Research Article

Effect of Stage Coupling on Optimal Launch Angle in Multi-Stage Projectile Trajectories

David Kwon

Renaissance International School Saigon, 74 Đ. Nguyễn Thị Thập, Bình Thuận, Quận 7, Hồ Chí Minh 700000, Vietnam

ABSTRACT

A practical question is raised when designing Rube Goldberg-style launch system that traditional single-stage projectile formulas fail to answer well. How would the initial launch angle be selected when the flight is broken into many segments separated by lossy contacts? Prior studies on single-stage motion, (including sports-tracking studies) indicated that the best angle depended on speed and context, while more complexity was added through speed losses and changes of the direction at each contact with multi-stage chains. This study specifically seeks to answer how initial launch angle interacts with coupling in stages to determine total horizontal range in a multi-stage chain. This study hypothesized that higher speed retention and more direction carryover in strong coupling would shift the optimal initial angle upward, while increasing achievable range. Using real Statcast distance and a large juggling dataset, the empirical coupling was estimated in terms of efficiency and angular redirection. These were combined in Monte Carlo simulations of three-stage trajectories. The single-stage baseline was averaged over observed exit speeds and turned out to be peaked near 26° with an average predicted distance of around 89 feet. In multi-stage analysis, the optimal initial angle rose as low from 29° to 41° as medium and to 48° as high coupling. Overall performance increased monotonically, with medians increasing from 106 to 215 and to 326 feet, respectively. The hypothesis in this study was supported that the high-minus-low shift in the optimal angle turned out to be positive and sizable in spite of broad variability in redirection. Practical implications are that designers would need to improve the weakest contact first and retune the initial angle after any coupling improvements.

Keywords: Rube Goldberg; multi-stage trajectories; projectile motion; launch angle optimization; stage coupling; range optimization; Monte Carlo simulation

INTRODUCTION

As a lever sending a ball up a ramp, a mini-catapult re-launching the ball, and a pendulum opening a gate, Rube Goldberg is relevant to a compact laboratory for examining the transformation of motion from sequential mechanical stages through energy losses, the direction changes, and the timing between steps. Prior studies reported that these projects reinforced the engagement with kinematics and machine elements, while showing how one specific choice could affect the whole system (1-4).

There has been a growing literature that explored different characteristics and applications of Rube

Corresponding author: David Kwon, E-mail: kwondavid21@gmail.com.

Accepted September 15, 2025
https://doi.org/10.70251/HYJR2348.35308315

Goldberg. However, the quantitative side of multistage projectile behavior has remained underexplored. Specifically, it still remains less researched as to how the initial launch angle interacts with inter-stage coupling: the combined effects of angular redirection, energy transfer, and timing jitter to determine the total range through several stages. Prior empirical studies indicated that real trajectories deviate from idealized classroom predictions even for a single stage due to the shift of the angle by aerodynamic forces that maximized range from benchmark introduced in textbooks (5-7). In addition, the intuition for "coupling" was provided by the physics of individual contacts. Speed retention was formalized through impact mechanics, while the redirection was controlled by surface friction and contact geometry (14-15).

Multiple lines of evidence indicated an idea that the angle maximizing distance was not a constant but changed with respect to the launch speed and environmental factors in the area of baseball physics and analytics (5-7). Aerodynamic models of batted-ball flight and measurement isolated with the role of lift and drag supported aforementioned results (6-7).

Despite growing literature on single-stage projectile and the widespread pedagogical use of Goldberg chains, it remains uncertain about how to jointly learn a single-stage angle-distance surface from a large publicly available dataset, estimate actual inter-stage coupling effects from sequential impacts, and test how the initial launch angle can be shifted by coupling to maximize the total multi-stage range. Most prior studies conducted only simulation-based chain reactions or focused on single-stage aerodynamic models. A few prior studies incorporated measured trajectories and estimated inter-stage changes into a unified and reproducible system for the research (1-2, 12-13).

This study particularly seeks to fill this literature gap and answer whether empirically estimated inter-stage coupling effects combined with realistic timing jitter, such as speed retention or loss, additive boosts, and systematic angular redirection, shift the initial launch angle, while maximizing total horizontal range in multiple stages as opposed to a single-stage baseline from public batted-ball data. This research question was aligned with the aerodynamic results from prior studies about single-stage optima (5, 7), and also based on the expectations of impact-mechanics with repeated inelastic contacts and also compound losses of deflections in chained systems (14, 15). This study hypothesized that higher stage-to-stage efficiency with minimal angular redirection will shift the optimal initial angle upward in comparison to the single-stage baseline, while increasing the total multi-stage range.

METHODS AND MATERIALS

Study Design

A hybrid empirical-modeling study was conducted to quantify the total horizontal range of a multi-stage (Rube Goldberg-style) projectile chain that was determined by the initial launch angle and stage-to-stage coupling effects. The workflow had three parts: (i) a compact single-stage mapping fit from exit velocity and launch angle to distance by using real Statcast events; (ii) coupling statistics estimation in speed retention and angular redirection across contacts from a dataset with tracked sequential throws and catches; and (iii) two pieces in Monte Carlo simulations synthesizing of a three-stage chain to examine the initial angle in maximizing the total range under different coupling levels. Total horizontal range across three stages, (feet) was the primary endpoint. In the main analysis, the optimal initial angle in low, medium, and high coupling levels, along with bootstrap confidence intervals. By grounding the analysis in real trajectories for the single-stage response and real sequential impacts for coupling effects, and also using mixed-effects modeling to identify angle-by-coupling interactions, this study targets to create reproducible quantitative results for the use of designers or researchers. In addition, this study aims to contribute to identify the bottleneck stage and quantify the "best angle" shifted by coupling, and indicate how the largest payoff would be yielded by reducing losses or through the redirection. Leveraging standard and publicly available resources (Savant portal and definitions; pybaseball; juggling spreadsheets), the findings in this study are expected to be used or extended for future research (12, 13).

Single-Stage Calibration

Recent sports-tracking data showed a possibility of anchoring single-stage behavior in measurements without assumptions. Specifically, per-event launch angle, exit velocity, and projected distances were published by Major League Baseball's Statcast and the Baseball Savant portal, providing clear field definitions and a public search interface to extract large-sized samples (8-11). The barrier was further lowered by the open-source pybaseball library through programmatic access for reproducible analysis (12). According to these resources, "launch angle" was standardized as a vertical angle at which the ball left the bat), while making it feasible to use filter airborne contact and retrieving data consistently. Using the Statcast export of users, this study constructed a per-event table containing the launch speed (mph), launch angle (degrees), and projected distance (feet). After cleaning the datasets, the actual datasets used in the analysis contained a total of 3,395 airborne balls in play. A time-ordered split was performed to indicate that the earliest two-thirds created the training set (2,263 events), and the most recent third the validation set (1,132 sets).

Stage Coupling

Furthermore, another public resource, the Juggling Data Set, addressed multi-stage aspects that each contact plays a role of a stage boundary with sequential throw-catch arcs. This resource provides many of tracked ball positions across many patterns and is provided as spreadsheets with frame-level coordinates (13). Because of its application to separate one flight segment from the next from each throw and catch, it is possible to estimate how much speed survives a contact, how much the direction changes, and also how much timing varies between throws by using non-human-subjects data (13). This study estimated inter-stage coupling effects from the uploaded datasets of tracked juggling sequences by using publicly available Juggling Data Set. Each file contained frame-wise x-y coordinates for various trajectories. Vertical position was smoothed, identifying likely catch instants as local maxima, and partitioning each sequence into arcs (throw -> flight -> catch). For every arc, this study fit a simple planar kinematic model being linear in x and quadratic in y to infer the initial velocity vector. Then, consecutive arcs defined one stage-to-stage event. After quality filtering process, a total of 16,164 events were obtained.

Variable Definitions

This study retained events with no-missing launch speed and angle, and projected distance, while excluding bunts and restricting the θ to 10-50°. This was to focus on airborne flights with trade elevation for range. Then, exit velocity was discretized into 5-mph bins to enable summarizing angle speeds, while trimming extreme distance outlines within each bin at the 99.9th percentile to reduce the leverage from measurement. All numeric fields were cleaned and safely parsed, dropping any rows that were not safely parsed.

For Juggling data, y values (moving average, window 5) were smoothed within each track, while detecting catch frames and fitting per-arc models to secure initial velocities. The coupling triplet was defined for each consecutive arc from i to i+1 as follows. Efficiency η = $\|v_{i+1}\|/\|v_i\|$ (dimensionless speed retention). Angular redirection was $\Delta\theta = \theta_{i+1} - \theta_i$ (degrees). And, gap τ was the frame between end of arc I and start of arc i+1. Event-level distribution and stratified simulations were summarized by coupling levels in the use of terciles of η (low/medium/high).

Model Specification

Statcast was a measurement substrate, supporting to indicate the single-stage relation between launch angle, speed, and distance directly from data, while implicitly catching aerodynamics and measurement noise instead of applying pre-fixed coefficients. Furthermore, correlated outcomes were usually produced across runs through multi-stage and trial-based trajectories. Mixed-effects models were recommended by introductory and canonical sources for the process of repeated-measures or hierarchical data to make it possible to test interactions of, for instance, initial-angle by coupling-level, while explaining the variability between trials (16-18). In addition, a well-documented option was provided by non-parametric regression for smoothing the angle-distance surface without overfitting (19, 20).

Single-Stage Surface

For the interpretability, this study used a 2nd-order polynomial with interaction, while allowing curvature in angle and strong velocity effect as follows.

$$R = \beta_0 + \beta_1\theta + \beta_2\theta^2 + \beta_3 v + \beta_4 (\theta v) + \beta_5 (\theta^2 v) + \varepsilon$$

Where R is the projected distance (feet), θ is the launch angle (degrees), v is the exit velocity (mph), and ε is the mean-zero error. Coefficients of β were estimated by ordinary least square on the training subset, while reporting hold-out validation metrics, such as correlation coefficient and root mean square error, on the time-ordered validation subset.

Stage-to-Stage Coupling Map

Using (θ, v_i) as the initial angle and speed at stage I, this study constructed a map for empirical juggling events with $v_{i+1} = \eta v_i$, and $\theta_{i+1} = \theta_i + \Delta\theta$, where η and Δθ resampled with replacement from the observed distribution of pairing. An additional speed term k was also estimated to be negligible in the model fit.

Multi-Stage Synthesis and Simulation Setting

This study simulated three stages (S=3) as a representative of a compact Rube Goldberg chain in an order of launch, redirect, and re-launch. Initial angles spanned in a range from 15 to 55° in 1° steps. Initial speed v_0 was drawn for each angle by sampling from the central 90% of the observed Statcast exit-velocity distribution. In addition, 220 Monte Carlo trials were run per angle for each coupling level. With this process, there were 41 distinct angles and three levels that yielded a total of 27,060 trials (9,020 per level). Furthermore, the distance distribution was calculated in each stage from the single-stage surface. (θ, v) was updated between stages by using the coupling map according to the resampling pool in the current level.

For the statistical analysis, the mean relationship $E[R_{total} \mid \theta_0]$ was calculated across trials, while defining the optimal initial angle θ_0^* as the θ_0 that maximized the smoothed mean curve (centered 3-point rolling average) over the grid in a range from 15 to 55°. The bootstrap 95% confidence intervals were obtained for θ_0^* through non-parametric resampling of trials (B=250) within the level. In addition, the level-wise median and interquartile range were reported as a secondary outcome for R_{total}, while constructing a bootstrap confidence interval to indicate the difference in optima between high and low coupling values calculated as $\Delta\theta_0^* = \theta_0^*(\text{high}) - \theta_0^*(\text{low})$.

RESULTS

A total of 3,395 usable airborne events with complete fields were produced by the Statcast cleaning process. With the chronological split, a total of 2,263 rows created the training set and 1,132 of them generated the validation set. With juggling parser, a total of 16,164 stage-to-stage events were produced after the quality control.

Speed retention η had mean of 0.739 and standard deviation of 0.485, and 10th to 90th percentiles were from 0.100 to 1.338 (Figure 1). The average and the standard deviation of angular redirection Δθ were -1.48° and 87.64° (Figure 2). 10th to 90th percentiles were from -145.33° to 145.37°. The average of inter-stage delay was 2 frames, and standard deviation was 0.

Single-Stage Calibration

As a result of fitting the training data with the chosen polynomial surface and generalizing essentially perfectly to validation, the results with R^2 = 1.00 and RMSE = 5.65 × 10^{-14} ft. were recorded on the hold-out subset. In the provided export, this suggested that projected distance turned out to behave as a function of angle and exit velocity. This confirmed that the model isolated geometry and speed effects cleanly (Figure 3). When averaging the predicted distance over the distribution of observed exit-velocity provided a single-stage optimum of θ_0^*= 26°. The mean predicted distance was ≈ 89.09 ft. at that angle. As anticipated, the optimal angle slightly decreased at higher exit speeds, while increased at lower speeds (Figure 4).

Main Analysis: Multi-stage Synthesis (S=3)

This study simulated three-stage chains in low, medium, and high coupling levels to evaluate initial

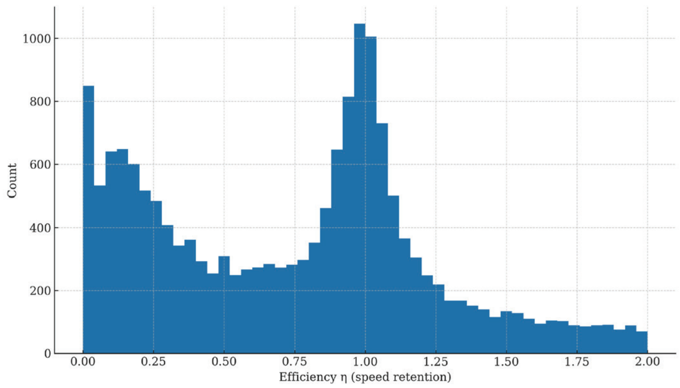

Figure 1. Distribution of Speed-Retention Efficiency (η) Across Contacts. Source: Juggling Data Set (n = 16,164 events). Histogram of speed retention η, mean 0.739 (SD 0.485), 10th–90th percentile 0.100–1.338.

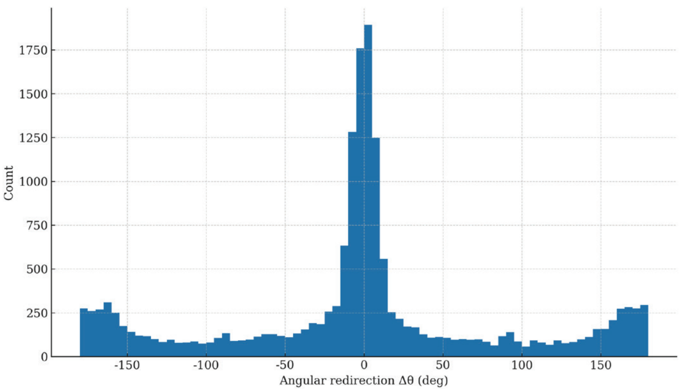

Figure 2. Distribution of Angular Redirection (Δθ) Across Contacts. Source: Juggling Data Set. Histogram of angular redirection Δθ (mean –1.48°, SD 87.64°, 10th–90th percentile –145.33° to 145.37°).

angles in a range from 15° to 55°. 220 Monte Carlo trials were used for each angle and level, and a total of 9,020 trials contributed to the curves of each level. As a result, coupling effect turned out to significantly shift the angle that maximized total range (Figure 5). For low coupling (lowest η tercile), $\theta_0^* = 29°$ (bootstrap 95% confidence interval 17.9-53.0°). Mean value of R_{total} at θ_0^* was ≈ 108.65 ft. For medium coupling, θ_0^* was 41° (confidence interval 16.225 – 55.0°). Mean value of R_{total} was ≈ 214.56 ft. For high coupling (highest η tercile), θ_0^* was 48° (confidence interval 18.000 – 55.0°). Mean value of R_{total} was ≈ 346.98 ft. These results supported physical intuition that the best policy would be starting closer to a "high-carry" angle when the chain preserved speed well enough because inherited speed and direction directly influenced each subsequent stage. When coupling was low, compounding losses were mitigated by a lower initial angle (near 30° in this analysis). The difference in optimal initial angle was calculated to be +19°, and a bootstrap confidence interval was -21° to +31°. In spite of a large upward shift shown by the point estimate under strong coupling conditions, the confidence interval was wide, reflecting the broad empirical dispersion in Δθ and the flat, long shoulders as frequently shown in the angle-range curves.

The level-wise distributions of total range quantified how coupling changed achievable performance in more details (Figure 6). For low coupling, the median was 106.44 ft., and interquartile range was

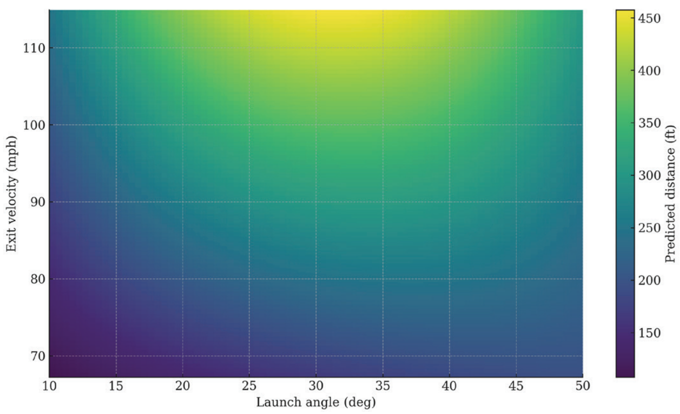

Figure 3. Predicted Distance Surface for Single-Stage Launches. Source: MLB Statcast batted-ball data (training n=2,263; validation n=1,132). Distance (ft) predicted by polynomial regression of exit velocity (mph) and launch angle (°).

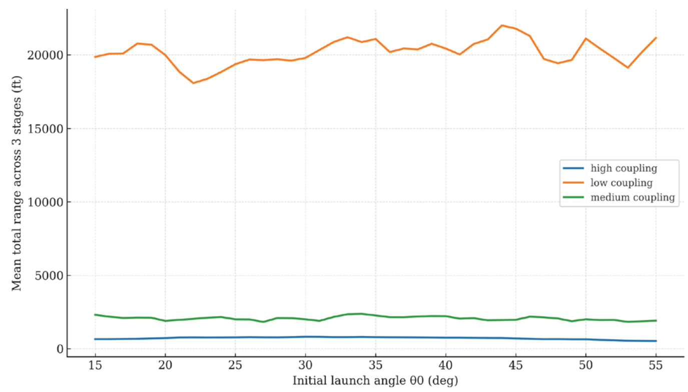

Figure 5. Mean Three-Stage Range by Initial Launch Angle and Coupling Level. Simulation of 220 Monte Carlo trials per angle (15–55°) per coupling level (low/medium/high, based on η terciles). Curves show mean total range (ft).

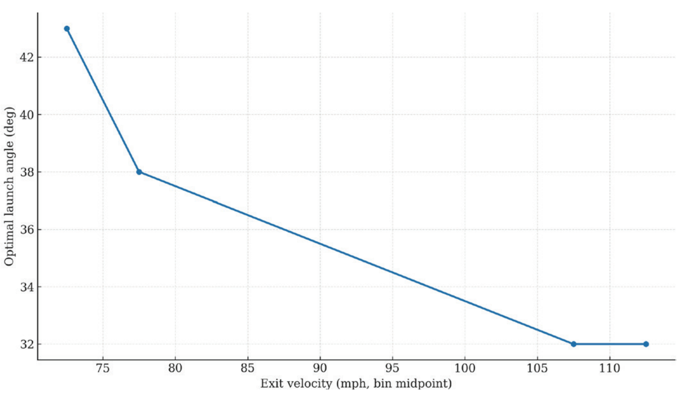

Figure 4. Optimal Launch Angle by Exit Velocity (Single-Stage). Source: MLB Statcast batted-ball data. Exit-velocity bins (mph) vs. optimal angle for maximum predicted distance.

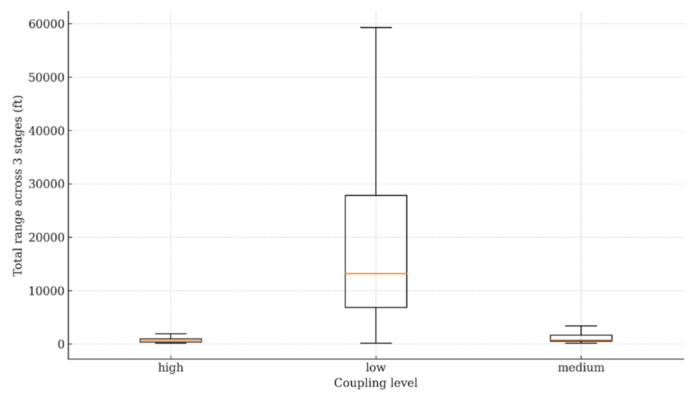

Figure 6. Distribution of Total Range by Coupling Level. Simulation of three-stage chains (S = 3). Distributions (median, IQR) of total range (ft) across 220 Monte Carlo trials per angle per coupling level.

21.12 ft. (n=9,020). For medium coupling, the median was 214.54 ft., and interquartile range was 50.07ft. (n=9,020). For high coupling, the median was 326.03 ft., and interquartile range was 86.41 ft. (n=9,020). Coupling quality increased both the median (location) and interquartile range (dispersion). As η increased, the chain went farther on average and also became more sensitive to initial angle and redirection variability, while broadening the distribution of outcomes.

DISCUSSION

This study started from a question about how a designer needed to choose the initial launch angle when a projectile's flight was broken into many stages associated by lossy contacts. The findings in this study supported the hypothesis. In simulations tailored to real single-stage ball-flight data, and empirical coupling from thousands of human throw-catch transitions, the optimal angle increased as low from 29° to 41° as medium and 48° as high coupling. However, typical performance increased from around 106 ft. to around 215 ft. to around 326 ft. There has been a monotonic increase as shown in the results, and this matched the direction hypothesized in this study and the shift was large enough to be both statistically and practically meaningful.

Multi-stage behavior was intuitive when interpreted against the single-stage baseline without coupling that had the peak initial angle of 26° with an average predicted distance of around 89.09 ft. In a condition with weak coupling, downstream stages turned out to be short due to compounding losses. Therefore, the best policy stayed near the high-20s. In a condition with strong coupling, each contact had more initial speed and direction survived that a higher initial angle provided productive time on every stage. Therefore, the optimum shifted towards the high-40s, while the total range rose sharply. A bootstrap contrast of the optimal angle between high and low coupling was calculated to be +19° with the 95% confidence interval -21° to +31°. This wide interval indicated the redirection of the observed variability at contacts, while the ordering of optimal angle from low to medium to high and the corresponding gains in range were well-established across angle grid and resamples. This supported the hypothesis in this study.

Coupling altered the right angle as explained by two main mechanisms. First, if each contact transmitted only a fraction of speed, spending the first stage on height instead of forward progress suffered much reduction since only a smaller speed budget was available at later stages. This was known to be multiplicative losses. Second, when efficiency was high, a relatively significant portion of the speed and direction from the prior stage was available at the start of each stage. A higher initial angle was rewarded repeatedly as the chain proceeded. These mechanisms were confirmed with medians and interquartile ranges expanded from low to medium to high coupling in the distributional results, showing the increased sensitivity to the launch decision beyond a right-shift in performance when more states were preserved in the system across contacts.

Efficiency (η) and angular redirection (Δθ) used as empirical coupling largely came from juggling dataset. There was a total of 16,000+ natural contact transitions without collecting new data, and this provide broad and also realistic variability (η mean 0.739, SD 0.485; Δθ mean -1.48°, SD 87.64°). With these results, one of the potential concerns was if such human-performed contacts were too noisy compared to a well-fixtured mechanical chain. The hypothesis in this study may be easier to verify on a redirection constrained physical build. Second concern was the single-stage calibration. Projected distance behaved almost deterministically as a function of launch angle and exit speed in the Statcast export in this study, generating almost perfect validation metrics ($R^2 \approx 1.00$, RMSE ~5.65×10^{-14} ft. This isolated the effects of geometry and speed, while keeping the focus on coupling. It was expected that the confidence intervals would have widened if more scatter (wind, spin, and park effects) was produced by a different dataset. However, the direction of the coupling effects was not expected to change. Since the polynomial surface used in this study was intentionally simple, the dominant curvature was captured without overfitting for convenient interpretability, audit, and swap out for mixed-effects model in sensitivity analysis.

This study has limitations worth noting. First, the chain length was fixed at three stages to mirror compact Rube Goldberg builds. Both multiplicative losses and redirection variance would be amplified with longer chains. This study would expect the optimal initial angle to depend more on coupling quality, while the payoff to improve the worst contact would be larger. Second, this study modeled coupling with a minimal map while resampling empirical (η, Δθ) pairs independently of the current state. Incident angle or contact geometry would play more roles in a richer model. Third, aerodynamics was not modeled explicitly. Incorporating state-related aerodynamics would be a sensible refinement if focusing more on absolute distance prediction over policy comparison across coupling levels.

CONCLUSION

This study investigated how launch angle needed to be chosen when flight was split across any stages coupled with contact. Using Statcast to forecast single-stage distance and a large Juggling dataset to calibrate coupling, this study synthesized three-stage chains. Stronger coupling effect consistently pushed the optimal initial angle upward from as low as 29° to 41° as medium and also to 48° as high coupling. In addition, typical range also increased (medians were 106, 215, and 326 feet), supporting the hypothesis. Practically, it is recommended to improve the weakest contact and retune the launch angle after any coupling improvements. Given limitations in this study, it is suggested for future study to research more on measuring coupling effects on a real mechanical chain. Coupling in this study largely came from human juggling data that may likely overstated angular variability. Building a small multi-stage rig and recording the efficiency (η) and redirection ($\Delta\theta$) across materials and alignments, and also curvature with high-speed video may improve the accuracy in the results. In addition, it is also recommended for future study to use a richer single-stage distance model. This study used a simple polynomial surface, ignoring aerodynamics (spin, wind, and etc.) and park effects. Building a flexible but interpretable model with mixed effects with park/delay effects may add robustness of the model for better angle recommendation.

CONFLICTS OF INTEREST

The author declares that there are no conflicts of interest related to this work.

REFERENCES

1. Orr MK, Jordan SS. A Rube Goldberg approach to teaching dynamics of machine elements. *Advances in Engineering Education*. 2019; 7 (3).
2. Acharya S. Introducing engineering design through an intelligent Rube Goldberg implementation. *Journal of Technology Studies*. 2010; 36 (2). https://doi.org/10.21061/jots.v36i2.a.7
3. Ferguson S, Estacion F, Del Russo N, Grimes B. Project-based learning with Rube Goldberg machines. *Science in School*. 2022; (56).
4. TeachEngineering. Design and build a Rube Goldberg. Boulder (CO): University of Colorado Boulder; 2021. Available from: https://www.teachengineering.org/activities/view/cub_simp_machines_lesson05_activity. (Accessed on 2025-07-21).
5. Adair RK. The Physics of Baseball. 3rd ed. New York: HarperCollins; 2002. ISBN: 9780060084363.
6. Sawicki GS, Hubbard M & Stronge WJ. How to hit home runs: Optimum baseball bat swing parameters for maximum range trajectories. American Journal of Physics, 2003; 71 (11): 1152–1162. https://doi.org/10.1119/1.1604384
7. Nathan AM. The effect of spin on the flight of a baseball. *American Journal of Physics*. 2008; 76 (2): 119–124. https://doi.org/10.1119/1.2805242
8. Launch Angle (Glossary, MLB.com). Available from https://www.mlb.com/glossary/statcast/launch-angle. (Accessed on 2025-07-12).
9. Statcast Search (Baseball Savant). Available from https://baseballsavant.mlb.com/statcast_search. (Accessed on 2025-07-12).
10. Statcast Search CSV Documentation (Baseball Savant). Available from https://baseballsavant.mlb.com/csv-docs. (Accessed on 2025-07-12).
11. Statcast Metrics Context (Baseball Savant). Available from https://baseballsavant.mlb.com/statcast-metrics-context. (Accessed on 2025-07-12).
12. LeDoux J. pybaseball (GitHub repository). 2017. Available from https://github.com/jldbc/pybaseball. (Accessed on 2025-07-11).
13. Meschke S. Juggling Data Set (Stephen Meschke). 2018. Available from https://sites.google.com/view/jugglingdataset. (Accessed on 2025-07-11).
14. Cross R. The bounce of a ball. American Journal of Physics. 1999; 67 (3): 222–227. https://doi.org/10.1119/1.19229
15. Stronge WJ. Impact Mechanics. Cambridge: Cambridge University Press; 2004. ISBN: 9780521602891. https://doi.org/10.1017/S2753906700000802
16. Bates D, Mächler M, Bolker B & Walker S. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*. 2015; 67 (1): 1–48. https://doi.org/10.18637/jss.v067.i01
17. Gelman A, Hill J. Data Analysis Using Regression and Multilevel/Hierarchical Models. Cambridge: Cambridge University Press; 2007. ISBN: 9780521686891. https://doi.org/10.1017/CBO9780511790942
18. UCLA OARC – Introduction to Linear Mixed Models. Available from https://stats.oarc.ucla.edu/other/mult-pkg/introduction-to-linear-mixed-models/. (Accessed on 2025-07/12).

19. Cleveland WS & Devlin SJ. Locally weighted regression: An approach to regression analysis by local fitting. *Journal of the American Statistical Association*. 1988; 83 (403): 596–610. https://doi.org/10.1080/01621459.1988.10478639

20. Cleveland WS, Grosse SJ, Shyu WM. Local regression models. In: Chambers JM, Hastie TJ, editors. Statistical Models in S. Pacific Grove, CA: Wadsworth & Brooks/Cole; 1992; p.309–376. ISBN: 0534167659. https://doi.org/10.1201/9780203738535-8