

# Pitch Effectiveness: A Physics- and Perception-Based Generative Model

A joint data-generating process for pitch outcomes, with Stuff and Location as functionals

Dr. Scott Spencer  
Columbia University (Faculty)  
scott.spencer@columbia.edu

December 8, 2024

## Table of contents

Abstract .....	1
1 Introduction .....	2
2 Units and constants .....	2
3 Data and notation .....	3
4 The joint likelihood .....	4
5 The outcome-chain links .....	5
5.1 Swing — perceived location, so flight enters through $\epsilon^{\text{loc}}$ .....	5
5.2 Called strike   take — true location, the umpire process .....	6
5.3 Whiff / foul / fair   swing — the core stuff node .....	6
5.4 Contact quality   fair — oblique-collision geometry .....	6
6 Physics forward map .....	7
6.1 Trajectory .....	7
6.2 Aerodynamics and induced movement .....	7
6.3 The expectation anchor .....	8
7 Eye-frame perception .....	8
8 Valuation .....	11
9 Hierarchy, priors, and the estimand .....	12
10 Reporting: Stuff, Location, and their interaction .....	12
11 A minimal version and a staging path .....	13
12 Validation, limitations, and discussion .....	13
13 References .....	14

## Abstract

Public pitch-quality models regress flight features on run value with gradient-boosted trees, producing point estimates with no data-generating process. We specify a Bayesian generative alternative: a per-pitch outcome chain — swing or take, called strike, whiff, foul, fair contact and its quality — whose conditional probabilities are driven by closed-form pitch physics and a mechanistic model of batter perception. Deception is defined as the deviation of the pitch from the batter’s anchored expectation of ordinary flight, projected onto the batter’s line of sight, where motion in depth is poorly discriminated; the platoon split then emerges from eye geometry rather than entering as a rule. Run value is a derived functional over a coupled run-expectancy model, so what are commonly called “Stuff” and “Location” become reportable posterior functionals of one process — with credible intervals — rather than separate additive surfaces. The physics is closed-form throughout and estimation is standard HMC; because the response maps are physical, the

model calibrated on professional data transports to amateur populations by re-estimating only the batter layer. A minimal single-node version is immediately fittable.

## 1 Introduction

We want a measure of pitch effectiveness that explains *why* a pitch succeeds, behaves sensibly outside the data it was trained on, and reports honest uncertainty. The common approach — the public “Stuff+” family spanning FanGraphs’ Stuff+ and Pitching+ (McGrattan 2023), PitchingBot (Grove 2023), Driveline’s model (Baseball 2021; Lambert 2024), and tjStuff+ (Nestico 2024) — pours pitch-flight features into a gradient-boosted tree (Chen and Guestrin 2016) and reads out a run-value number. That object has no data-generating process: it is a conditional-expectation surface fit with a point-prediction loss, it cannot extrapolate beyond its training support, it treats velocity, spin, and movement as independent inputs when movement is the aerodynamic *consequence* of release speed and spin, it returns no uncertainty, and its target confounds the batter, count, park, and umpire with the pitch. (Of the family, PitchingBot comes closest to a process model — it factorizes outcomes into swing, whiff, and contact stages (Grove 2023) — though its inputs remain unconstrained features and its outputs point estimates.)

This paper specifies the alternative: a single Bayesian generative model of the chain a pitch actually goes through. Given the pitch’s flight and how the batter *sees* it, the batter swings or takes; a take is called a ball or strike by the umpire; a swing misses, fouls, or makes fair contact, whose quality maps to a batted-ball outcome. Every one of these events advances the count or base-out state, and the run value of the pitch is the resulting change in expected runs to the end of the half-inning — a ball, a called strike, and a whiff are valued directly by their count transition; fair contact is valued through its batted-ball quality. Location and flight are modeled **jointly** — a batter does not experience them separately — so what is commonly called “Stuff” and “Location” become reportable functionals of one coherent process rather than two additive surfaces (Section 10). Two facts make this practical rather than aspirational: the physics is closed-form from the release-state measurement, so no trajectory solver runs inside the sampler; and the model’s sophistication lives entirely in those deterministic maps, leaving every stochastic node an ordinary, well-understood density — logistic for the discrete events, wrapped or skewed Gaussian for the batted-ball channels, negative binomial for runs — the standard building blocks that sample efficiently under HMC.

The paper proceeds top-down. Section 4 states the joint likelihood and Section 5 the four conditional links that carry the substance; the deterministic maps that feed them follow — trajectory and aerodynamics in Section 6, the batter’s expectation and eye-frame perception in Section 7. Section 8 couples the chain to run values, Section 9 gives the pooling structure and the estimand, and Section 10 the reporting decomposition. Section 11 then extracts a minimal fittable version and a staging path, and Section 12 collects the validation targets and limitations. Readers who prefer the generative order can read Section 6–Section 7 before Section 4–Section 5.

## 2 Units and constants

We work in SI throughout, so the aerodynamic and collision equations take their natural form and estimated coefficients are physically interpretable. Constants: gravitational acceleration  $g = 9.81 \text{ m/s}^2$ ; ball radius  $R = 0.0364 \text{ m}$ , mass  $m = 0.145 \text{ kg}$ ; air density  $\rho \approx 1.23 \text{ kg/m}^3$ ; cross-section  $A = \pi R^2$ ; and the aerodynamic prefactor  $K = \frac{1}{2} \rho A / m \approx 0.0177 \text{ m}^{-1}$ . (Implementation kernels may operate natively in field units — the collision kernel uses mph and inches — with conversions confined to their interfaces.)

Field coordinates follow the tracking convention:  $x$  lateral,  $y$  from the plate toward the pitcher (depth),  $z$  vertical (up), with  $\hat{z}$  the upward unit vector; the front of the plate sits at  $y_{\text{plate}} = 0.43 \text{ m}$  and the front of the rubber at  $18.44 \text{ m}$ .

### 3 Data and notation

Each row is a pitch  $i$ , thrown by pitcher  $j[i]$  to batter  $b[i]$ , with umpire  $u[i]$ , catcher  $c[i]$ , and count / base-out state  $s[i]$ . Bracketed indices are index maps in the multilevel-model convention (Gelman and Hill 2007):  $b[i]$  is the batter facing pitch  $i$ , so a batter-level parameter enters the likelihood as, e.g.,  $\alpha_{b[i]}$ , mirroring its implementation (`alpha[batter[i]]`). The measured **release state** is the nine-parameter (9P) description – position, velocity, and acceleration – from which all plate-state quantities are derived by integration (Section 6). Position and velocity evolve along the flight, so their release values carry the 0 subscript; the 9P fit treats acceleration as a single constant vector per pitch, so  $\mathbf{a}$  carries none. (Physically the aerodynamic acceleration is not constant – drag and Magnus scale with the decaying speed – constant- $\mathbf{a}$  is the tracking system’s least-squares approximation; see Section 6.1.)

One data convention matters here: tracking systems report the 9P snapshot at a standardized  $y = 50$  ft (15.24 m) from the plate – a PITCHf/x-era convention kept by TrackMan and by the public Hawk-Eye/Statcast schema (Kagan and Nathan 2017) – not at the hand. True release depth is located by the extension,  $r_y = 18.44$  m –  $e$  (front of the rubber minus extension), and  $\mathbf{r}_0, \mathbf{v}_0$  are reconstructed there by running Equation 9 backward from the snapshot. We take  $t = 0$  at this reconstructed release, so  $t^f$  is the batter’s true reaction budget – which is why  $e$  is a load-bearing input rather than a redundant column: at equal speed, longer extension shortens  $t^f$  (the “perceived velocity” effect).

When a feed omits  $e$ , two fallbacks preserve the time origin. If release height and side are recorded – they usually are, since the systems measure release directly – release depth is recovered by back-solving the trajectory for the time at which the back-extrapolated height matches the measured release height,  $\mathbf{B}_z(t^*) = z^{\text{rel}}$ ; the trajectory alone carries no information about where on the quadratic release occurred, so this anchor is essential. Failing that,  $e$  becomes a latent with a pooled prior (population mean, per-pitcher partial pooling), so the few milliseconds it contributes to  $t^f$  enter as honest uncertainty rather than a silent bias. We additionally observe the realized outcome path and, on fair contact, the batted ball.

Symbol	Meaning	Unit
$\mathbf{r}_0 = (r_x, r_y, r_z)_i$	release position	m
$\mathbf{v}_0 = (v_x, v_y, v_z)_i$	release velocity; $v = \ \mathbf{v}_0\ $ release speed	m/s
$\mathbf{a} = (a_x, a_y, a_z)_i$	acceleration, constant over the flight in the 9P fit	m/s <sup>2</sup>
$\mathbf{B}(t)$	ball position $t$ seconds after release ( $B$ for ball; derived, Section 6.1)	m
$\omega, \phi$	spin rate, spin-axis angle	rad/s, rad
$e$	release extension from the front of the rubber; locates release depth $r_y = 18.44$ m – $e$	m
$a_v, a_h$	vertical / horizontal approach angle at the plate	rad
$t_i^f$	time of flight, release to plate	s
$\mathbf{x}_i^{\text{Pl}} = (x_{\text{plate}}, z_{\text{plate}})$	plate location	m

Symbol	Meaning	Unit
$h_b$	batter height ( $\rightarrow$ ANSUR-II anatomy: eye, shoulder, reach)	m
$\mathbb{I}_i^{\text{sw}}, \mathbb{I}_i^{\text{cs}}, \mathbb{I}_i^{\text{wh}}, \mathbb{I}_i^{\text{fl}}, \mathbb{I}_i^{\text{fr}}$	event indicators: swing; called strike take; whiff / foul / fair  swing	—
$v_f, LA, \theta_{\text{spray}}$	exit velocity, launch angle, spray angle	m/s, rad, rad
$o_i$	terminal batted-ball outcome (out, 1B, 2B, 3B, HR)	—

Derived perceptual quantities that enter the outcome-chain probabilities (Section 5) as predictors — the line of sight  $\hat{s}(t)$  and the perceptual errors  $\varepsilon_i^{\text{loc}}, \varepsilon_i^{\text{time}}$  — are defined in Section 6–Section 7. Throughout,  $\Lambda(\cdot)$  is the logistic CDF; a hat denotes a unit vector ( $\hat{z}, \hat{v}, \hat{\omega}, \hat{s}$ ), with two exceptions:  $\hat{\mathbf{B}}$ , the batter’s expected trajectory (Section 6.3), and  $\hat{\mathbf{a}}_M$ , the Magnus-predicted acceleration.

## 4 The joint likelihood

The data factor into the per-pitch **outcome chain**, the **valuation** likelihood on observed runs, and the **anatomy** likelihood that ties body dimensions to batter height, all sharing the parameter vector  $\Theta$ :

$$\mathcal{L}(\Theta) = \underbrace{\prod_{i=1}^N \mathcal{L}_i^{\text{chain}}}_{\text{response chain}} \times \underbrace{\prod_m \mathcal{L}_m^{\text{runs}}}_{\text{valuation}} \times \underbrace{\mathcal{L}^{\text{anat}}}_{\text{anatomy}}. \quad (1)$$

Each pitch contributes one factor along its realized path, with a contact-quality density  $f_i^{\text{bb}}$  appended to the fair-ball branch:

$$\begin{aligned} \mathcal{L}_i^{\text{chain}} = & \underbrace{(\pi_i^{\text{sw}})^{\mathbb{I}_i^{\text{sw}}} (1 - \pi_i^{\text{sw}})^{1 - \mathbb{I}_i^{\text{sw}}}}_{\text{swing / take}} \cdot \underbrace{[(\pi_i^{\text{cs}})^{\mathbb{I}_i^{\text{cs}}} (1 - \pi_i^{\text{cs}})^{1 - \mathbb{I}_i^{\text{cs}}]}_{\text{called strike | take}}]^{1 - \mathbb{I}_i^{\text{sw}}} \\ & \cdot \underbrace{[(\pi_i^{\text{wh}})^{\mathbb{I}_i^{\text{wh}}} (\pi_i^{\text{fl}})^{\mathbb{I}_i^{\text{fl}}} (\pi_i^{\text{fr}} f_i^{\text{bb}})^{\mathbb{I}_i^{\text{fr}}}]^{\mathbb{I}_i^{\text{sw}}}}_{\text{whiff / foul / fair (+ quality) | swing}} \end{aligned} \quad (2)$$

where  $f_i^{\text{bb}} = f^{\text{bb}}(v_{f,i}, LA_i, \theta_{\text{spray},i} \mid \cdot)$  is the batted-ball density (Section 5.4). Figure 1 diagrams the chain. The entire Stuff-vs-Location question reduces to *which flight features enter which  $\pi$ , and how*; the four links below are the substance of the model.

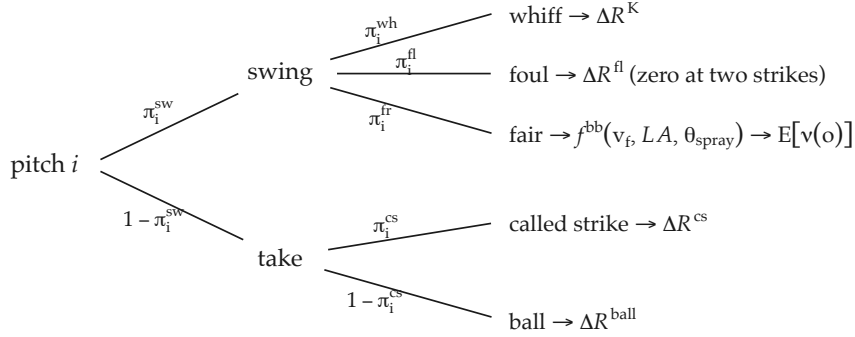


Figure 1: The per-pitch outcome chain of Equation 2. Each pitch follows exactly one path; the branch probabilities are the four links of Section 5, and every leaf is a count or base-out transition valued in runs through Section 8. Stuff-versus-location is a question about which flight features enter which branch probability.

## 5 The outcome-chain links

The perceptual errors  $\varepsilon_i^{\text{loc}}$  (a plate-plane mislocation) and  $\varepsilon_i^{\text{time}}$  (a timing error) are deterministic functions of the trajectory and the batter’s eye (Section 7).

### 5.1 Swing — perceived location, so flight enters through $\varepsilon^{\text{loc}}$

$$\pi_i^{\text{sw}} = a_{s[i]}^{\text{cnt}} a_{b[i]}^{\text{bat}} \Lambda_{\text{bilog}} \left( \mathbf{x}_i^{\text{pl}} + \varepsilon_i^{\text{loc}}; \text{edges} ( b[i] ), \sigma_{b[i]}^{\text{sw}} U_i^{\text{sw}}, \rho^{\text{sw}} \right). \quad (3)$$

The batter swings according to the *perceived* location  $\mathbf{x}^{\text{pl}} + \varepsilon^{\text{loc}}$  read against a batter-specific swing zone, modeled as a **bivariate logistic**: a soft rectangle whose *edges* mark the decision boundary, whose edge *sharpness*  $\sigma^{\text{sw}}$  sets how abruptly swing probability falls across that boundary, and whose spatial correlation  $\rho^{\text{sw}} \in (-1, 1)$  couples the width and height judgments — the two directions are not independent perceptual channels, and the independent product  $(1 - \Lambda_x)(1 - \Lambda_z)$  is recovered as the  $\rho^{\text{sw}} = 0$  special case.

Nor is the sharpness uniform over the zone: localization degrades with eccentricity from the perceiver’s eye, and one-sidedly — an eye placed high and to one side reads the near region almost without error while the far region blurs. A stepwise rule would be the natural description, but HMC wants a continuous, differentiable surface, so each direction is gated with a softplus:

$$U(\mathbf{x}; \mathbf{E}, \lambda) = 1 + \lambda \log \left( 1 + \sqrt{s_\iota \left( \tilde{x} - \tilde{x}^E \right)^2 + s_\iota \left( z^E - z \right)^2} \right), \quad (4)$$

where  $\mathbf{E}$  is the perceiver’s eye,  $\tilde{x}$  the handedness-mirrored lateral coordinate oriented toward the outside edge,  $s_\iota(\cdot)$  a softplus with smoothness  $\iota$  — each term contributes only once the ball is beyond the eye in its direction — and  $\lambda$  a per-perceiver amplitude;  $\iota$  and the  $\lambda$ ’s are estimated (lengths inside the logarithm are in meters — the implicit unit scale is absorbed into  $\lambda$  and  $\iota$ ). For the batter,  $U_i^{\text{sw}} = U \left( \mathbf{x}_i^{\text{pl}} + \varepsilon_i^{\text{loc}}; \mathbf{E}_{b[i]}, \lambda^{\text{sw}} \right)$  with the eye of Section 7, which sits above the zone on the batter’s own side: localization is sharp up-and-in and degrades toward low-and-away. The whole is scaled by count- and batter-level propensities  $a^{\text{cnt}}, a^{\text{bat}} \in (0, 1]$ . Because  $\varepsilon^{\text{loc}}$  is a deterministic function of induced movement, swing probability depends on location **and** flight jointly — chase emerges from deceptive movement rather than a separate term (Gray 2013). The bivariate-logistic form is a signal-detection model of the swing decision: the half-widths  $w^{\text{sw}}$  are the batter’s criteria and  $1/(\sigma^{\text{sw}} U^{\text{sw}})$  its sensitivity  $d'$  — highest near the batter’s eye, falling toward low-and-away. Better hitters carry higher  $d'$  through more sharply and *asymmetrically* placed

criteria – covering the lower zone more, mirroring the umpire’s own reluctance to call high strikes – rather than through greater patience, and  $d'$  tracks batting performance; so  $w^{\text{sw}}$  (with the upper and lower  $z$  edges allowed to differ) and  $\sigma^{\text{sw}}$  are batter-level hierarchical parameters (Gray 2010).

## 5.2 Called strike | take – true location, the umpire process

$$\pi_i^{\text{cs}} = \Lambda_{\text{bilog}} \left( \mathbf{x}_i^{\text{pl}}; \text{edges} ( u[i], s[i], c[i] ), \sigma^* U_i^{\text{cs}}, \rho \right). \quad (5)$$

The same bivariate-logistic zone as the swing decision (Equation 3), read from the other viewpoint. The umpire’s task is retrospective – judging where the ball crossed rather than predicting where it will be – so he reads the *true* location, and the flight-induced perceptual errors of Section 7 do not enter; instead the zone *edges* shift by umpire, count, and catcher (framing), the directions couple through his own spatial correlation  $\rho$ , and the edge sharpness carries the same eccentricity inflation (Equation 4), anchored now at the umpire’s eye. He sets up over the catcher’s inside shoulder, eye near the **upper-inside corner** of the zone – so his eye, like the batter’s, mirrors with batter handedness – and  $U_i^{\text{cs}} = U ( \mathbf{x}_i^{\text{pl}}; \mathbf{E}_u, \lambda^{\text{cs}} )$ : a ball above or inside eye level is read against the zone almost without error, while uncertainty grows as the ball moves low and away. Both eyes thus sit on the batter’s side of the plate and above the middle of the zone, so the two perceivers blur the *same* corner – low-and-away is doubly fogged, for the call and for the swing decision alike, one reason it is the corner pitchers target. Conditioning on these factors – rather than smoothing a marginal spline over location, whose apparent shape is largely an artifact of where pitches are thrown – yields a sharper, more honest decision surface and absorbs any expanded or contracted zone natively, with uncertainty.

## 5.3 Whiff / foul / fair | swing – the core stuff node

Contact probability is driven by the timing error, the reaction budget, and reach:

$$\pi_i^{\text{con}} = 1 - \pi_i^{\text{wh}}, \quad \text{logit } \pi_i^{\text{con}} = \alpha_{b[i]} - b_\tau | \varepsilon_i^{\text{time}} | - b_r \delta_i^{\text{reach}} - b_v / t_i^f. \quad (6)$$

Here  $t_i^f$  is the reaction budget (less time to react  $\Rightarrow$  more whiffs),  $\delta_i^{\text{reach}}$  is the shoulder-to-ball reach deficit (anatomy), and  $\varepsilon_i^{\text{time}}$  is the timing error from misjudged motion in depth (Section 7), the dominant driver. Behavioral data fix the scale: the timing margin for error is on the order of 9 ms, yet a mis-set speed expectation alone can produce tens to over 100 ms of swing-timing error – exceeding the true fast-slow arrival gap – so the prior-driven part of  $\varepsilon_i^{\text{time}}$  (sequencing and count; Section 6.3) can dominate the geometric part (Gray 2002b, 2002a). Conditional on contact, the fair/foul split is a second logit on the same timing error and the contact offset (Section 5.4),  $\log ( \pi_i^{\text{fl}} / \pi_i^{\text{fr}} ) = \eta_{b[i]} + \eta_\tau | \varepsilon_i^{\text{time}} | + \eta_D \| \mathbf{D}_i \|$ , with  $\pi_i^{\text{fl}} + \pi_i^{\text{fr}} = \pi_i^{\text{con}}$  so the three swing outcomes sum to one.

## 5.4 Contact quality | fair – oblique-collision geometry

The batter’s swing is summarized by swing speed  $v_{\text{bat},b}$  and attack angles ( $AA_b, AA_{h,b}$ ) – batter-level centers whose per-swing scatter is marginalized rather than sampled. The timing and space errors displace the contact point on the bat from the sweet spot; write the contact offset

$$\begin{aligned} D_{v,i} &= \mu_{D_v} + \beta_\tau \varepsilon_i^{\text{time}} + \beta_a (a_{v,i} - AA_b), \\ D_{h,i} &= \mu_{D_h} + \beta_h (a_{h,i} - AA_{h,b}) + \beta_\ell \varepsilon_{x,i}^{\text{loc}}, \end{aligned} \quad (7)$$

so the offset grows when timing is wrong, when the bat is mislocated, or when the pitch’s approach angles ( $a_v, a_h$ ) mismatch the bat’s attack angles. The oblique-impact map  $\mathcal{C}$  – normal restitution  $e_A$ ; tangential restitution  $e_T$  in the grip regime, capped in slip by Coulomb friction  $\mu_{\text{fric}}$ , the two blended continuously – converts the incoming ball state and bat state into the outgoing batted ball (Cross and Nathan 2006; Stronge 2018):

$$(v_f, LA, \theta_{\text{spray}})_i = \mathcal{C}(\mathbf{v}_{1,i}, \boldsymbol{\omega}_{1,i}, v_{\text{bat},b}, AA_b, AA_{h,b}, \mathbf{D}_i), \quad (8)$$

where the incoming ball velocity  $\mathbf{v}_{1,i}$  and spin  $\boldsymbol{\omega}_{1,i}$  at contact come from the trajectory (Section 6) – including the spin’s surface velocity at the contact point, the term whose omission historically forced ad-hoc spin-attenuation factors: larger  $\|\mathbf{D}_i\|$  tilts the contact normal off the line of approach and lowers the exit velocity relative to the swing speed delivered (the *smash factor*,  $v_f/v_{\text{bat}}$ ).  $\mathcal{C}$  is evaluated in a contact frame ( $\hat{\mathbf{e}}_x$  toward the pitcher,  $\hat{\mathbf{e}}_y$  up,  $\hat{\mathbf{e}}_z$  toward third base for a right-handed batter), the tracking-frame ball state rotated at the interface; the contact normal is a line-of-centers construction anchored on the approach direction – not the bat axis – tilted by  $\mathbf{D}$ . The observation density factorizes across channels,

$$v_f^{\text{obs}} \sim \text{SkewT}(v_f, \sigma_{v_f}, \dots), \quad LA^{\text{obs}} \sim \mathcal{N}_{\text{wrap}}(LA, \sigma_{LA}), \quad \theta_{\text{spray}}^{\text{obs}} \sim \mathcal{N}_{\text{wrap}}(\theta_{\text{spray}}, \sigma_{\theta}).$$

Gaussian channels with wrapped residuals for the angles suffice on a quality-filtered sample; here the batted-ball population is deliberately unfiltered – weak contact is precisely the signal that stuff suppresses – so the exit-velocity channel is skewed and heavy-tailed to represent mishits rather than filter them.  $f_i^{\text{bb}}$  in Equation 2 is this joint density, and the batted ball maps to a terminal outcome,  $\Pr(o | v_f, LA, \theta_{\text{spray}})$ , feeding the valuation (Section 8).

## 6 Physics forward map

### 6.1 Trajectory

Plate-state quantities are not separately measured; they are derived by integrating the release state forward in time. Write  $\mathbf{B}(t)$  for the **ball position**  $t$  seconds after release –  $B$  for ball, kept distinct from the plate location  $\mathbf{x}^{\text{pl}}$  and the batter index  $b$ . Under the 9P fit’s constant acceleration, integrating twice gives  $\mathbf{B}$  as a quadratic in time, and setting its depth component to the front of the plate gives the time of flight  $t^f$  as that quadratic’s root:

$$\mathbf{B}(t) = \mathbf{r}_0 + \mathbf{v}_0 t + \frac{1}{2} \mathbf{a} t^2, \quad \mathbf{B}_y(t^f) = y_{\text{plate}} \Rightarrow t^f, \quad (9)$$

giving plate location  $\mathbf{x}_i^{\text{pl}} = (\mathbf{B}_x, \mathbf{B}_z)(t^f)$ , the approach angles  $a_v, a_h$  from  $\dot{\mathbf{B}}(t^f)$ , and the reaction budget  $t_i^f$ . (For exposition we use the constant-acceleration form; the same kinematics with the velocity-dependent drag and Magnus accelerations below give the exact path at little additional cost.)

### 6.2 Aerodynamics and induced movement

A pitch deviates from a gravity-only path because of aerodynamic forces, chiefly the Magnus force from the interaction of spin and airflow (Nathan 2008). The Magnus acceleration is

$$\mathbf{a}_M = K C_L v^2 \frac{\hat{\boldsymbol{\omega}} \times \hat{\mathbf{v}}}{\|\hat{\boldsymbol{\omega}} \times \hat{\mathbf{v}}\|}, \quad C_L(S) = \frac{S}{0.40 + 2.32 S}, \quad S = \frac{R \omega_{\perp}}{v}, \quad \omega_{\perp} = \lambda \omega, \quad (10)$$

where the lift coefficient  $C_L$  depends on the spin factor  $S$  through the transverse (*active*) spin  $\omega_{\perp}$ : the active-spin fraction  $\lambda \in (0, 1]$  is the share of total spin perpendicular to  $\mathbf{v}$  (parallel gyrospin produces no deflection), and it enters inside the nonlinear lift curve rather than as an outside multiplier. The spin axis  $\phi$  is a single angle with  $\phi = 0$  at pure backspin (Magnus force pointing up) (Nathan 2018). Removing gravity from the measured acceleration gives the **aerodynamic acceleration**

$$\mathbf{a}^{\text{aero}} = \mathbf{a} + g \hat{\mathbf{z}} = \underbrace{\hat{\mathbf{a}}_M(v, \omega, \phi, \lambda)}_{\text{Magnus}} + \underbrace{\tilde{\mathbf{a}}}_{\text{non-Magnus residual}}, \quad (11)$$

the residual  $\tilde{\mathbf{a}}$  capturing seam-shifted wake, grip, and departures of the empirical  $C_L$  from Equation 10.

The trajectory consumes  $\mathbf{a}^{\text{aero}}$  whole, and the deviation  $\mathbf{d}(t)$  of Section 6.3 uses only its split along versus across the velocity – not the Magnus/residual split, which does no work there. It earns its place three other ways. First, it supplies the spin sense  $q^\omega$  that drives the perceptual spin illusion (Section 7). Second, it is the **identification instrument** for that illusion: two pitches can carry the *same* measured deviation  $\mathbf{d}$  with very different spin signatures – one Magnus-efficient (large  $\omega_\perp$ ), one seam-shifted-wake-dominant (large  $\tilde{\mathbf{a}}$ , small  $q^\omega$ ) – and any systematic outcome gap between such pairs is perceptual, not kinematic. Third, it carries the transport: where measured accelerations are noisy or missing (amateur tracking), movement is reconstructed from  $(v, \omega, \phi, \lambda)$  through Equation 10 with  $\tilde{\mathbf{a}}$  as the honestly uncertain remainder.

### 6.3 The expectation anchor

Humans carry a hardwired internal model of gravity used for interception timing (McIntyre et al. 2001), and an experienced batter is calibrated to how ordinary pitches fly – which includes the drag that every pitch shares. We therefore take the batter’s *expected* trajectory to be the observed early motion continued under the full ordinary physics, gravity plus drag, from a pickup time  $t_0$ :

$$\hat{\mathbf{B}}(t) = \mathbf{B}(t_0) + \dot{\mathbf{B}}(t_0) (t - t_0) + \frac{1}{2} [ -g \hat{\mathbf{z}} + \mathbf{a}^{\text{drag}} ] (t - t_0)^2, \quad \mathbf{d}(t) = \mathbf{B}(t) - \hat{\mathbf{B}}(t). \quad (12)$$

Here  $\mathbf{a}^{\text{drag}} = ( \mathbf{a}^{\text{aero}} \cdot \hat{\mathbf{v}}(t_0) ) \hat{\mathbf{v}}(t_0)$  is the along-path component of the aerodynamic acceleration at pickup. The deviation  $\mathbf{d}(t)$  is then exactly the *induced* movement: the double integral of the velocity-perpendicular (lift) remainder,  $\mathbf{a}^{\text{lift}} = \mathbf{a}^{\text{aero}} - \mathbf{a}^{\text{drag}}$ . This is the same induced movement that flight-based metrics already quantify – induced break is conventionally measured against a spinless ball *with* drag – now with a mechanism for *why* it matters: the anchor carries everything common to ordinary flight, charging the pitch only with its pitch-specific lift.

The anchor is a modeling choice with two leaner alternatives, and the behavioral evidence splits by channel. For crossing *height*, batters measured in simulators extrapolate at the ball’s current rate – a constant-velocity rule that ignores even gravity’s well-experienced drop, biasing estimates high when the ball is far and low when near (Gray 2002a). For interception *timing*, the internal gravity model is hardwired (McIntyre et al. 2001). So a constant-velocity anchor, a gravity-only anchor, and the full-physics anchor adopted here differ only in the  $(t - t_0)^2$  coefficient – none adds a parameter – and they are directly comparable on held-out predictive density; we take the full physics as the core and carry the other two as baselines. Batters also do not use the anchor in isolation: they integrate it with a situational expectation (count, arsenal) in proportion to each source’s reliability, leaning on the expectation when the ball is seen only briefly or arrives fast (Gray and Cañal-Bruland 2018). A **staged extension** formalizes this as a precision-weighted blend,

$$\hat{\mathbf{B}}^{\text{blend}}(t) = (1 - w_i) \hat{\mathbf{B}}(t) + w_i \hat{\mathbf{B}}^{\text{sit}}(t), \quad w_i = \Lambda ( c_0 - c_1 t_i^f ), \quad (13)$$

where  $\hat{\mathbf{B}}^{\text{sit}}$  is the count-conditional expected trajectory and the weight  $w_i$  on the situational expectation grows as the reaction budget shrinks – so the deception leverage of an off-expectation pitch rises with velocity. The core model sets  $w = 0$ : the flight anchor is parameter-free, requires no arsenal or sequence data, and therefore transports to amateur settings where those are unreliable.

## 7 Eye-frame perception

Place the batter’s eye  $\mathbf{E}$  from anatomy, handedness, and stance. The line of sight and the split of the induced deviation into **depth** vs **transverse** components are

$$\hat{\mathbf{s}}(t) = \frac{\mathbf{B}(t) - \mathbf{E}}{\|\mathbf{B}(t) - \mathbf{E}\|}, \quad d_{\parallel}(t) = \mathbf{d}(t) \cdot \hat{\mathbf{s}}(t), \quad \mathbf{d}_{\perp}(t) = \mathbf{d}(t) - d_{\parallel}(t) \hat{\mathbf{s}}(t). \quad (14)$$

The line of sight is instantaneous:  $\hat{\mathbf{s}}(t)$  rotates as the ball approaches – fastest near the plate, as the range collapses – so Equation 14 re-splits the deviation at every moment. Over the flight, the family of sight rays sweeps a surface with its apex at the eye. For the batter’s *expected* path  $\hat{\mathbf{B}}$  that surface is, to first order, exactly a plane (a line and an external point span one; the induced deviation is centimeters against a roughly 17 m flight). Call it the **gaze plane**: the plane containing the eye and the anchor trajectory – an earlier sketch of this construction, as a batter’s *plane of sight*, appears in Spencer (2023). Every expected sight ray – hence every instantaneous depth axis – lies in this plane, which gives the decomposition a planar reading: deviation **normal to the gaze plane** is pure bearing change, the signal the eye reads best; deviation **within the plane** divides between the radial (depth) direction and an in-plane, elevation-like transverse direction, the depth share growing as the ray turns oblique near the plate. Because the eye sits to the batter’s side, the gaze plane is tilted off the pitch’s vertical plane, so the same physical break loads these components differently by break direction and handedness – the geometry behind the platoon result below (Figure 2).

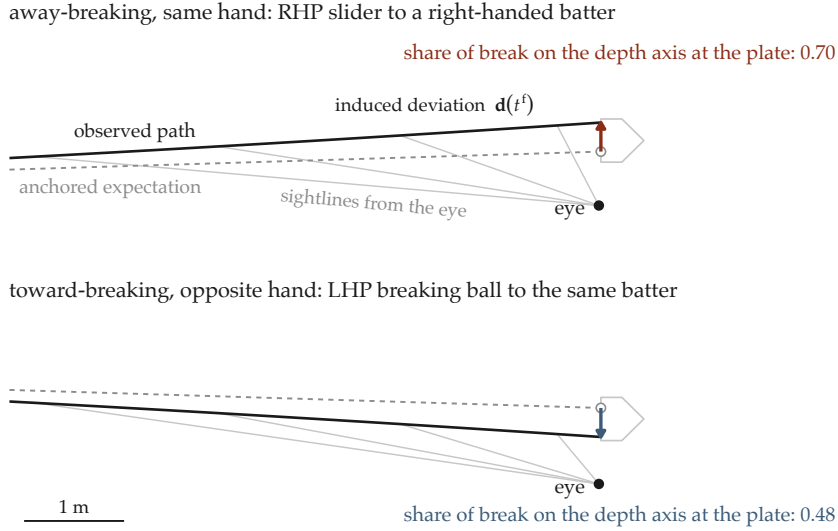


Figure 2: Bird’s-eye view of the final six metres of two mirrored breaking balls to a right-handed batter, computed from the constant-acceleration trajectory and drawn to scale (release speed 36.5 m/s; 0.28 m of horizontal break). The batter’s eye sits on the third-base side of the plate. The induced deviation  $\mathbf{d}(t^f)$  of the away-breaking pitch carries 0.70 of its length along the terminal sightline – depth, which the eye reads poorly – against 0.48 for the mirrored toward-breaking pitch located inside. The platoon split is this geometry, not a rule.

**Gated accumulation along the flight.** The eye resolves transverse (bearing) motion well and motion in depth poorly (Gray 2017b), and the batter can act on what he perceives only up to a commit point a fixed lead before arrival,  $t_c = t^f - \Delta$ ,  $\Delta \approx 175$  ms. The principled form of each perceptual error is therefore a *gated accumulation* of deviation along the flight. With the running projections  $u_{\parallel}(t) = \mathbf{d}(t) \cdot \hat{\mathbf{s}}(t)$  and  $\mathbf{u}_{\perp}(t) = \mathbf{d}_{\perp}(t)$  from Equation 14,

$$\varepsilon_i^{\text{time}} = \frac{1}{|\dot{r}(t^f)|} \int_{t_0}^{t^f} w_{\parallel}(t) \dot{u}_{\parallel}(t) dt, \quad \varepsilon_i^{\text{loc}} = \Pi_{xz} \int_{t_0}^{t^f} w_{\perp}(t) \dot{\mathbf{u}}_{\perp}(t) dt, \quad (15)$$

with  $\Pi_{xz}$  the projection to the plate plane and  $r(t) = \|\mathbf{B}(t) - \mathbf{E}\|$  the range to the eye. The gates encode what the batter can correct. Depth:  $w_{\parallel}(t) \equiv 1 - \text{motion-in-depth divergence}$  sits below its coarse discrimination thresholds (120–220 ms) for the whole flight (Gray and Regan 2006), so none of it is ever corrected. Transverse:  $w_{\perp}(t) = \Lambda((t - t_c)/s)$  – bearing change seen before commit is tracked out; only the post-commit accrual stands. Three consequences. With a hard gate the transverse error is the post-commit accrual  $\Pi_{xz}[\mathbf{d}_{\perp}(t^f) - \mathbf{d}_{\perp}(t_c)]$ . Because  $\Delta$  is a fixed *time*, the post-commit share grows as  $t^f$  shrinks ( $\Delta(2t^f - \Delta)/(t^f)^2$  under constant acceleration), so velocity mechanically steals correction time – a velocity-by-movement interaction the model *produces* rather than fits. And  $\Delta$  is estimable, with the observed commit lead as its prior. Computationally there is no differential equation to solve: every integrand is a closed-form function of the already-solved trajectory; the hard-gate version needs no quadrature at all, and the smooth gate needs a 5–10-node Gauss–Legendre rule. Figure 3 shows the depth-loading along the flight for the two pitches of Figure 2.

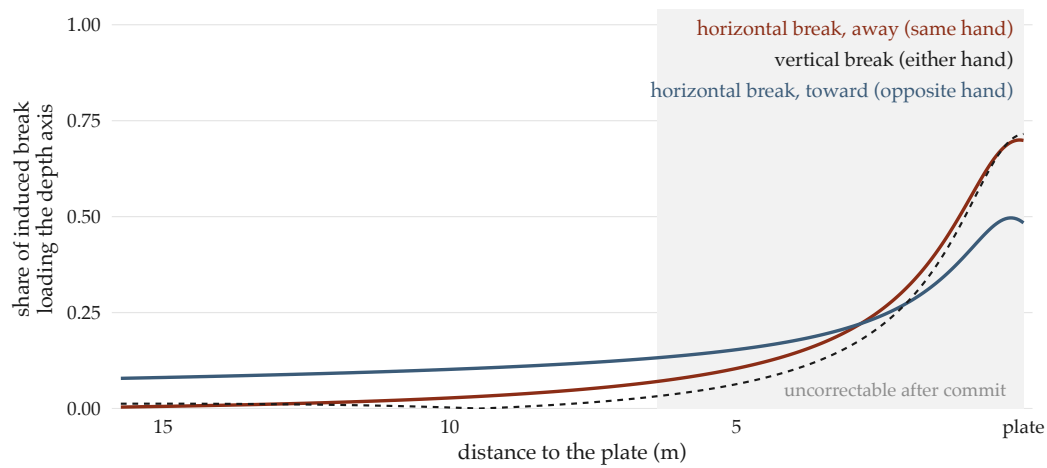


Figure 3: Share of the induced break loading the depth axis along the flight, for the two pitches of Figure 2. The share is small early – when bearing change is visible and correctable – and rises steeply inside the commit window (final  $\approx 175$  ms, shaded), more steeply for the away-breaking pitch. Vertical break (dashed) behaves like the away case because the ball drops below the eye late in flight. The velocity-by-movement interaction of Equation 15 is visible directly: a faster pitch compresses the same late rise into less correctable time.

**Simplification, and what it costs.** For the depth channel,  $w_{\parallel} \equiv 1$  makes Equation 15 telescope exactly to the endpoint (since  $\mathbf{d}(t_0) = \mathbf{0}$ ):  $\varepsilon^{\text{time}} = d_{\parallel}(t^f) / |\dot{r}(t^f)|$ , with no further approximation – the assumption lives in the gate, justified by the thresholds. For the transverse channel the core model replaces the post-commit accrual with a constant attenuation of the endpoint,  $\kappa_{\perp} \Pi_{xz} \mathbf{d}_{\perp}(t^f)$ . The information cost, in order: (i) the per-pitch post-commit share varies with  $t^f$ , and a single  $\kappa_{\perp}$  flattens it, so the velocity-by-movement interaction must be absorbed by link coefficients instead of emerging; (ii)  $\Delta$  is no longer estimable –  $\kappa_{\perp}$  stands in as the average uncorrected share; (iii) with full-path tracking, per-pitch accrual profiles (true late break) would differ and the endpoint would discard them – though under the 9P constant-acceleration fit that profile is common to all pitches by construction, so this third cost is currently zero and becomes real only with full-flight data. With those caveats, the working (core) form of the errors is:

$$\begin{aligned}
\varepsilon_i^{\text{loc}} &= \kappa_{\perp} \Pi_{xz} \mathbf{d}_{\perp}(t^f) + \kappa_{\omega z} q_i^{\omega} \hat{\mathbf{z}}, \\
\varepsilon_i^{\text{time}} &= \kappa_{\parallel} \frac{d_{\parallel}(t^f)}{|\dot{r}(t^f)|} + \delta_{\tau} + \kappa_{\omega\tau} q_i^{\omega}, \\
\kappa_{\parallel} &\approx 1, \quad \kappa_{\perp} < 1.
\end{aligned} \tag{16}$$

where the depth deviation  $d_{\parallel}$  converts to a time-of-arrival error through the closing speed  $\dot{r}$ . The transverse gain  $\kappa_{\perp}$  now reads as the average uncorrected share of transverse deviation; perceptual discrimination thresholds (Weber fractions  $\approx 0.1$ – $0.2$  in restrained-lab conditions rising to  $0.5$ – $0.8$  for batters in-stance, larger for amateurs) inform its prior.

The core model sets  $\delta_{\tau} = \kappa_{\omega z} = \kappa_{\omega\tau} = 0$ : the geometric channel —  $\mathbf{d}$  projected on the sightline — carries deception by itself. The added terms are **staged extensions**, psychophysically motivated (Gray and Regan 2006) but kept out of the core because they are nearly collinear with the geometric channel and must earn their identification. First, direction-of-motion-in-depth and time-to-passage are not discriminated independently — they are perceptually entangled, with a systematic bias toward perceiving approaching objects as arriving *sooner* than they do, and the absolute resolution of time-to-passage is coarse (discrimination thresholds of order  $120$ – $220$  ms, comparable to the entire flight): hence the constant bias  $\delta_{\tau}$ , and hence also  $\kappa_{\parallel} \approx 1$  — depth/timing is essentially uncorrectable. Second, **ball spin biases the perceived direction of motion in depth independently of the actual deflection**: backspin shifts the perceived trajectory toward the sky, topspin toward the ground (the “rising fastball” and extra-drop illusions) (Gray and Regan 2006; Gray 2017a). This acts even when the physical deviation  $\mathbf{d}$  is held fixed, entering through  $q_i^{\omega} \in [-1, 1]$ , the vertical sense of the transverse spin — under the spin-axis convention of Section 6.2,  $q^{\omega} = \cos \phi$ :  $+1$  pure backspin/ride,  $-1$  pure topspin/drop,  $0$  pure sidespin — with perceptual gains  $\kappa_{\omega z}, \kappa_{\omega\tau}$ .

**Identifying the extensions.** Because  $q^{\omega}$  correlates with the Magnus lift already inside  $\mathbf{d}$ , the illusion gains are estimated under zero-centered priors: “perception = physics” is the null, and posterior mass away from zero is the evidence that the illusion adds anything beyond the anchor-extrapolation error. Three structural contrasts can move them off that null. (i) *Cross-axis*: horizontal break carries the geometric channel with  $q^{\omega} \approx 0$ , pinning  $\kappa_{\perp}$  with the illusion switched off, so the illusion gains are identified from the vertical residual. (ii) *Cross-outcome*:  $\kappa_{\omega z}$  loads on the swing/chase node and  $\kappa_{\omega\tau}$  on the whiff node — separate likelihood terms. (iii) The *Magnus/seam-shifted-wake instrument* of Section 6.2: equal- $\mathbf{d}$ , different-spin pitch pairs whose outcome gaps are attributable only to perception. The bias  $\delta_{\tau}$  is identified through the asymmetry of  $|\varepsilon^{\text{time}}|$  between earlier- and later-than-expected arrivals.

**Two conditions for deception.** A pitch beats the hitter when its induced movement is *both* (i) unexpected under the flight anchor (Equation 12) *and* (ii) projected onto the depth axis (Equation 14), where it is uncorrectable. Transverse-projected deviation fails condition (ii) and is largely tracked out. Because the eye sits to the batter’s own side, a same-side away-breaking ball loads  $d_{\parallel}$  far more than a toward-breaking ball, so the platoon and same-/opposite-hand breaking-ball splits emerge from Equation 16 with no handedness rule — a built-in validation target, operationalized through the platoon intercepts of Section 9.

## 8 Valuation

Run expectancy  $R(s)$  is defined over base-out  $\times$  count states (Tango, Lichtman, and Dolphin 2014), with a count-to-terminal recursion mapping per-pitch ball/strike probabilities into walk and strikeout probabilities. The observed-runs likelihood is

$$\mathcal{L}_m^{\text{runs}} : \quad \text{runs}_m \sim \text{NegBinomial} ( R(s_m), \xi ) . \quad (17)$$

The recursion’s per-count ball/strike probabilities are not free parameters: they are the outcome chain’s own marginals — take, called-strike, swing, and whiff probabilities averaged over the reference pitch and location distribution at each count. This is the coupling that makes Equation 1 a genuinely joint likelihood rather than a typographic product: the chain’s parameters enter the valuation likelihood through the recursion, the observed-runs data sharpen them in return, and run-value uncertainty propagates into the estimand instead of entering as a fixed label — the defect of regression-on-run-value pipelines noted in Section 1.

Per-pitch expected run value is the **derived functional** threading the chain (Equation 2) through the count-transition values  $\Delta R$  and the contact linear weights  $\nu(o)$  (the run value of terminal outcome  $o$ ;  $w$  is reserved for the gates and weights above):

$$\begin{aligned} \text{xRV}_i = & (1 - \pi_i^{\text{sw}}) [ \pi_i^{\text{cs}} \Delta R_i^{\text{cs}} + (1 - \pi_i^{\text{cs}}) \Delta R_i^{\text{ball}} ] \\ & + \pi_i^{\text{sw}} [ \pi_i^{\text{wh}} \Delta R_i^K + \pi_i^{\text{fr}} \mathbb{E}_{f^{\text{bb}}} [ \nu(o) ] + \dots ] . \end{aligned} \quad (18)$$

## 9 Hierarchy, priors, and the estimand

The **batter population** is partially pooled — swing zone  $w^{\text{sw}}$ , reach, swing speed  $v_{\text{bat}}$ , attack angle  $AA$ , the contact intercept  $\alpha_b$ , and the perceptual gains  $\kappa$  — as are the **umpire** and **catcher** called-zone effects; season effects absorb drift. Platoon intercepts (batter-hand  $\times$  pitcher-hand) are included **as a test statistic, not an explanation**: the eye-frame geometry of Section 7 is supposed to produce the platoon split on its own, so under a correct model these intercepts concentrate near zero, and posterior mass away from zero measures exactly the platoon structure the geometry fails to capture. Body dimensions (eye, shoulder, reach) are tied to batter height through an anatomy regression on military anthropometric survey data (Gordon et al. 2014), contributing  $\mathcal{L}^{\text{anat}}$ . Empirically, swing-decision sensitivity  $d'$ , reliance on the prior, the timing-error scale, and susceptibility to sequencing all vary with playing level, which is what licenses transport across populations: hold the physics and perceptual mechanism fixed and re-estimate the batter population (lower  $d'$ , larger Weber gains, slower bat) for amateur or high-school use (Gray 2010, 2002b, 2002a). Of note, handedness mirroring of the geometry (eye position, spin-axis sign) is in scope, since the platoon validation in Section 7 requires both hands.

The **pitcher is not a fitted random effect**: a pitcher is summarized by the flight inputs he produces, so his grade is a functional of the model, not a parameter. For a pitch shape  $\varsigma$  (its release state),

$$\text{Stuff}(\varsigma) = \mathbb{E}_{s, \mathbf{x}^{\text{pl}} \sim \text{ref}, b \sim \text{ref-pop}} [ \text{xRV}(\varsigma, \mathbf{x}^{\text{pl}}, s, b) ], \quad (19)$$

integrating location over a fixed reference distribution and the batter over a reference population. That integral is the Stuff/Location separation, with the interaction retained. Report the posterior mean with a credible interval; rescale to mean 100, SD 10 only for display.

## 10 Reporting: Stuff, Location, and their interaction

Modeling jointly does not preclude the familiar separate grades — the separation is a *reporting* decomposition of Equation 18, not a modeling choice. From one model:

- **Stuff** — value holding the shape, averaging location over a neutral reference;
- **Location** — value attributable to where the pitch is thrown, given the shape;
- **Interaction** — how much the locating amplifies or wastes the stuff.

The two main effects are always recoverable from a model containing the interaction, while the interaction is unrecoverable from two main-effect-only (additive) models. Here the decomposition is exact – the three pieces sum to total xRV – rather than assumed, and one can always collapse back to the two conventional grades.

## 11 A minimal version and a staging path

The full chain need not be fit at once, and the smallest model that still embodies the data-generating process is one GLM. Fit the whiff node (Equation 6) alone, on swings, with the core form of the timing error (Equation 16 at  $\delta_\tau = \kappa_{\omega z} = \kappa_{\omega \tau} = 0$ ). Its inputs reduce to the gravity-removed acceleration  $\mathbf{a}^{\text{aero}}$  and the pickup velocity direction (both from the 9P state), the flight time  $t^f$ , fixed anatomical eye constants  $\mathbf{E}$ , and the reach geometry – because under constant acceleration the deviation collapses to  $\mathbf{d}(t^f) = \frac{1}{2} \mathbf{a}^{\text{lift}} (t^f - t_0)^2$ . One pooled batter intercept and three positive coefficients complete the model; the perceptual gain  $\kappa_{\parallel}$  enters only as a product with  $b_\tau$  and is not separately estimated. Deliberately absent: platoon terms, pitch-type labels, and spin. The platoon split emerging through  $\varepsilon^{\text{time}}$  alone is therefore the immediate, falsifiable test, and a pitch is graded by its trajectory, not its name.

From there, each stage adds one assembled component and one new testable claim: (1) the swing/take and called-strike links (Equation 3, Equation 5), which make the model generative for called-strikes-plus-whiffs and, through Section 8, for count-transition run values; (2) contact quality through the collision map (Equation 7–Equation 8); (3) the staged extensions – the situational blend (Equation 13), the spin-illusion and bias terms of Equation 16, and the windowed transverse accumulation of Equation 15 with its estimable commit lead  $\Delta$ .

## 12 Validation, limitations, and discussion

**What would falsify the mechanism.** The model carries its tests inside it. The platoon intercepts of Section 9 concentrate near zero if the eye-frame geometry is right, and measure what it misses if not. The spin-illusion gains are estimated against a “perception = physics” null (Section 7); posterior mass at zero retires that extension. The commit lead  $\Delta$ , once the windowed transverse channel is fit, should agree with the behavioral commit point rather than wander freely. The three expectation anchors – constant-velocity, gravity-only, and the full-physics core – are directly comparable on held-out predictive density. And the chain invites ordinary posterior predictive checks: swing and called-strike frequencies by count and location, whiff rates by pitch shape, exit-velocity–launch-angle surfaces by contact offset.

**Limitations.** The 9P fit imposes a common within-flight accrual profile, so the windowed accumulation of Equation 15 adds information only when full-flight tracking data replace the constant-acceleration summary. The batter’s eye and stance are placed from anatomy rather than measured, so stance variation loads on the perceptual gains. The core sets the situational weight to zero, which understates deception for pitchers with strong sequencing patterns – a deliberate choice, since arsenal priors do not transport to thin amateur data. Batter intent beyond the count (situational swing plans) is absorbed by the count propensities rather than modeled. And handedness mirroring of the geometry is assumed exact.

**What the structure buys.** Mechanism: a grade decomposes into named physical causes – reaction budget, depth-projected deviation, collision offset – rather than feature attributions. Transport: the physics and perception maps are population-invariant, so moving from a professional calibration to an amateur population re-estimates only the batter layer. Honesty: every output is a posterior functional with a credible interval, and every simplification in the document is stated together with its information cost. The nearest-term extension with the highest payoff is

full-flight tracking, which converts the windowed accumulation of Equation 15 from an assumption-driven refinement into a measurement.

### 13 References

- Baseball, Driveline. 2021. "Pitch Design: What Is Stuff+? Quantifying Pitches with Pitch Models." *Driveline Baseball*.
- Chen, Tianqi, and Carlos Guestrin. 2016. "XGBoost: A Scalable Tree Boosting System." In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–94. San Francisco California USA: ACM. <https://doi.org/10.1145/2939672.2939785>.
- Cross, Rod, and Alan M Nathan. 2006. "Scattering of a Baseball by a Bat." *American Journal of Physics* 74 (10): 896–904. <https://doi.org/10.1119/1.2209246>.
- Gelman, Andrew, and Jennifer Hill. 2007. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press.
- Gordon, Claire C, Cynthia L Blackwell, Bruce Bradtmiller, Joseph L Parham, Patricia Barrientos, Stephen P Paquette, Brian D Corner, et al. 2014. "2012 Anthropometric Survey of U.S. Army Personnel." U.S. Army Natick Soldier Research, Development and Engineering Center.
- Gray, Rob. 2002a. "Behavior of College Baseball Players in a Virtual Batting Task." *Journal of Experimental Psychology: Human Perception and Performance* 28 (5): 1131–48. <https://doi.org/10.1037//0096-1523.28.5.1131>.
- . 2002b. "Markov at the Bat." *Psychological Science* 13 (6): 542–47.
- . 2010. "Expert Baseball Batters Have Greater Sensitivity in Making Swing Decisions." *Research Quarterly for Exercise and Sport* 81 (3): 373–78. <https://doi.org/10.1080/02701367.2010.10599685>.
- . 2013. "Being Selective at the Plate: Processing Dependence Between Perceptual Variables Relates to Hitting Goals and Performance." *Journal of Experimental Psychology: Human Perception and Performance* 39 (4): 1124–42. <https://doi.org/10.1037/a0030729>.
- . 2017a. "A Baseball Batter's Perception of Spin & How It Relates to Hitting Performance." *The Perception & Action Podcast*.
- . 2017b. "Pitch Tunneling & Perceptually Equivalent Pitches." Educational. *Perception & Action Podcast*.
- Gray, Rob, and Rouwen Cañal-Bruland. 2018. "Integrating Visual Trajectory and Probabilistic Information in Baseball Batting." *Psychology of Sport & Exercise* 36 (May): 123–31. <https://doi.org/10.1016/j.psychsport.2018.02.009>.
- Gray, Rob, and David M Regan. 2006. "Unconfounding the Direction of Motion in Depth, Time to Passage and Rotation Rate of an Approaching Object." *Vision Research* 46 (15): 2388–2402. <https://doi.org/10.1016/j.visres.2006.02.005>.
- Grove, Cameron. 2023. "PitchingBot Pitch Modeling Primer." *Sabermetrics Library*. <https://library.fangraphs.com/pitching/pitchingbot-pitch-modeling-primer/>.
- Kagan, David, and Alan M Nathan. 2017. "Statcast and the Baseball Trajectory Calculator." *The Physics Teacher* 55 (3): 134–36. <https://doi.org/10.1119/1.4976652>.
- Lambert, Jack. 2024. "Revisiting Stuff+: An Update on Driveline's Methodology to Quantifying Pitch Design." *Driveline Baseball*.

McGrattan, Owen. 2023. "Stuff+, Location+, and Pitching+ Primer." *Sabermetrics Library*. <https://library.fangraphs.com/pitching/stuff-location-and-pitching-primer/>.

McIntyre, J, M Zago, A Berthoz, and F Lacquaniti. 2001. "Does the Brain Model Newton's Laws?"

Nathan, Alan M. 2008. "The Effect of Spin on the Flight of a Baseball." *American Journal of Physics* 76 (2): 119–24. <https://doi.org/10.1119/1.2805242>.

———. 2018. "Determining the 3D Spin Axis from TrackMan Data (Updated, March 31, 2015 and August 31, 2018)." *Physics of Baseball*, 1–8.

Nestico, Thomas. 2024. "Modelling tjStuff+ V3.0." *Medium*.

Spencer, Scott. 2023. "Pitch Selection to Maximize Motion-in-Depth Variation." Blog. <https://ssp3nc3r.github.io/post/pitch-selection-to-maximize-motion-in-depth/>.

Stronge, W. J. 2018. *Impact Mechanics*. Second edition. Cambridge, United Kingdom: Cambridge University Press.

Tango, Tom M., Mitchel G. Lichtman, and Andrew E. Dolphin. 2014. *The Book: Playing the Percentages in Baseball*. Place of publication not identified: TMA Press.