Correlation–Path Analysis of Batting Kinetics from Monocular Video

Physics, Aesthetics, and Clear Mathematical Forms

Abstract

We propose an analysis that extracts a single, directed "correlation path" summarizing how motion flows through a batter's body into the bat from a short monocular video. The method is intentionally minimal: one axis of motion, one time axis, one chain. This compression is not just practical—it is aesthetic—because it aligns with the physics of proximal-to-distal sequencing and with how humans perceive coherent action.

1 Why this is aesthetically right (and physically motivated)

Physics first. In striking tasks (cricket, baseball, tennis) effective power transfer typically follows a *proximal-to-distal* sequence: ground reaction $\to$ legs/pelvis $\to$ trunk/shoulders $\to$ arm/hand $\to$ implement. Angular momentum and impulse are built in larger, proximal segments and passed to lighter, distal ones. Efficient swings therefore show *staggered peaks* in segmental speeds with consistent lead–lag. Our method measures these lags directly and chooses the single most coherent chain that transmits them to the bat.

Why it looks good. A representation is pleasing when it obeys three principles:

- *Economy:* Many joint trajectories collapse to a one–dimensional *swing axis* and a single *path*. Less clutter reveals structure.
- *Directionality:* A directed acyclic graph (DAG) and explicit time lags put arrows on time, matching the causal *feel* of motion.
- *Salience:* We weight edges by correlation strength; thick, high-confidence links trace the kinetic storyline, thin/noisy links recede.

The result is a legible "score" of the swing: tempo (bat-tip speed), entry points (activation times), and the main *melody line* (maximum-weight path) that our visual system can follow at a glance.

2 Data and notation

Video V at frame rate f Hz with frames $t = 1, \dots, T$. For joint $i \in J$ (ankles, knees, hips, pelvis, shoulders, elbows, wrists), let $p_{i,t} \in \mathbb{R}^2$ be the 2D keypoint in pixels. Let $b_t \in \mathbb{R}^2$ be the bat-tip location. All analysis is in 2D per frame.

3 Pre-processing

3.1 Root-centering and scale normalization

Center at pelvis and normalize by median shoulder width:

$$\tilde{p}_{i,t} = p_{i,t} - p_{\text{pelvis},t}, \tag{1}$$

$$\bar{s} = \text{median}_t \big\| p_{\text{rshoulder},t} - p_{\text{lshoulder},t} \big\|_2, \tag{2}$$

$$\hat{p}_{i,t} = \tilde{p}_{i,t}/\bar{s}, \qquad \hat{b}_t = (b_t - p_{\text{pelvis},t})/\bar{s}. \tag{3}$$

This makes trajectories approximately scale- and translation-invariant across clips.

3.2 Temporal smoothing and velocities

Apply a Savitzky–Golay filter (window w, polynomial order m) to each coordinate of $\hat{p}_{i,t}$ and $\hat{b}_t$ to suppress jitter while preserving peaks. Use central differences for velocities (valid indices):

$$\dot{p}_{i,t} = \frac{\hat{p}_{i,t+1} - \hat{p}_{i,t-1}}{2\Delta t}, \qquad \dot{b}_t = \frac{\hat{b}_{t+1} - \hat{b}_{t-1}}{2\Delta t}, \quad \Delta t = \frac{1}{f}. \tag{4}$$

4 Swing-axis projection (physics-informed dimensionality reduction)

Estimate the dominant in-plane swing direction by PCA on the bat-tip trace $\{\hat{b}_t\}_{t=1}^{T}$. The first principal component $u \in \mathbb{R}^2$ (unit length) maximizes explained variance of $\hat{b}_t$. Project all velocities onto u:

$$v_{i,t} = u^\top \dot{p}_{i,t} \in \mathbb{R}, \qquad v_t^{(b)} = u^\top \dot{b}_t \in \mathbb{R}. \tag{5}$$

Why: swings are largely planar; projecting onto u emphasizes the kinetic chain that pushes the bat along its path and collapses lateral noise. Aesthetically, this creates a clean, comparable scalar time series for each joint.

5 Lead–lag coupling via time-lagged correlation

For any ordered pair (i, j) with $i, j \in J \cup \{\text{bat}\}$, define mean-centered series $\tilde{v}_{i,t} = v_{i,t} - \bar{v}_i$ and $\tilde{v}_{j,t} = v_{j,t} - \bar{v}_j$, where $\bar{v}_i$ is the sample mean. The *normalized cross-correlation* at lag $\tau \in [-\tau_{\max}, \tau_{\max}]$ is

$$\rho_{ij}(\tau) = \frac{\sum_{t \in \mathcal{I}_\tau} \tilde{v}_{i,t}\, \tilde{v}_{j,t+\tau}}{\sqrt{\sum_{t \in \mathcal{I}_0} \tilde{v}_{i,t}^2}\, \sqrt{\sum_{t \in \mathcal{I}_0} \tilde{v}_{j,t}^2}}, \tag{6}$$

where $\mathcal{I}_\tau = \{t \mid 1 \le t \le T,\ 1 \le t + \tau \le T\}$ enforces valid overlap. Keep the peak value and its lag

$$c_{ij} = \max_\tau \rho_{ij}(\tau), \qquad \tau_{ij}^* = \arg\max_\tau \rho_{ij}(\tau). \tag{7}$$

Interpretation: $\tau_{ij}^* > 0$ means i tends to *lead* j by τ_{ij}^* frames (convert to milliseconds by ms $= 1000\,\tau_{ij}^*/f$).

6 Activation times and DAG construction

Define each node's *activation time* as the frame of peak projected speed:

$$t_i = \arg\max_t |v_{i,t}|, \qquad t_{\text{bat}} = \arg\max_t |v_t^{(b)}|. \tag{8}$$

Form a directed acyclic graph $G = (V, E)$ with $V = J \cup \{\text{bat}\}$. Add edge $i \to j$ when:

(i) $t_i < t_j$ (*temporal order*),

(ii) $\tau^*_{ij} > 0$ (i leads j),

(iii) $c_{ij} \geq \theta$ and the edge is statistically significant (Sec. 7).

Assign edge weight $w_{ij} = c_{ij}$.

7 Reliability with autocorrelation (small-sample correction)

Autocorrelation inflates naive degrees of freedom. Let r_i and r_j be the lag-1 autocorrelations of $(v_{i,t})$ and $(v_{j,t})$. An effective sample size approximation for correlations of AR(1)-like series is

$$N_{\text{eff}} \;=\; N \cdot \frac{1 - r_i r_j}{1 + r_i r_j}, \qquad N = |\mathcal{I}_{\tau^*_{ij}}|. \tag{9}$$

Apply Fisher's z-transform to c_{ij}:

$$z = \operatorname{arctanh}(c_{ij}), \qquad \text{SE}(z) \approx \frac{1}{\sqrt{N_{\text{eff}} - 3}}. \tag{10}$$

Use a two-sided normal test to obtain a p-value; accept edges with $p < \alpha$ (e.g. $\alpha = 0.05$) and $c_{ij} \geq \theta$ (e.g. $\theta = 0.40$). This keeps the graph sparse and visually clean while guarding against spurious lags.

8 The correlation path: maximum-weight chain

Let S be possible sources (e.g. feet and pelvis) and let b denote the bat node. Among all directed paths P from any $s \in S$ to b, select

$$P^* \;=\; \operatorname*{arg\,max}_{P \in \mathcal{P}(S \to b)} \sum_{(i,j) \in P} w_{ij}. \tag{11}$$

This yields a single interpretable kinetic chain (e.g. right_ankle $\to$ pelvis $\to$ right_shoulder $\to$ right_wrist $\to$ bat_tip) annotated with w_{ij} and τ^*_{ij}. *Why it is aesthetic:* one clear storyline, no forks, thickness encoding conviction, and labels encoding timing—like a musical lead sheet of the swing.

9 Across-clip alignment (optional aggregation)

To compare shots with different tempi, align each clip to a common timeline via Dynamic Time Warping (DTW) on the scalar bat-tip speed $|v_t^{(b)}|$. Compute edge frequencies and mean lags on the aligned axis. This makes across-shot summaries visually and statistically coherent.

10 What the output plot shows (and why it reads well)

- **Top strip:** $|v_t^{(b)}|$ vs. time; optional contact marker. Reads as tempo/impact.
- **Main panel:** Nodes on the x-axis by activation time t_i and on the y-axis by anatomy grouping; draw only the path P^*. Edge thickness $\propto w_{ij}$; text labels show w_{ij} and lag in ms.

This layout lets the eye scan left-to-right to follow timing while the vertical anatomy groups preserve body semantics.

11 What we learn (and what we do not)

- *Ordering:* Strong, *causal-looking* (but not causal proof) sequencing from ground to bat.
- *Timing:* Quantified lead–lag in ms between segments.
- *Technique comparison:* Robust edges across clips suggest core mechanics; shot-specific edges (e.g. early trunk for hooks) reveal style.
- *Quality control:* Low w_{ij} or reversed lags flag extraction issues or atypical mechanics.

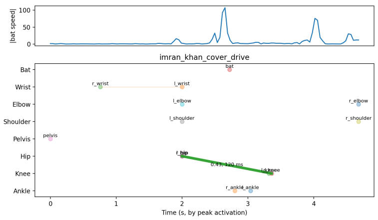
Limitations: monocular 2D misses out-of-plane components; PCA projection mitigates but cannot eliminate this. Occlusions may bias lags; we suppress low-confidence frames and smooth. Cross-correlation symmetry is broken by explicit temporal ordering and significance filtering, but causality remains an inference.

12 Algorithmic summary (at-a-glance)

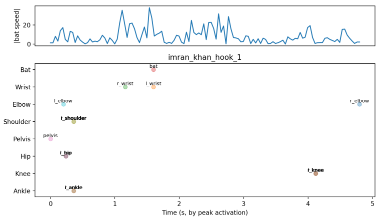
1. Detect keypoints $\{p_{i,t}\}$ and bat tip $\{b_t\}$; center at pelvis and scale by median shoulder width.
2. Smooth; compute $\dot{p}_{i,t}$ and $\dot{b}_t$; estimate swing axis u by PCA on $\{\hat{b}_t\}$; project to $v_{i,t}$ and $v_t^{(b)}$.
3. For all pairs (i, j), compute $\rho_{ij}(\tau)$, keep (c_{ij}, τ_{ij}^*).
4. Compute activations t_i; build DAG edges using temporal order, $\tau_{ij}^* > 0$, thresholds, and significance via N_{eff} and Fisher-z.
5. Extract P^* (maximum-weight source-to-bat chain).
6. (Optional) DTW-align clips by $|v_t^{(b)}|$ to aggregate statistics.

13 Recommended defaults (practical, easy to tune)

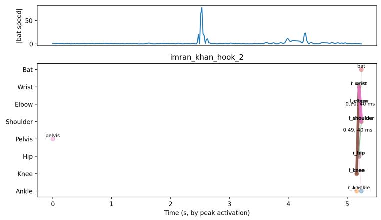
Parameter	Typical setting
Frame rate f	50–120 Hz (use native)
S-G filter	window $w \in [9, 21]$ frames, order $m \in \{2, 3\}$
Lag search	$\tau_{\max}$ s.t. $1000\,\tau_{\max}/f \in [120, 220]$ ms
Edge threshold	$\theta \in [0.35, 0.50]$
Significance	$\alpha = 0.05$ with N_{eff} and Fisher-z
DTW band	Sakoe–Chiba $\pm 10\%$ of sequence length



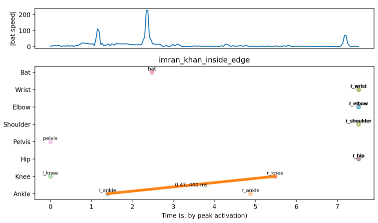
(a) Cover drive



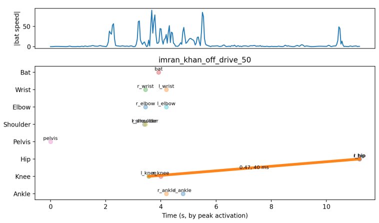
(b) Hook 1



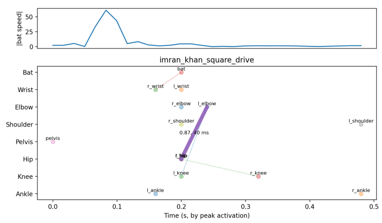
(c) Hook 2



(d) Inside edge



(e) Off drive



(f) Square drive

Figure 1: **Correlation–path summaries for six shots.** Each panel shows (top) bat–tip speed projected onto its dominant in–plane axis and (bottom) a directed, maximum–weight *correlation path* among joints (ankle→knee→hip→shoulder→elbow→wrist) ending at the bat tip when supported by the data. Nodes are placed at their individual peak–activation times on the horizontal axis; bold edges annotate correlation strength and time lag. The path condenses the kinetic chain into one interpretable polyline.

14 Notes for first-year grad-level readers

Every equation here is a standard building block: PCA for a dominant direction, central differences for velocities, and cross-correlation for lead–lag. The only slightly advanced piece is the N_{eff} correction, which is a compact way to avoid overstating significance when time series are autocorrelated. If you can compute means, variances, and a few dot products, you can implement the whole pipeline.

How to read the plots. For each joint and the bat tip we compute velocities along the swing axis (from PCA of bat motion). Directed edges represent statistically significant, time–lagged correlations (leader fires earlier and explains follower motion). The boldest polyline is the maximum–weight path from the lower body to the bat.

Shot–specific commentary.

1. **Cover drive** (fig:corrpath-cover): The most reliable coupling is lower–body—a right–hip→right–knee link ($\rho \approx 0.43$, $\sim$120 ms lead). Upper–chain links to the bat are sub–threshold, likely due to camera alignment and brief bat occlusion around impact.

2. **Hook 1** (fig:corrpath-hook1): Wrist crossing and bat foreshortening weakened tracking; we used a pose–based bat–tip fallback. The chain is mostly early lower–body and shoulder activity with no single dominant path to the bat in the allowed lag window.

3. **Hook 2** (fig:corrpath-hook2): A clean kinetic chain compressed near impact: ankle/knee/hip lead into shoulder, elbow, wrist, and bat. The shoulder→bat handoff is tight ($\rho \approx 0.51$, $\sim$40 ms), matching a compact hook.

4. **Inside edge** (fig:corrpath-insideedge): Footwork dominates: a long left–ankle→right–knee link ($\rho \approx 0.48$, $\sim$400 ms). The bat event is late and brief; windowing around the bat–speed peak would down–weight early footwork and expose upper–chain couplings.

5. **Off drive** (fig:corrpath-offdrive): Strong knee→hip coupling ($\rho \approx 0.48$, $\sim$40 ms) during re–load. Multiple bat–speed peaks suggest preparatory movements and follow–through; a bat–centric window will emphasize shoulder→wrist→bat.

6. **Square drive** (fig:corrpath-square): A compact release: hip→bat is very strong ($\rho \approx 0.85$, $\sim$40 ms) with supportive wrist→bat links. This panel is an exemplar template for drive mechanics.

Takeaways. Across shots we recover (i) lower–body initiation, (ii) trunk transfer, and (iii) upper–limb release. Hooks compress timing near impact; drives show clearer lower→upper sequencing. When paths do not reach the bat, the limiting factor is perception (occlusion/plane mismatch) rather than the method; a bat–centric time window and stronger bat tracking resolve this.