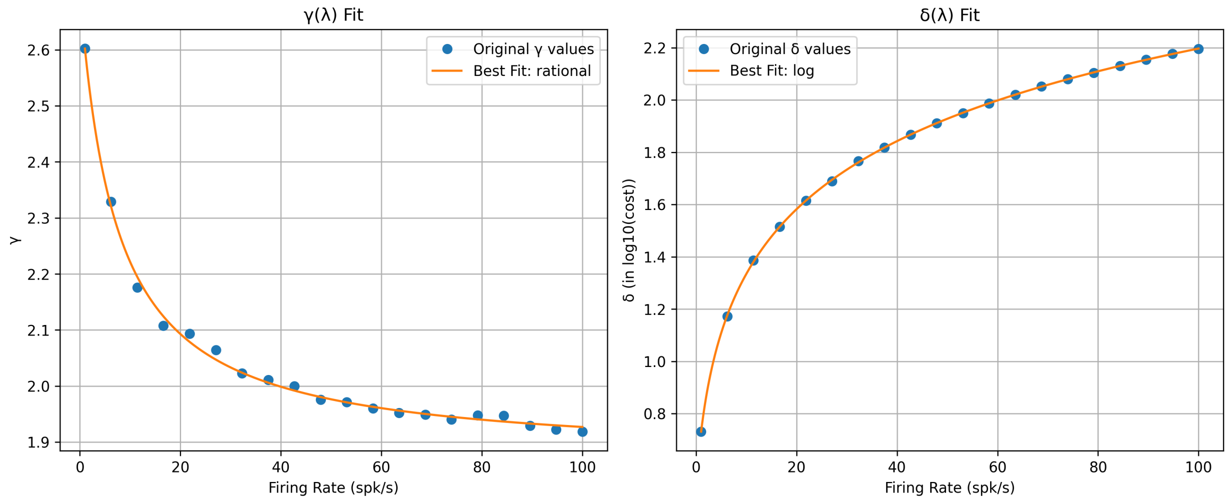

Figure 2: Best fitting functions for $\gamma(\lambda)$ and $\delta(\lambda)$. The rational model closely follows the decay trend of γ, while the logarithmic model captures the saturation behavior of δ.

Theoretical Bounds

Although we ultimately estimated $\gamma(\lambda)$ and $\delta(\lambda)$ empirically by fitting sigmoids, we also derived mathematical expressions to understand their limiting behavior.

Lower Bound (per spike): To understand the minimal possible SPKD between two spike trains drawn from Poisson processes, we considered the expected absolute difference in spike counts between two independent Poisson random variables $N_1, N_2 \sim \text{Poisson}(\lambda T)$. The distribution of the difference $N_1 - N_2$ follows a *skellam distribution*:

$$P(k) = e^{-2\lambda T} \left(\frac{\lambda T}{\lambda T} \right)^{k/2} I_{|k|}(2\lambda T)$$

where I_k is the modified Bessel function of the first kind. The expected absolute difference is:

$$\mathbb{E}[|N_1 - N_2|] = \sum_{k=-\infty}^{\infty} |k| P(k)$$

Since this sum has no closed form, we verified the commonly used approximation:

$$\mathbb{E}[|N_1 - N_2|] \approx \sqrt{\frac{4\lambda T}{\pi}}$$

To do this, we computed the exact expectation numerically by adding $|k|P(k)$ using up to 200 Bessel terms. We then compared this exact value to the normal approximation across a range of firing rates. The comparison (Figure 3) showed that the approximation is very accurate even for moderate values of λ, and the normalized version:

$$\gamma(\lambda) \approx \frac{\mathbb{E}[|N_1 - N_2|]}{\lambda T} \approx \sqrt{\frac{4}{\lambda T \pi}}$$

is a good match to the lower bound observed in our SPKD simulations.

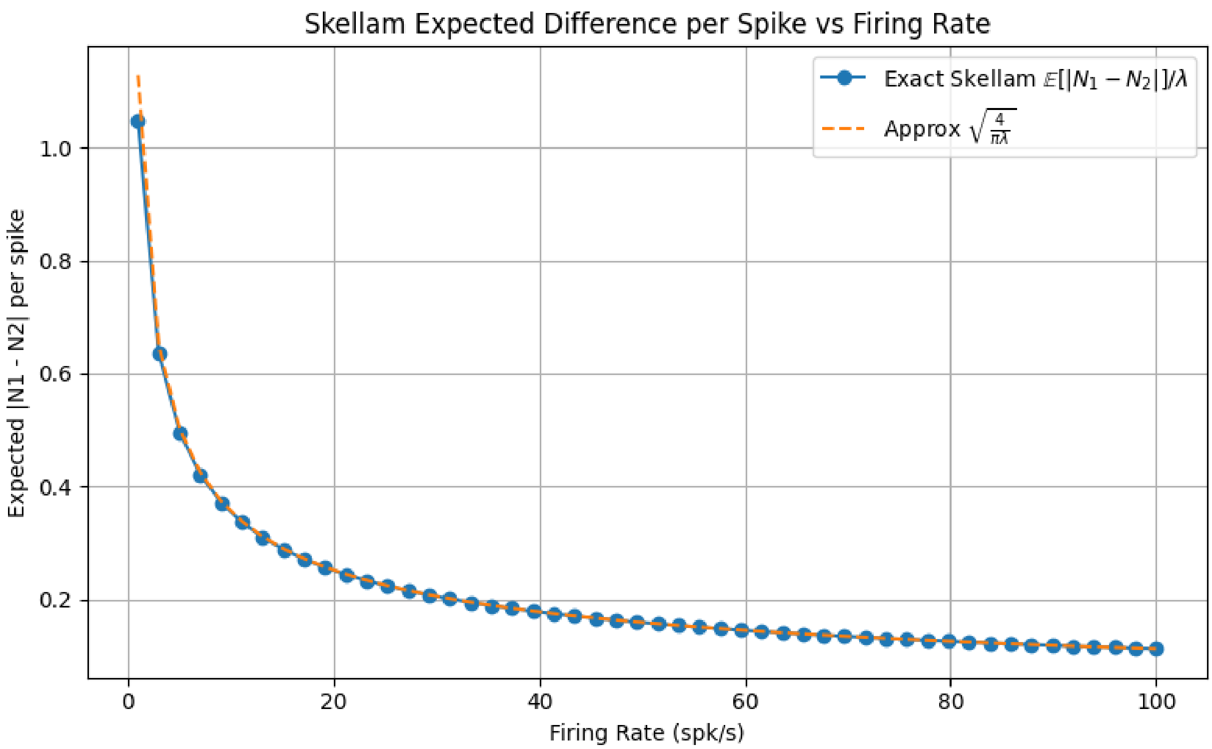

Figure 3: Comparison between the exact Skellam expectation and the normal approximation $\sqrt{4/(\pi\lambda)}$. The agreement confirms the validity of the approximation across a wide range of firing rates.

Upper Bound (per spike): In the extreme case where $q \to \infty$, the shifting of spikes becomes prohibitively expensive. In this case, the optimal Victor–Purpura alignment simply deletes all spikes in one train and adds all spikes from the other. Since both N_1 and N_2 have expected values λT, the total number of operations is:

$$\mathbb{E}[N_1 + N_2] = 2\lambda T$$

Dividing by the expected number of spikes gives:

$$\delta(\lambda) \approx \frac{2\lambda T}{\lambda T} = 2$$

These theoretical bounds are useful for anchoring the sigmoid model and are closely matched with the limits learned from the data.

Final Model Equation

Using the best-fit functions and the the learned parameters for $\gamma(\lambda)$ and $\delta(\lambda)$, the final form of our model is:

$$\text{SPKD}(q;\lambda)=\sqrt{\frac{4}{\pi\lambda}}+\frac{2-\sqrt{\frac{4}{\pi\lambda}}}{1+\exp\left[-\left(\frac{6.074}{\lambda+7.299}+1.870\right)\left(\log_{10}q-(0.396\cdot\log(\lambda+1.506)+0.367)\right)\right]}$$

This expression predicts the expected SPKD per spike for a given cost q and the firing rate λ. The functions $q_0(\lambda)$ and $k(\lambda)$ are still learned empirically per rate.

Validation

To test the model's ability to generalize, we validated it on firing rates that were not used during training. We selected ten intermediate firing rates between 5 and 95 spk/s and compared simulated SPKD values to model estimates across a range of cost values.

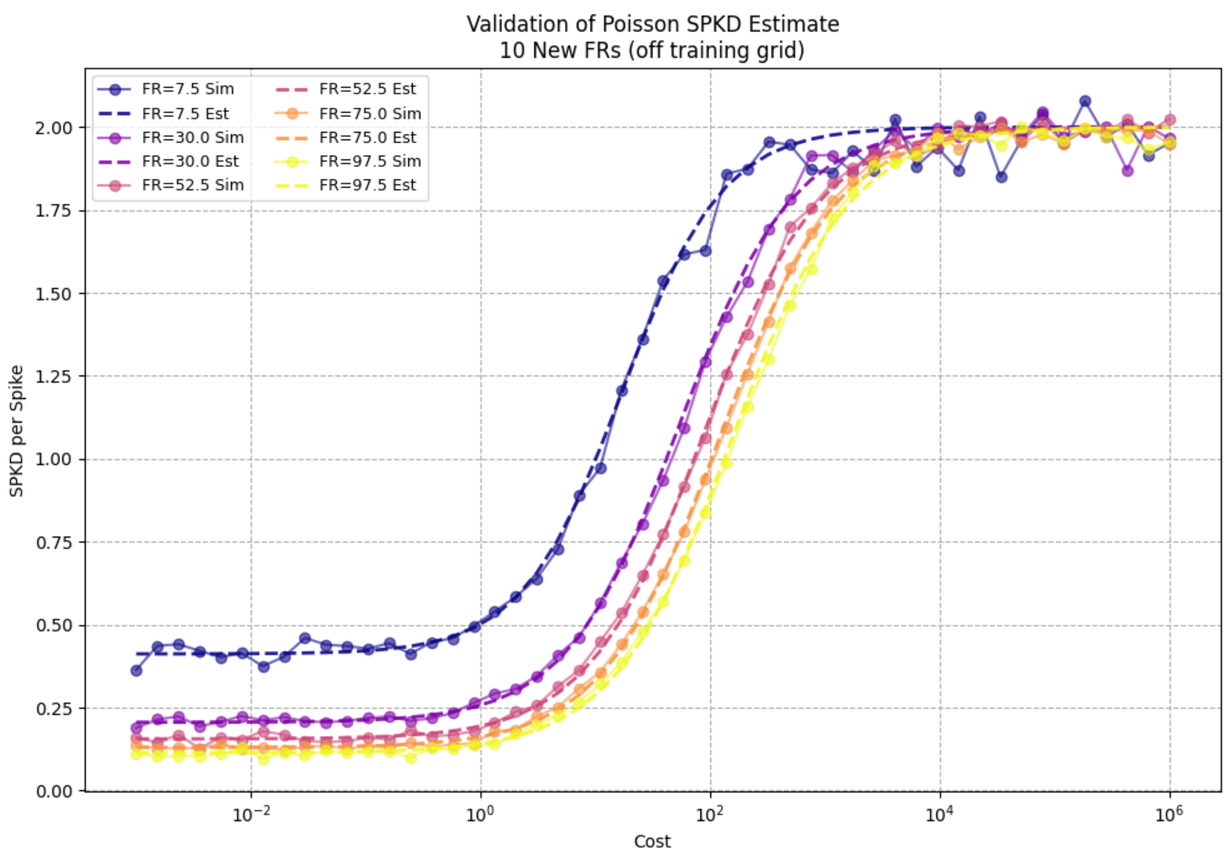

Figure 4: Validation of the Poisson SPKD model on firing rates not used during training. Each solid curve represents SPKD values from simulation, while dashed curves show model predictions. The model remains accurate across the entire cost domain.

As shown in As shown in Figure 4, the predicted values closely match the simulated values across all firing rates and cost ranges, confirming the model's ability to generalize beyond the training set.