Lazy ring + one directed shortcut: Bimodality at $K = 2$ vs. $K = 4$ and the jump-over mechanism (Exact FPT via AW inversion + trajectory decomposition)

1 Problem restatement and objective

In the lazy ring with one directed shortcut (selfloop rule) we previously observed (for $K = 2$, i.e. only ± 1 moves) that a double peak in the first-passage/first-absorption pmf $f(t)$ only appears for sufficiently small shortcut strength β. As β increases, the shortcut is taken more often, mass concentrates on a faster time scale, and the slow peak is suppressed, so bimodality disappears.

Here we fix the start node to n_0 and set the absorbing target to the "farthest" (approximately antipodal) node, then scan the $K = 4$ case (allowed ring moves $\pm 1, \pm 2$). The key new feature at $K = 4$ is that a single ring move can *jump over* the absorbing target without landing on it (e.g. from target $- 1$ to target $+ 1$), which may alter the bimodal mechanism.

The goal is two-layered:

- **Layer A (where bimodality occurs):** use an exact/high-precision method to scan parameters and apply one fixed bimodality criterion to flag true double peaks.
- **Layer B (why bimodality occurs):** for representative bimodal cases, decompose trajectories into *mutually exclusive* classes based on shortcut usage and jump-over events, and quantify how those classes populate the first peak / valley / second peak.

2 Model definition (fixed convention)

2.1 State space and ring moves

We use paper-style indexing $1, \ldots, N$ with all additions taken modulo N. Let $K = 2k$ be the ring degree. At each step, with probability q the walker makes a ring move and with probability $1 - q$ it stays (lazy self-loop). Given that a ring move occurs, the displacement is chosen uniformly from $\{\pm 1, \ldots, \pm k\}$; hence each ring edge has probability q/K.

2.2 One directed shortcut (selfloop rule)

We add a single directed shortcut src $\to$ dst. Under the *selfloop rule*, at the source we take probability mass from the self-loop and reassign it to the shortcut:

$$p = \beta(1 - q), \qquad 0 \le p \le 1 - q \quad (\Leftrightarrow \ 0 \le \beta \le 1).$$

Explicit one-step probabilities:

- for $i \neq$ src:
$$P(i \to i) = 1 - q, \qquad P(i \to i \pm r) = \frac{q}{K}, \quad r = 1, \ldots, k;$$
- for $i =$ src:
$$P(\text{src} \to \text{src}) = 1 - q - p, \quad P(\text{src} \to \text{dst}) = p, \quad P(\text{src} \to \text{src} \pm r) = \frac{q}{K}.$$

2.3 Absorption and first-absorption time

When a step ends at target, absorption occurs with probability ρ; otherwise the walker survives and continues from target. In this report we take $\rho = 1$ (strict first passage), so the first-absorption time T equals the first hitting time of target.

2.4 Geometry (paper-like, with antipodal target)

We fix

$$n_0 = 1, \qquad \text{target} = n_0 + \left\lfloor \frac{N}{2} \right\rfloor, \qquad \text{src} = n_0 + 5, \qquad \text{dst} = \text{target} + 1,$$

with modular arithmetic on the ring. The choice dst = target + 1 avoids the special "shortcut points directly to the absorber" mechanism (dst = target), which is discussed separately in `valley_dst_report`.

2.5 Relation to other reports in this repo

Other reports (e.g. `valley_report` / `valley_dst_report`) also use a "add an outgoing edge at src and renormalize" construction, often in a non-lazy setting. That convention differs from the present selfloop rule (which preserves the ring-move probability q). Accordingly, in this report all scans and statistics refer to the **lazy + selfloop** convention unless explicitly stated.

3 Bimodality criteria (fixed, paper-style)

We use the same criterion as in existing reports:

- **Strict local peak:** $f(t) > f(t-1)$, $f(t) > f(t+1)$, and $f(t) \geq h_{\min}$, with boundary convention $f(0) = f(T+1) = 0$.
- **Paper bimodality:** at least two peaks, and the 2nd-highest peak height $\geq 1\%$ of the highest.
- **Macro bimodality (optional):** additionally require $t_2/t_1 \geq 10$ for the two dominant peaks.

Defaults: $h_{\min} = 10^{-7}$, relative-height threshold 1%, macro ratio threshold 10.

4 Two path events and mutually-exclusive classes

4.1 Shortcut traversal count C

Let C be the number of times a trajectory traverses the directed edge src $\to$ dst. We use the coarse indicator $C \geq 1$ (and also record the $0/1/\geq 2$ breakdown for robustness).

4.2 Jump-over count J (new at $K = 4$)

For $K = 4$ (allowed ± 2 ring moves), a ring step can cross the target without landing. We count jump-over events **only on ring moves** (not on self-loops or shortcut steps). A step from x to $x \pm r$ ($r \geq 2$) is a jump-over if the target lies strictly between departure and arrival along that direction:

$$\text{target} \in \{x \pm 1, \ldots, x \pm (r-1)\}.$$

Equivalently, for a $+r$ move define the directed ring distance $d_+(x, \text{target}) = (\text{target} - x) \bmod N$; jump-over occurs iff $1 \leq d_+ \leq r-1$ (and similarly for $-r$ using $d_-(x, \text{target}) = (x - \text{target}) \bmod N$). For $K = 2$ we always have $J \equiv 0$.

4.3 Four mutually-exclusive classes (A/B/C/D)

Using the indicators $(C \geq 1)$ and $(J \geq 1)$ we define:

$$\text{A: } (C = 0, J = 0), \quad \text{B: } (C \geq 1, J = 0), \quad \text{C: } (C = 0, J \geq 1), \quad \text{D: } (C \geq 1, J \geq 1),$$

which are mutually exclusive and exhaustive. For $K = 2$, the classification degenerates to shortcut vs. nonshortcut.

5 Peak–valley–peak windows and conditional statistics

Given an exact $f(t)$, we locate the first two peaks in time order $(t_1 < t_2)$ under the paper criterion, and define the valley bottom

$$t_v = \arg\min_{t_1 < t < t_2} f(t).$$

We set a symmetric window half-width

$$\Delta = \max\left(1, \lfloor 0.05\,(t_2 - t_1)\rfloor\right),$$

and label each Monte Carlo trajectory by the nearest center among $\{t_1, t_v, t_2\}$ if it lies within distance Δ (ties are assigned to "other"). For each window W we report:

- window mass $\Pr(T \in W)$;
- $\Pr(C \geq 1 \mid T \in W)$ and $\Pr(J \geq 1 \mid T \in W)$ (Wilson 95% intervals recorded in `*.summary.json`);
- class fractions for A/B/C/D.

6 Exact $f(t)$ computation: AW inversion vs. time-domain iteration

6.1 AW inversion for the main model

For the main model (point absorption at target, no "jumpover absorbs" modification), we compute the generating function and invert it via Abate–Whitt (AW).

Defect-free lazy ring (spectral form) The defect-free eigenvalues are

$$\lambda_\ell = (1 - q) + \frac{2q}{K} \sum_{r=1}^{k} \cos\left(\frac{2\pi r \ell}{N}\right), \qquad \ell = 0, \ldots, N - 1.$$

Define the translation-invariant resolvent

$$Q(d, z) = \frac{1}{N} \sum_{\ell=0}^{N-1} \frac{\cos(2\pi \ell d/N)}{1 - z\lambda_\ell}, \qquad |z| < 1,$$

where d is the ring displacement.

Rank-one defect (selfloop shortcut) and point absorption The selfloop shortcut modifies only the outgoing probabilities at src, i.e. a single-column rank-one defect relative to the defect-free ring. We evaluate the corresponding resolvent entries $S_{i,j}(z)$ using a Sherman–Morrison update expressed in terms of $Q(\cdot, z)$. For imperfect absorption $\rho \in (0, 1]$, the first-absorption generating function is

$$\tilde{f}(z) = \sum_{t \geq 1} f(t) z^t = \frac{\rho\, S_{\text{target},n_0}(z)}{(1 - \rho) + \rho\, S_{\text{target,target}}(z)}.$$

For $\rho = 1$, $\tilde{f}(z) = S_{\text{target},n_0}(z)/S_{\text{target,target}}(z)$.

AW (FFT) coefficient recovery We sample $z_m = re^{2\pi im/L}$ and approximate

$$f(t) \approx r^{-t}\frac{1}{L}\sum_{m=0}^{L-1}\tilde{f}(z_m)\,e^{-2\pi imt/L}, \qquad t = 1,\ldots,T_{\max},$$

implemented via FFT. We use $r = 10^{-4/T_{\max}}$ and L the next power of two exceeding $T_{\max}$ by a small pad, consistent with other AW scripts in this repo. For the representative plots we take $T_{\max} = 4000$ to cover both dominant peaks and the valley.

6.2 Time-domain master-equation iteration (scans and nonlocal control)

Layer A scans require suppressing tail mass (e.g. survival $\lesssim 10^{-14}$), so we use time-domain master-equation iteration (absorbing flux recursion) to obtain numerically exact $f(t)$ for the fixed bimodality criterion. The counterfactual control "jumpover → absorb" makes absorption a nonlocal event (crossing without landing), so it is also computed in the time domain rather than by the point-absorption AW formula.

7 Layer A: where bimodality truly occurs (parameter scans)

7.1 N scan at fixed $\beta = 0.02$

Fix $q = 2/3$, $\beta = 0.02$ (so $p = \beta(1-q)$), and the geometry above. Scan even $N = 10..200$:

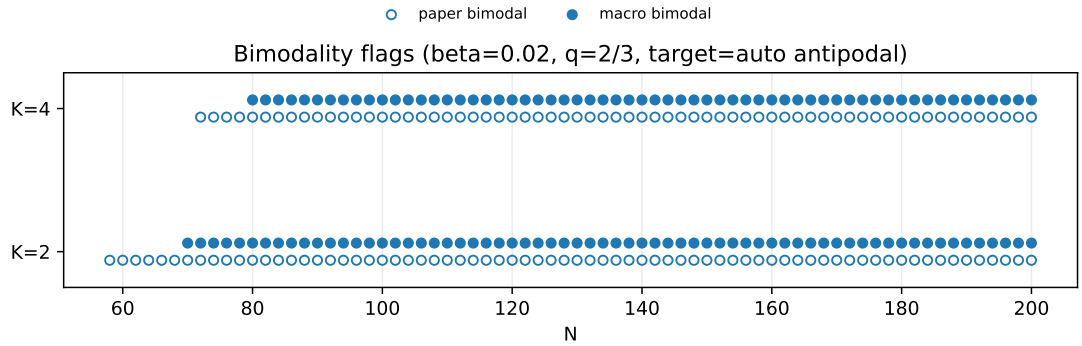


Figure 1: Bimodality flags for $\beta = 0.02$ at $K = 2$ vs. $K = 4$ (hollow = paper, filled = macro).

7.2 β scan at fixed $N = 100$

Fix $N = 100$, $q = 2/3$ and scan β:

8 Layer B: mechanism via trajectory decomposition

8.1 Representative point: $N = 100, \beta = 0.02$ ($K = 2$ vs. $K = 4$)

We first show the exact $f(t)$ (log scale) with peak/valley/peak markers:

Using the same window definition, we show window-conditional class fractions:

8.1.1 Quantitative window summaries

To make the mechanism claims auditable, we tabulate the window-conditional statistics (generated from `*.summary.json`):

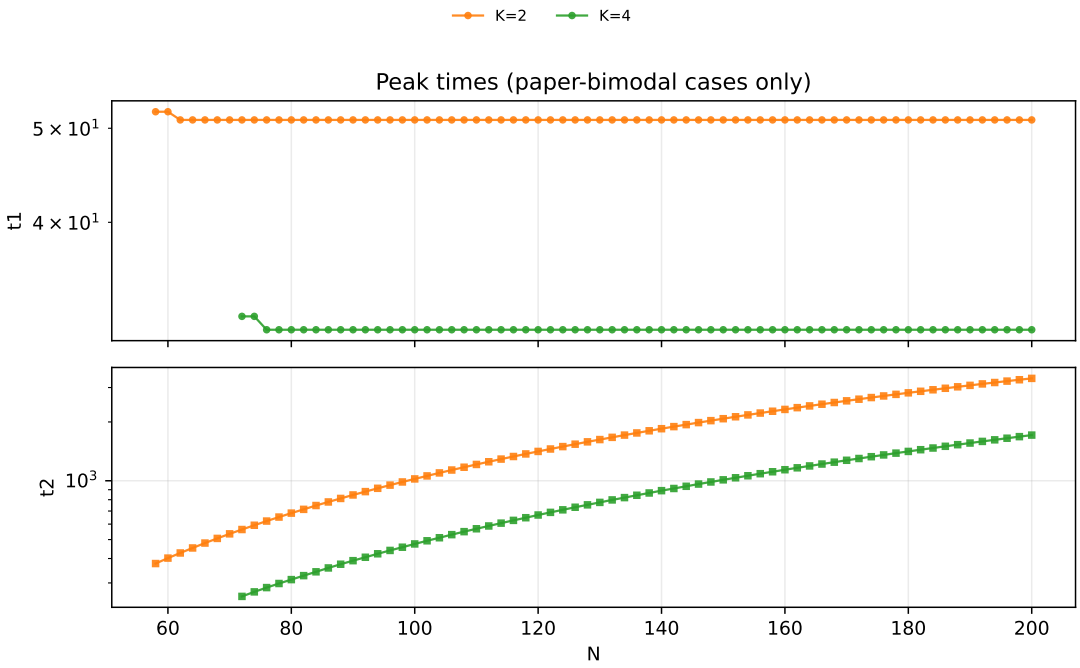


Figure 2: Peak times t_1, t_2 (log scale), restricted to paper-bimodal cases.

β	p	paper	macro	t_1	t_2
0	0	0	0	–	–
0.0005	0.000166667	1	1	54	1243
0.001	0.000333333	1	1	54	1236
0.002	0.000666667	1	1	54	1222
0.005	0.00166667	1	1	53	1183
0.01	0.00333333	1	1	53	1124
0.02	0.00666667	1	1	51	1022
0.03	0.01	1	1	50	935
0.05	0.0166667	1	1	48	780
0.07	0.0233333	0	0	–	–
0.1	0.0333333	0	0	–	–
0.15	0.05	0	0	–	–
0.2	0.0666667	0	0	–	–
0.3	0.1	0	0	–	–
0.5	0.166667	0	0	–	–
1	0.333333	0	0	–	–

Table 1: Beta scan at N=100, K=2 (±1), q=2/3, target=auto (ãntipodal), selfloop rule p=beta(1-q): paper/macro flags and peak times.

β	p	paper	macro	t_1	t_2
0	0	0	0	–	–
0.0005	0.000166667	0	0	–	–
0.001	0.000333333	1	1	32	515
0.002	0.000666667	1	1	32	513
0.005	0.00166667	1	1	32	506
0.01	0.00333333	1	1	32	495
0.02	0.00666667	1	1	31	475
0.03	0.01	1	1	31	457
0.05	0.0166667	1	1	30	424
0.07	0.0233333	1	1	29	395
0.1	0.0333333	1	1	28	356
0.15	0.05	1	1	27	295
0.2	0.0666667	0	0	–	–
0.3	0.1	0	0	–	–
0.5	0.166667	0	0	–	–
1	0.333333	0	0	–	–

Table 2: Beta scan at N=100, K=4 ($\pm1, \pm2$), q=2/3, target=auto (ãntipodal), selfloop rule p=beta(1-q): paper/macro flags and peak times.

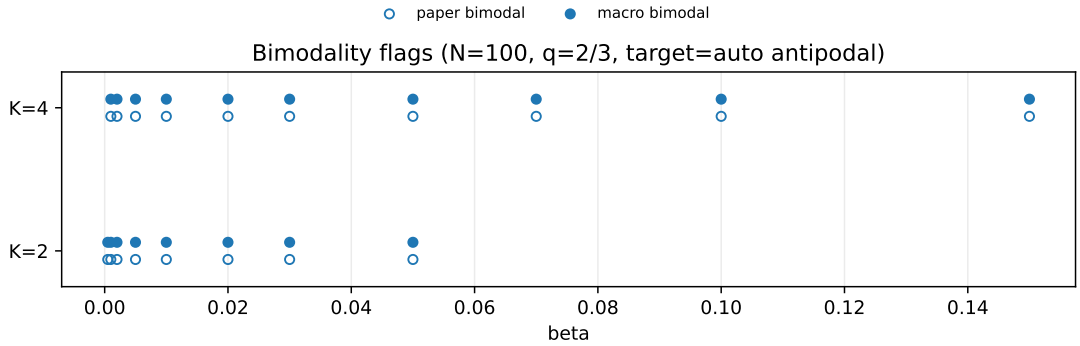


Figure 3: Bimodality flags vs. β at $N = 100$ for $K = 2$ and $K = 4$.

window	t_c	n	$\Pr(T \in W)$	$\Pr(C \geq 1 \mid W)$	$\Pr(J \geq 1 \mid W)$	A	B	C	D
peak1	51	4161	0.021	1.000	0.000	0.000	1.000	0.000	0.000
valley	291	3330	0.017	0.908	0.000	0.092	0.908	0.000	0.000
peak2	1022	5712	0.029	0.263	0.000	0.737	0.263	0.000	0.000

Table 3: N=100, K=2 (±1), $q = 2/3, \beta = 0.02$: window-conditional statistics (MC) ($\Delta = 48$). A/B/C/D: A=($C = 0, J = 0$), B=($C \geq 1, J = 0$), C=($C = 0, J \geq 1$), D=($C \geq 1, J \geq 1$).

window	t_c	n	$\Pr(T \in W)$	$\Pr(C \geq 1 \mid W)$	$\Pr(J \geq 1 \mid W)$	A	B	C	D
peak1	31	1825	0.009	1.000	0.302	0.000	0.698	0.000	0.302
valley	95	1658	0.008	0.935	0.370	0.052	0.578	0.013	0.356
peak2	475	5980	0.030	0.125	0.280	0.643	0.076	0.231	0.049

Table 4: N=100, K=4 ($\pm1, \pm2$), $q = 2/3, \beta = 0.02$: window-conditional statistics (MC) ($\Delta = 22$). A/B/C/D: A=($C = 0, J = 0$), B=($C \geq 1, J = 0$), C=($C = 0, J \geq 1$), D=($C \geq 1, J \geq 1$).

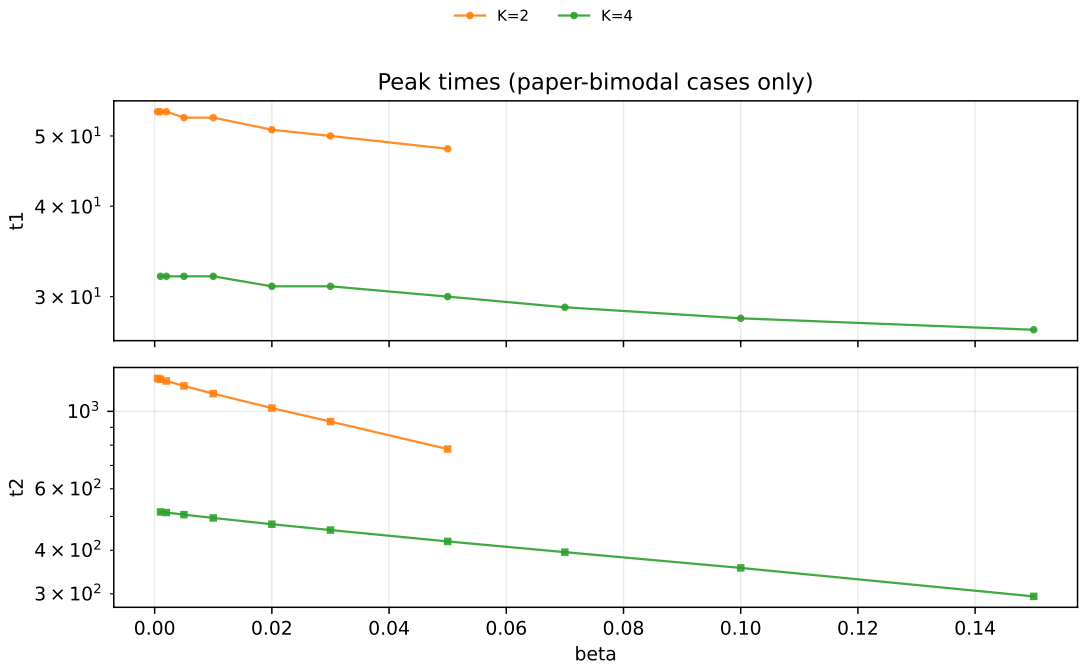


Figure 4: Peak times t_1, t_2 vs. β at $N = 100$ (paper-bimodal cases only).

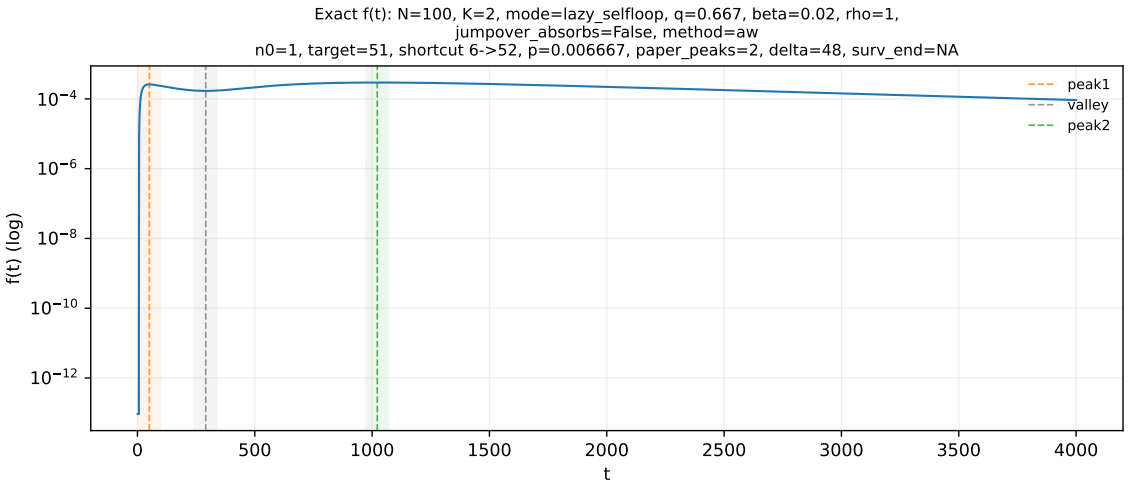


Figure 5: $K = 2$: exact $f(t)$ via AW inversion ($q = 2/3, \beta = 0.02$).

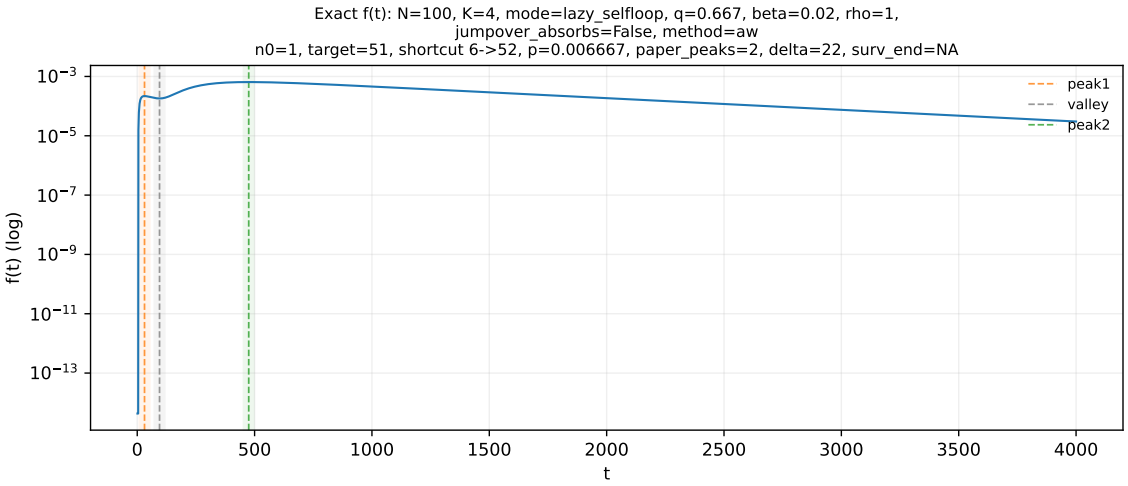


Figure 6: $K = 4$: exact $f(t)$ via AW inversion ($q = 2/3, \beta = 0.02$).

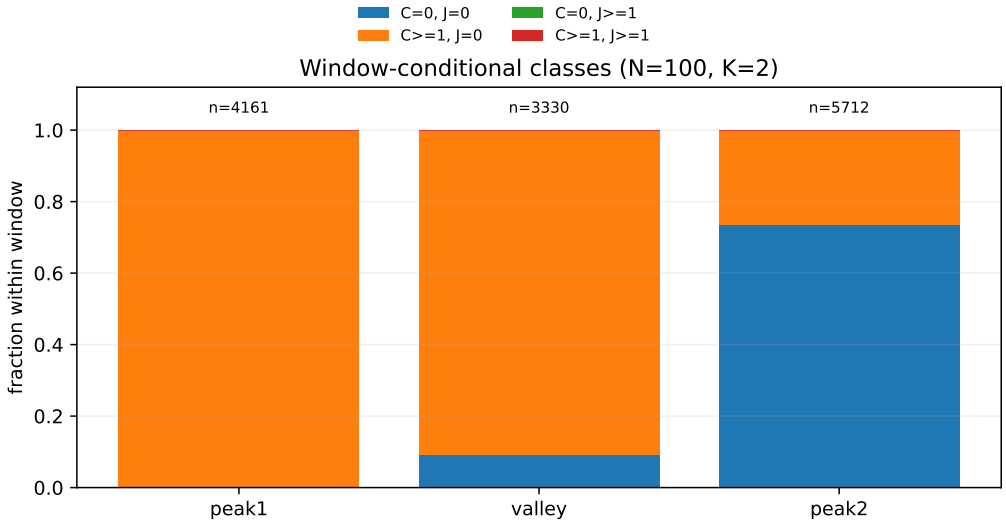


Figure 7: $K = 2$: A/B/C/D class fractions within peak1/valley/peak2. Here $J \equiv 0$ so it reduces to shortcut vs. nonshortcut.

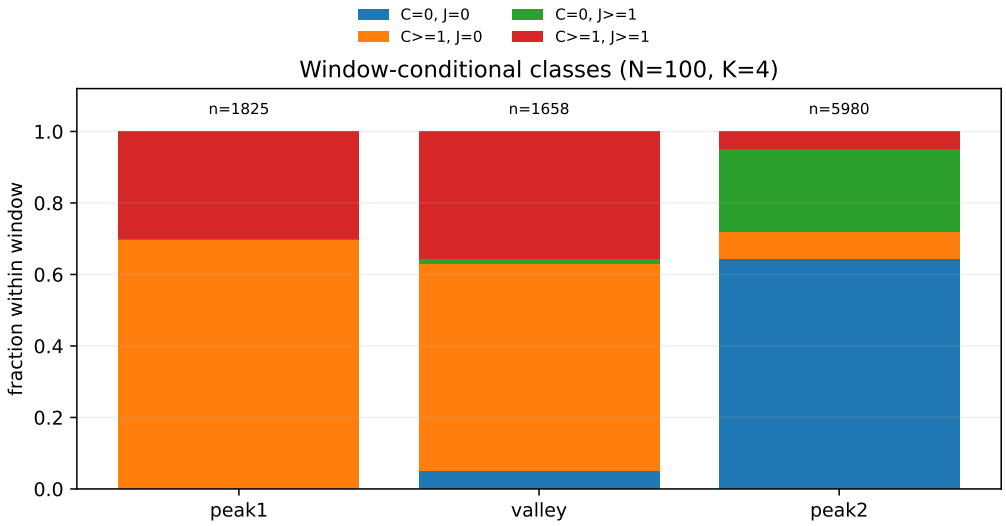


Figure 8: $K = 4$: A/B/C/D class fractions within peak1/valley/peak2.

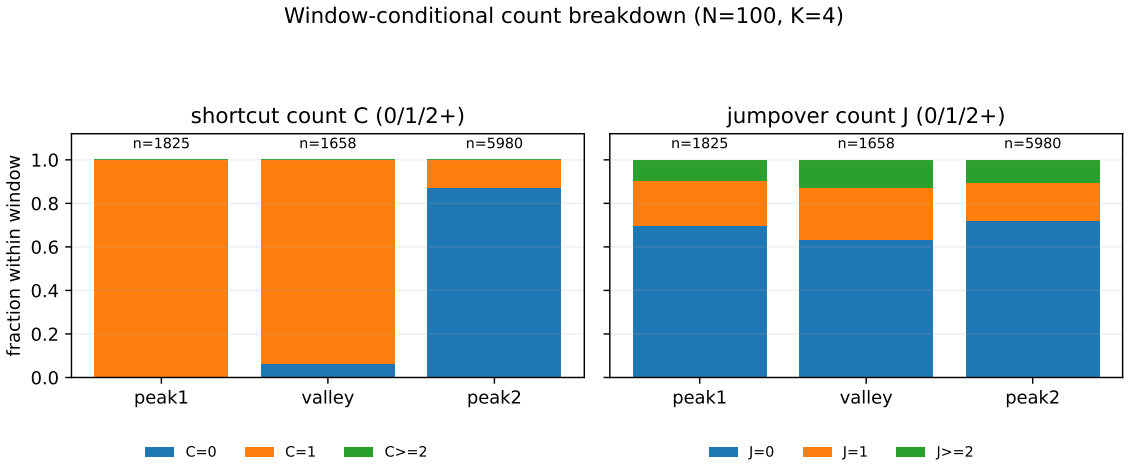


Figure 9: $K = 4$: $0/1/\geq 2$ breakdown for C and J within the three windows (to test whether repeated overshoots accumulate near the slow peak).

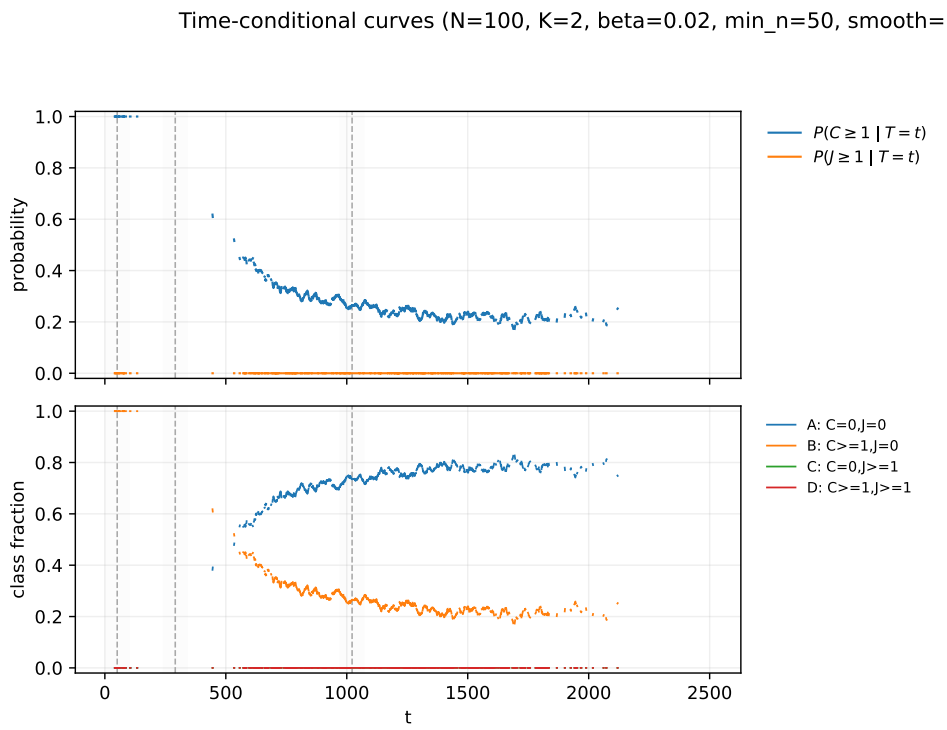


Figure 10: $K = 2$: time-conditional curves $\Pr(C \geq 1 \mid T = t)$ and class weights vs. t (MC).

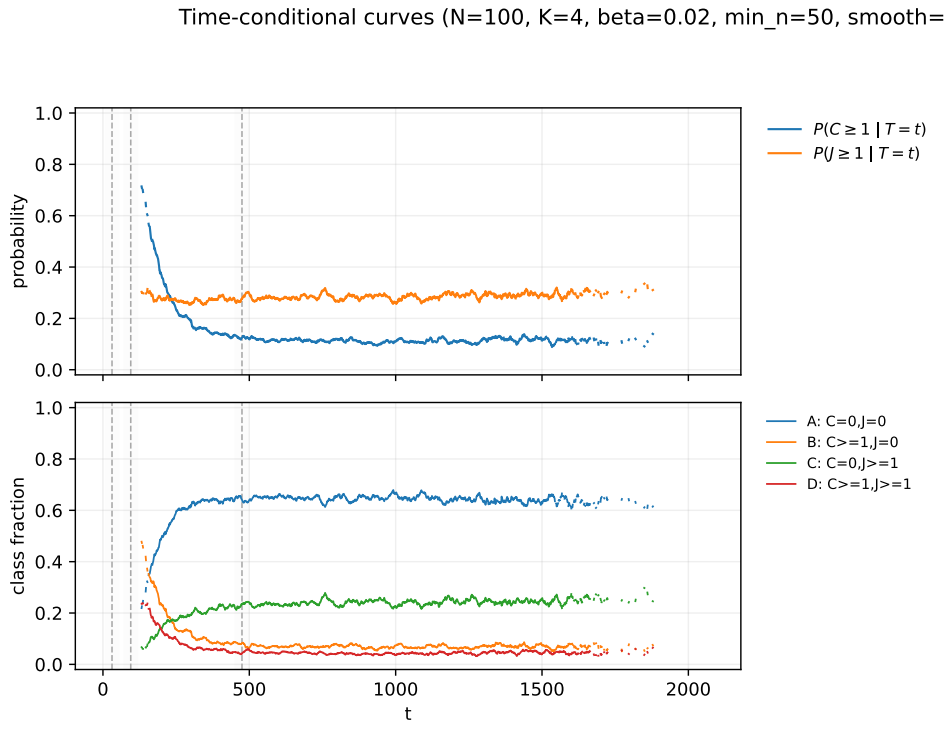


Figure 11: $K = 4$: time-conditional curves $\Pr(C \geq 1 \mid T = t)$, $\Pr(J \geq 1 \mid T = t)$ and class weights vs. t (MC).

8.1.2 Mechanism takeaways

- **Peak 1 is shortcut-dominated:** in both $K = 2$ and $K = 4$, within the peak1 window we have $\Pr(C \geq 1 \mid W) \approx 1$, corresponding to trajectories that reach src early and use the shortcut. This produces the early peak (here $t_1 = 51$ for $K = 2$ and $t_1 = 31$ for $K = 4$).
- **The valley is a class-weight transition:** in the valley window, the $C = 0$ fraction rises; for $K = 4$ the jump-over probability also increases, indicating that "near-target but overshot" trajectories become common.
- **Peak 2 is dominated by nonshortcut trajectories:** in the peak2 window, $C = 0$ becomes the majority class; for $K = 4$ this remains true (A+C dominates B+D), identifying peak2 as a slow subpopulation not effectively using the shortcut.
- **The key $K = 4$ difference is the jump-over subfamily:** in $K = 4$ peak2, the nonshortcut family splits into A ($C = 0, J = 0$) and C ($C = 0, J \geq 1$), with C carrying a substantial fraction of the slow-peak window mass. This directly shows jump-over is a real delay mechanism inside the slow component, not a rare artifact.

In mixture form,

$$f(t) = \sum_{(c,j)\in\{0,1\}^2} \Pr(C \geq 1 = c,\ J \geq 1 = j)\, f(t \mid C \geq 1 = c,\ J \geq 1 = j),$$

which reduces to a two-class mixture for $K = 2$ ($J \equiv 0$), but requires the (C, J) refinement for $K = 4$.

Why jump-over changes the slow time scale For $K = 2$, any attempt to "cross" the target necessarily lands on it (step length is 1), so once a trajectory enters the target neighborhood absorption is essentially determined by whether it lands on the target. For $K = 4$, when the walker sits at target ± 1, a ∓ 2 ring move sends it to the opposite side without absorption; this inserts an additional return cost into near-absorbing local paths (possibly repeatedly), delaying a subset of trajectories and reshaping the slow component.

8.2 Counterfactual control: treat jump-over as absorption

To separate "step-length distribution changes" from the nonlocal "missed absorption by overshoot" mechanism, we compare the baseline $K = 4$ model to a counterfactual where any jump-over event triggers absorption.

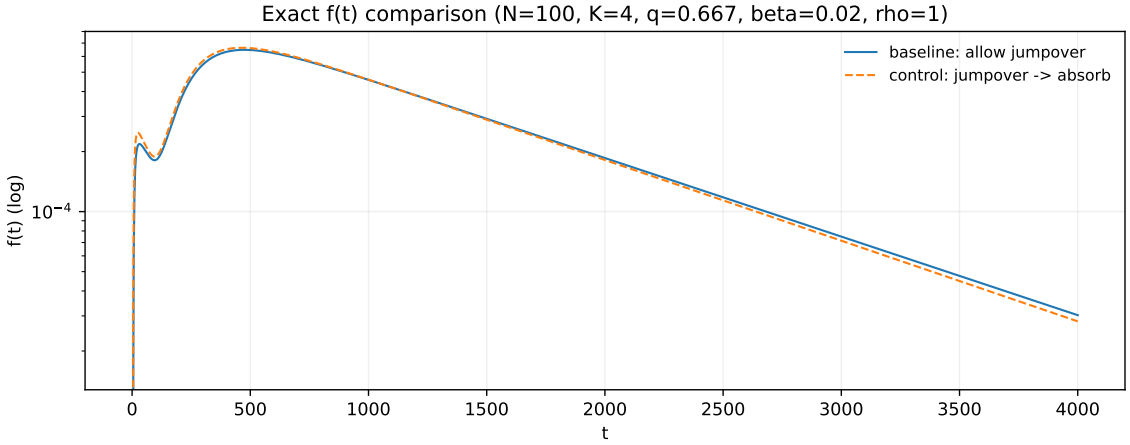


Figure 12: $K = 4$ at $N = 100, q = 2/3, \beta = 0.02$: exact $f(t)$ overlay (time-domain master equation): baseline (allow jump-over) vs. control (jump-over $\rightarrow$ absorb).

8.3 Larger shortcut strength: $K = 4, N = 100, \beta = 0.15$

To align with the intuition that larger β suppresses the slow peak, we also show a representative larger-β $K = 4$ case:

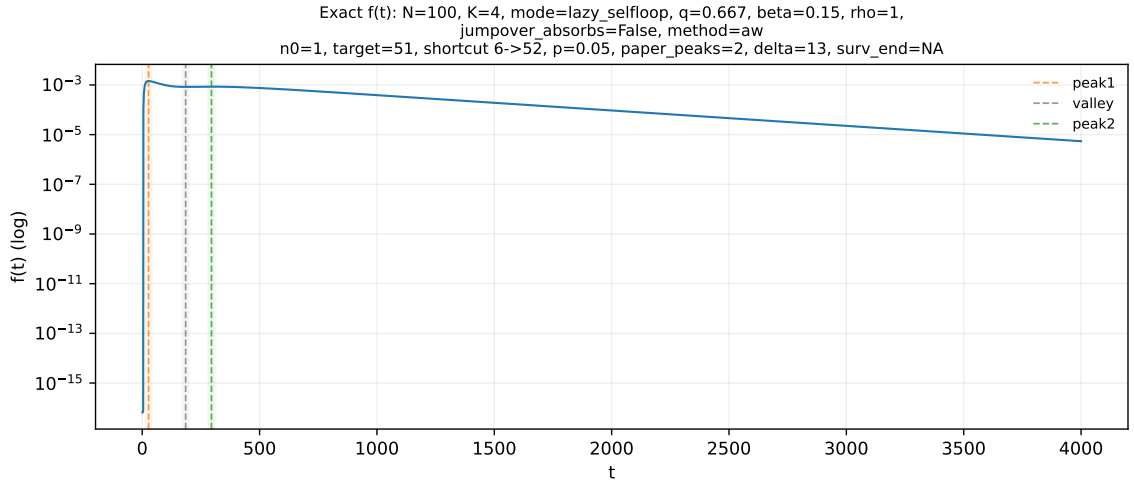


Figure 13: $K = 4$: exact $f(t)$ via AW inversion ($N = 100, q = 2/3, \beta = 0.15$).

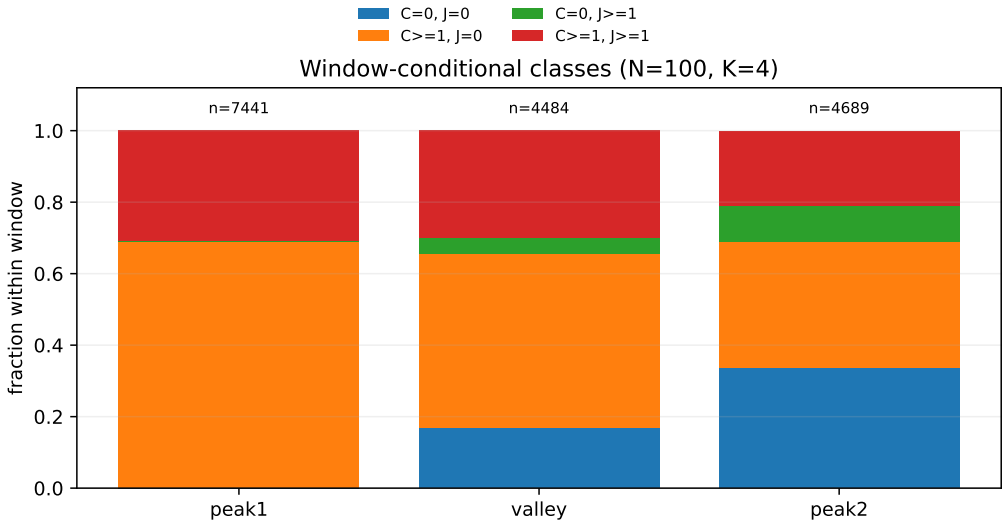


Figure 14: $K = 4$: window-conditional A/B/C/D fractions ($N = 100, q = 2/3, \beta = 0.15$).

9 Reproducibility: scripts and commands

All results are generated in this report directory using:

- exact distributions + scans + MC: `code/jumpover_bimodality_pipeline.py`
- scan plots: `code/plot_scan_summaries.py`
- beta-scan tables: `code/make_beta_scan_table.py`
- window-stat tables: `code/make_case_window_table.py`

```
# representative case (exact + MC), AW inversion
python3 code/jumpover_bimodality_pipeline.py analyze \
  --N 100 --K 4 --beta 0.02 --q 0.6666667 \
  --exact-method aw --max-t-plot 4000 \
```

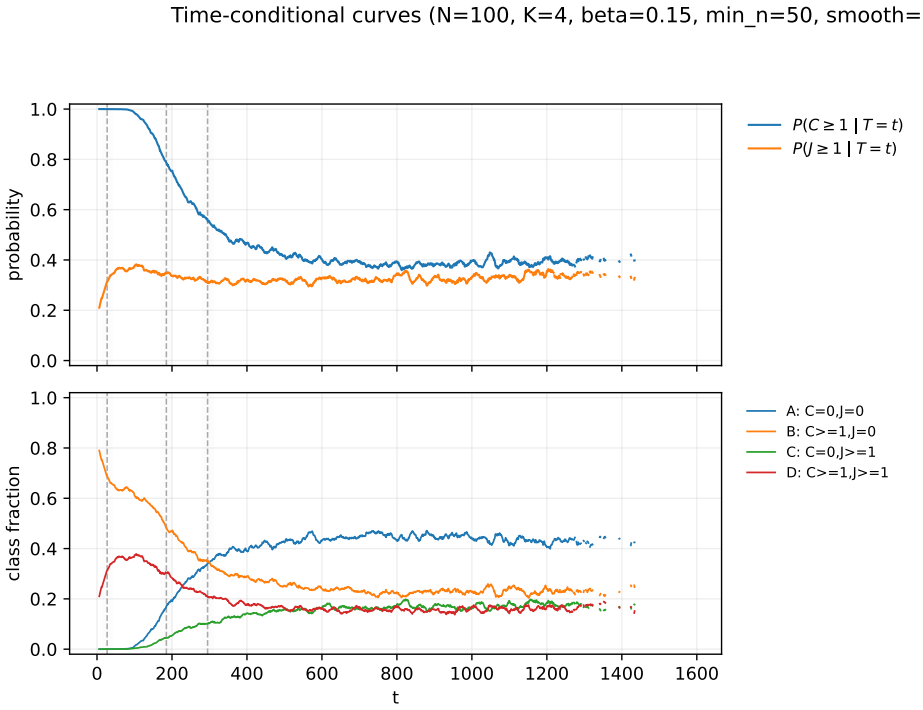


Figure 15: $K = 4$ ($\beta = 0.15$): time-conditional curves (compare to $\beta = 0.02$).

```
  --target auto --sc-src auto+5 --sc-dst auto_target+1 \
  --n-walkers 200000 --seed 0 \
  --out-prefix data/cases/N100_K4_beta002/N100_K4_beta002

# window-stat table from summary.json
python3 code/make_case_window_table.py \
  --summary data/cases/N100_K4_beta002/N100_K4_beta002.summary.json \
  --out tables/case_window_stats_N100_K4_beta002_en.tex \
  --caption "N=100, K=4 (\pm1,\pm2), q=2/3,beta=0.02: \
  window-conditional statistics (MC)" \
  --label tab:case-window-N100-K4-beta002-en
```