9TH ANNUAL
MIT SLOAN
SPORTS ANALYTICS CONFERENCE
FEBRUARY 27 - 28, 2015 BOSTON CONVENTION AND EXHIBITION CENTER

# Move or Die:
# How Ball Movement Creates Open Shots in the NBA

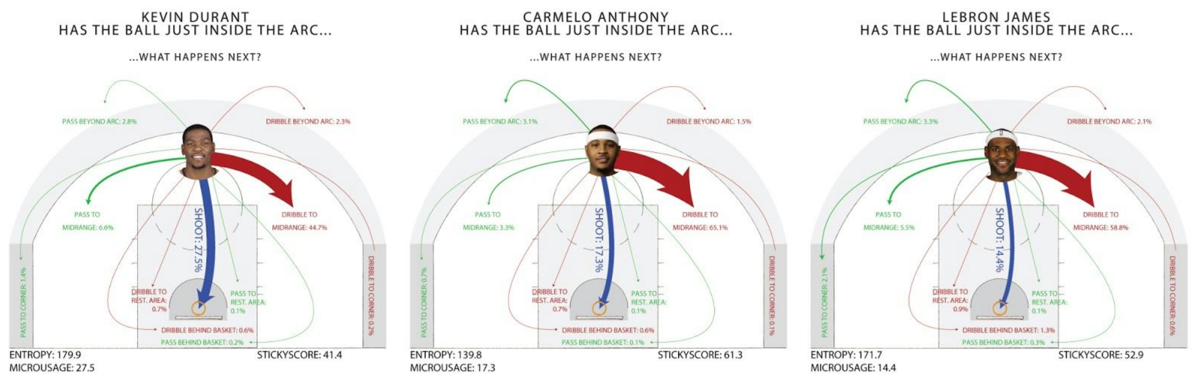
Alexander D'Amour, Daniel Cervone, Luke Bornn, and Kirk Goldsberry
Harvard University
Cambridge, MA 02139
Email: damour@fas.harvard.edu

## Abstract

Throughout the NBA, coaches consistently stress the importance of ball movement in a functional and efficient offense: "either you move it or you die," according to San Antonio Spurs coach Gregg Popovich. In theory, teams with more effective ball movement find open shots more efficiently, and are more difficult to defend against. However, even with the advent of tracking data that provides a complete catalogue of offensive movement in the NBA, basketball analysts have yet to quantify ball movement in a way that predicts offensive success. We show that a critical role of ball movement is to introduce unpredictability into the offense, and that this unpredictability is both theoretically justified and correlates well with opportunities for open shots. To explore this relationship, we introduce entropy and opportunity metrics, and show how teams and players balance capitalizing on immediate shot opportunities with entropic strategies that can force the defense to yield better scoring opportunities down the line. Our approach uses Markov modeling in novel ways and opens the door for a long-overdue game-theoretic understanding of basketball analytics.

## 1 "You move it or you die"

On June 8, 2014 the Miami Heat defeated the San Antonio Spurs in game 2 of the NBA Finals. After the game, the Spurs' coach, Gregg Popovich, suggested one key reason for the loss was his team's uncharacteristically stagnant offensive ball movement stating, "either you move it or you die." For years basketball coaches have preached about the importance of ball movement; however, to this point basketball analysts have struggled to measure movement in a way that connects with offensive success. In short, contemporary basketball analytics fail to adequately characterize the mechanics of effective offensive basketball. While many conventional analytics are able to discern which offensive teams are most and least effective, no current metrics offer any explanatory power to describe how tactical movements correspond with offensive effectiveness.




New optical tracking data allow us to measure players' speeds and distances traveled---and we have long been counting passes, touches, and dribbles---but this new data has not clarified the link between ball movement and scoring. While eventual champions with 62 wins, the San Antonio Spurs led the NBA in 2013-14 in the average speed at which their players moved; the Philadelphia 76ers, who lost 63 games, were second. The Spurs were also in the top 5 in terms of both passes and touches per offensive possession, but so were the Utah Jazz and Milwaukee Bucks, who combined for only 40 wins that season. League-wide, we fail to detect any significant correlations between conventional, aggregate ball movement indicators such as speed, distance, passes, and touches and measures of scoring efficiency, either on a per-possession or per-game basis. It seems, therefore, that ball and player movement are not predictive of a team's success. Was Popovich's statement misfounded, therefore, and is extra movement really just wasted energy?

A stagnant offense is easy to spot, and analytically, it is easy to define metrics that fall to zero in these situations. However, a dynamic offense is less easy to quantify because not all movement is productive. An offense involving endless passing around the perimeter certainly moves the ball, but does not constitute an effective attack. This is largely why quantifying the value of movement is difficult. In this paper, we suggest that these previous analyses of ball movement have failed because they misconstrue its purpose. We contend that a major purpose for ball movement is to inject unpredictability into the offense, which plays a significant game-theoretic role. Unlike distance traveled or average velocity, this unpredictability is not directly observable in the data; we present an **entropy** metric that quantifies the unpredictability of an offense at any given moment to fill this gap.

We also suggest that previous analyses have used the wrong measures of success. Wins, points, shooting percentages and other box-score summaries of offensive production depend highly on individual players' shooting abilities, and while ball movement can create favorable shooting situations, it cannot change a player's innate shooting ability. Because not all teams have players who can cash in on such opportunities, we propose instead that offensive movement should be judged based on its ability to create open shots, whether or not these shots are taken.

Our analysis of basketball offense, through the lens of entropy and opportunity, highlights the situations, players, and teams for which movement creates better offensive outcomes. We find that well-calibrated entropy leads to more and better offensive opportunity, although teams vary in their ability to turn those opportunities into points.

## 2 Game Theory and the Role of Randomness

Approaching basketball analysis from just one side of the ball, whether offense or defense, is naive because it ignores the fundamental game theoretic logic that underlies decision-making in adversarial situations. Thus, a realistic picture of basketball offense requires an analysis of how teams navigate the best responses of the defense to any strategy they are likely to play. Games that are entertaining to watch have a rock-paper-scissors structure where for each single strategy that one team could pursue, the other team has a definitive answer that can shut it down. This ensures that there is no single, fixed optimal strategy for either team to follow. To operate optimally in such a game, both teams need to randomize between their options so that occasionally the opponent will guess wrong and provide an opportunity that can be exploited. The optimal mix between offensive strategies depends on both how well the defense is able to predict the randomly chosen strategy and on how well-off each strategy leaves the offense, averaging over how often the defense is able to muster the appropriate response.

A full game theoretic analysis of basketball requires enumerating the full set of strategies, i.e. plays, that an offense can pursue and the set of defensive responses to such plays and calculating the expected value for each combination. This requires an enormous number of heroic assumptions unless one has access to both coaches' proverbial offensive and defensive playbooks. Here we have a more modest goal. We quantify the unpredictability of a team's offensive strategy without making statements about the optimal level of randomness the team should seek. Depending on the particular situation, the optimal level of randomness may be high or low. However, we do note that well-calibrated randomness plays a large role in teams' ability to perform better or worse than the sum of their parts and that for many underperforming teams, the instinct appears to be to under-randomize, although in some cases, over-randomization is a problem as well.

## 3 Markov Modeling: An Engine for Metrics

Throughout this paper, we apply a model of a basketball possession as a discrete-state, continuous-time Markov chain. We construct a simplified state space $S$ that represents relevant situational attributes of the offense's structure. We take $S$ as the combination of the identity of the ballcarrier, the identities of his teammates, the position in one of seven disjoint regions of offensive half-court (a diagram of this is included in the technical appendix), and an indicator of whether or not a defender is within 5 feet of him. We assume that, to first order, the future evolution of the possession only depends on its current state (the Markov assumption). Although this coarse representation of a possession drops a large amount of detail from the full-resolution data, it nonetheless allows us to capture many critical aspects of offensive mechanics.

Markov modeling has been popularly used to compute expected point values in baseball, football, and basketball. However, Markov models can be used to generate a large number of other metrics which, to date, have been largely overlooked with the exception of [3]. Other related works can be found in [1,2,5,6,8]. The metrics that follow are all derived from the Markov transition model. See the appendix for details on estimating and computing summaries from Markov models.

# 4 Entropy

We use **entropy** to quantify the unpredictability of an offense. Mathematically, entropy is defined in terms of a probability distribution, and measures how little information we have about an outcome before we see it. In basketball, the entropy at a given moment measures the unpredictability of the future evolution of the possession. If a fan takes a sip of their beer during a high entropy situation, they are likely to be disoriented when they look back at the play. If a defender loses focus for a second during a high entropy situation, he is less likely to recover. Even for an attentive defender, a high entropy situation forces him to hedge between the variety of offensive options he is likely to face.

In this paper, we define entropy as a situational property, unique to the state and personnel characteristics of a particular moment in a possession, and always changing as the game moves. Situations where a single player is likely to simply shoot or hold the ball without a credible threat of pursuing other options have low entropy, while situations where the player is equally likely to quickly pursue passing, driving, and shooting options have high entropy. Movement is necessary but not sufficient to create entropy -- performing the same set sequence of passes and runs every possession has the same average entropy as simply holding the ball. One of the highest entropy situations by our calculation is Kevin Garnett in the center 3 area. Specifically, Garnett rarely follows the same path in such a circumstance, and exploring what the possession looks like 1-2 seconds later, we rarely see the same outcome.

An example state is LeBron James undefended in the corner 3 region, with Chalmers, Wade, Allen, and Bosh on the court with him. Given any 5 player lineup, which we denote $T$, there are thus $5 * 7 * 2 = 70$ possible states the offense can occupy; we notate these using $s(i, r, d, T)$ where $i \in T$ is the ballcarrier, $r \in R$ is the court region he occupies, and $d = 1$ if he is closely defended and $0$ otherwise. For each of these states we calculate entropy using the probability distribution over states $s(i', r', d', T)$ the possession can reach one second after being in state $s(i, r, d, T)$. The probability of such a transition $s(i, r, d, T) \to S(i', r', d', T)$ occurring in one second we denote by $p_{i,r,d,T}(i', r', d')$. It's possible that state $s(i, r, d, T)$ can also transition into a shot attempt or turnover within a one-second time window, and we denote the probability of this as $p_{i,r,d,T}(end)$. These probabilities (which sum to 1) are used to calculate the entropy for each state,

$$ent(i, r, d, \mathcal{T}) = -\Big[ \sum_{\substack{i' \in \mathcal{T} \\ r \in \mathcal{R} \\ d \in \{0,1\}}} p_{i,r,d,\mathcal{T}}(i', r', d') \log(p_{i,r,d,\mathcal{T}}(i', r', d')) \Big] - p_{i,r,d,\mathcal{T}}(\text{end}) \log(p_{i,r,d,\mathcal{T}}(\text{end})).$$

This formula is identical to quantifications of unpredictability in many other scientific areas, such as Shannon entropy in information theory, Gibbs entropy in statistical thermodynamics, and von Neumann entropy in quantum physics. Entropy and closely related concepts have been previously applied to basketball analyses [3,9], however our calculation is the first to include spatio-temporal and defender information, and bases uncertainty on a richer set of decision possibilities, including holding the ball for a specific length of time in one area, and movement to different areas of the court by the same ballcarrier.

In addition to entropy, we also derive two metrics aimed primarily at player evaluation that are intuitive indicators of how players can reduce offensive entropy: **MicroUsage**, equal to the average probability that the ballhandler will shoot or turn the ball over in the next second, and **StickyScore**, equal to the average probability that the ballhandler will hold the ball in the same position for the next second. These quantities are functionally related to the entropy in a given situation, and having a high value for either quantity implies that a situation has low entropy.

# 5 Opportunity

The theoretical advantage of entropy is to put the defense off-balance, forcing them to hedge between preventing multiple high-value scoring situations. It cannot improve a shooter's ability to knock down an open shot. An appropriate evaluation of the value of entropy requires that we separate these two effects.

For each of the states $s(i, r, d, T)$ used in the Markov model, we quantify two distinct opportunities for offensive success. The first of these is immediate opportunity, which is the expected points resulting in a shot attempt from this state, $opp(i, r, d, T) = E(points \mid shot\, from\ s(i, r, d, T))$. Crucially, the immediate opportunity of any state depends on the skill of player $i$ shooting from region $r$. For instance, a state representing Kyle Korver in the center 3 area will have higher opportunity that one corresponding to Tyson Chandler in the center 3 area.

In addition to immediate opportunity, we calculate several measures of future opportunity based on the probability of reaching states that tend to be valuable for all offenses regardless of skill. Here, we focus on the probability of obtaining an open looks at the basket. Specifically, for each state $s(i,\ r,\ d,\ T)$ our measurements of long-run opportunity are:

- **Open look probability**: the probability of the offense eventually reaching a state consisting of a player undefended in a valuable shooting area.
- **Open corner 3 look probability**: the probability of the offense eventually reaching a state consisting of a player undefended in a corner 3 area.

The immediate opportunity $(opp)$ of a state derives from the corresponding player's shooting ability from that region, while measures of long-run opportunity derive from non-shooting transitions out of that state. However, the two quantities do not necessarily trade off each other. For instance, when a point guard such as Stephen Curry or Chris Paul possesses the ball in the center 3 area, because both are good shooters there is relatively high immediate opportunity, but they also orchestrate evolutions of the offense that yield good future opportunities (for themselves or others).

# 6 Results

Using the above metrics, we are able to identify several strong relationships between entropy and opportunity that persist across the league. We summarize these results below.

*Entropy and immediate opportunity trade off.* In general, we find that entropy and immediate opportunity trade off, and that this negative correlation persists within and between game situations and teams. This is not paradoxical, but rather intuitive. In situations with high immediate opportunity, the ballhandler has a high likelihood of scoring if he attempts a shot. In an efficient offense, he should be attempting such shots with high probability, and this behavioral predictability reduces the state's entropy value. Consistent with a game-theoretic perspective, the entropy/immediate opportunity tradeoff represents teams seeking an optimal mix between points obtained now against a defense that is prepared and points obtained later against a defense that is potentially off-balance.

Figure 2 summarizes the relationship between entropy and immediate opportunity across all ball-handling situations in the league. Each point in the figure shows the entropy and immediate opportunity when a particular player is handling the ball in a specific state, averaged over all lineups in which he appears. While data strongly cluster by state, the entropy/immediate opportunity tradeoff persists both within states and across states.

Figure 3 summarizes the aggregate relationship between entropy and immediate opportunity by team. A team's position in the entropy/opportunity landscape reveals a good deal about their offensive philosophy. Teams that have skilled shooters and emphasize a "good-to-great" philosophy, where players are encouraged to forego good shots if they see potential for great shot opportunities down the line, appear in the top right corner. Unsurprisingly, both San Antonio and Miami appear here. On the other hand, teams with offenses built around great contested-shot-makers, for example Golden State with Steph Curry and Portland with LaMarcus Aldridge, appear in the top-left corner. These teams, perhaps wisely, do not emphasize entropy because their offensive strategies often pay off even when the defense is able to mount the optimal response. Teams that lack star players who can capitalize on shooting situations appear toward the bottom of the plot. Notably, Philadelphia in the bottom right runs a variant of the Spurs offense, although with the available personnel, obtaining a "great" shot requires significantly more randomization.

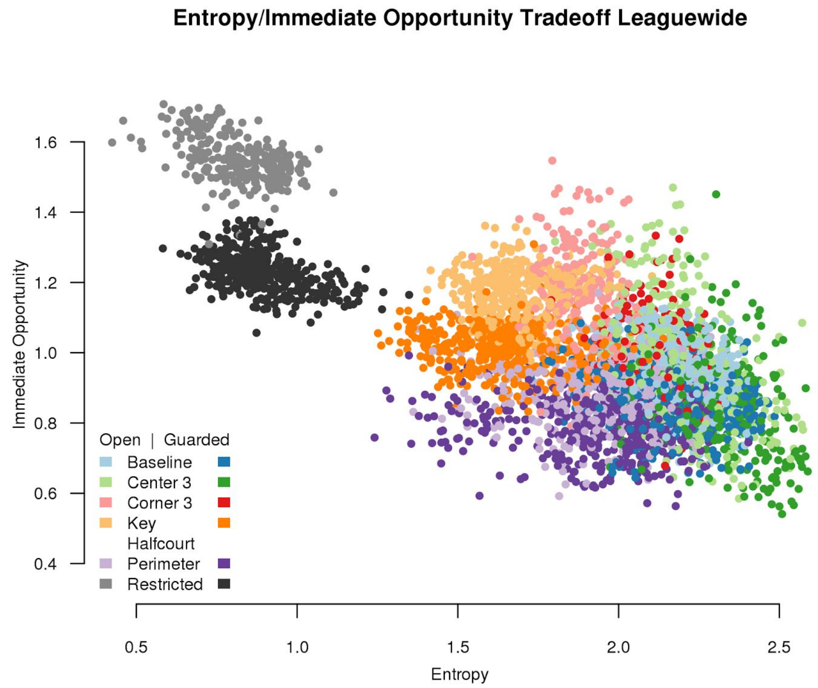



**Figure 2.** Summary of the relationship between entropy and immediate opportunity for all ballhandling situations in the NBA. Generally, entropy and immediate opportunity trade off.

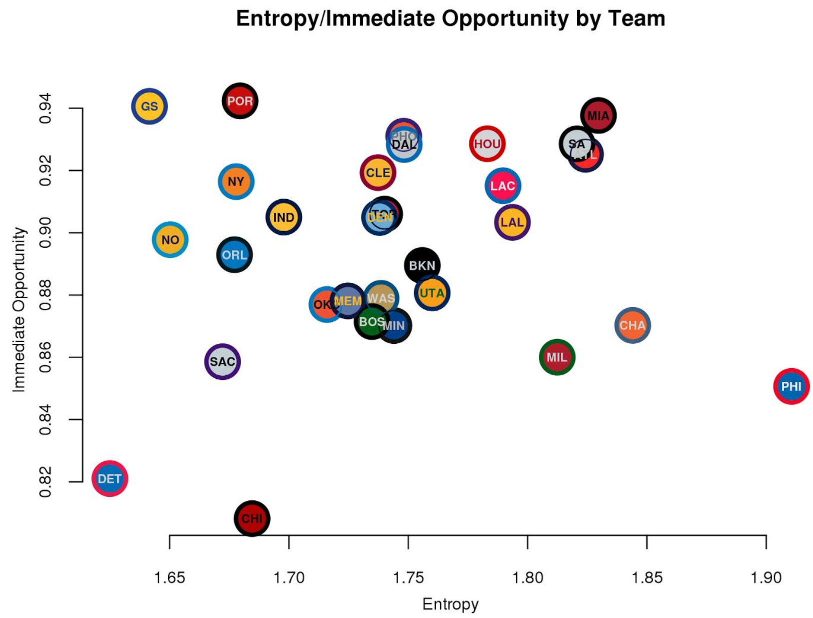



**Figure 3** Immediate opportunity and entropy aggregated by team. Different teams navigate this tradeoff in different ways. High entropy/high opportunity teams have "good-to-great" shooting philosophies while Low entropy/high opportunity teams rely on shot-makers who can reliably score contested shots.

*Entropy is associated with future opportunity.* Although entropy negatively correlated with immediate opportunity, it is highly positively correlated with future opportunity, reinforcing the notion that teams use ball movement to trade off immediately available points for more favorable opportunities down the line. Figures 4-7 summarize the relationships between entropy and our metrics open (probability of a future open shot) and cor3 (probability of a future open corner 3). As in Figure 1, the left plot summarizes the league-wide association for all ballhandlers in all situations, while the right plot summarizes the aggregate team-wise relationship.

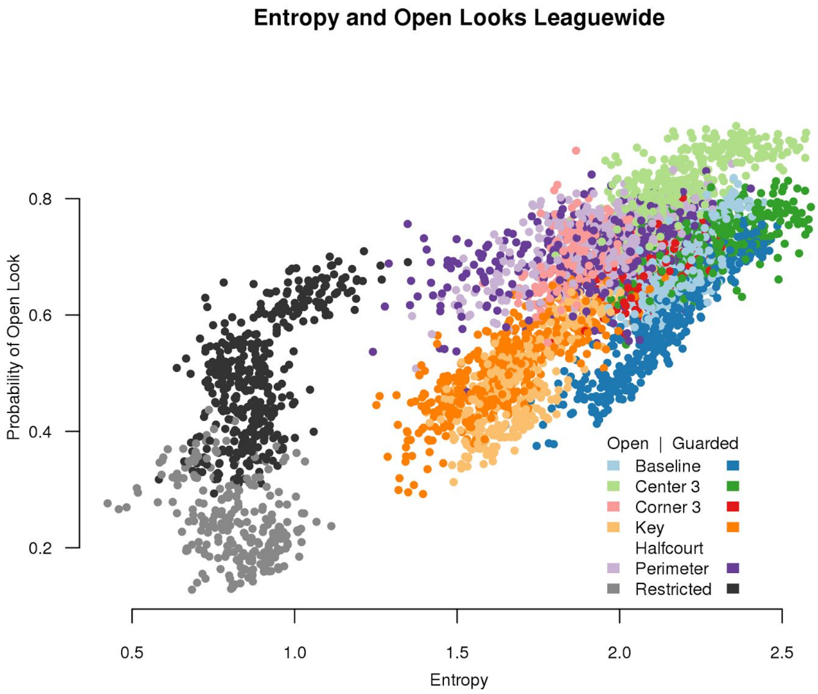



**Figure 5** Relationship of entropy and the probability that an offense will generate an open look in the future.. In general, these are positively correlated, exemplifying the game-theoretic value of entropy.

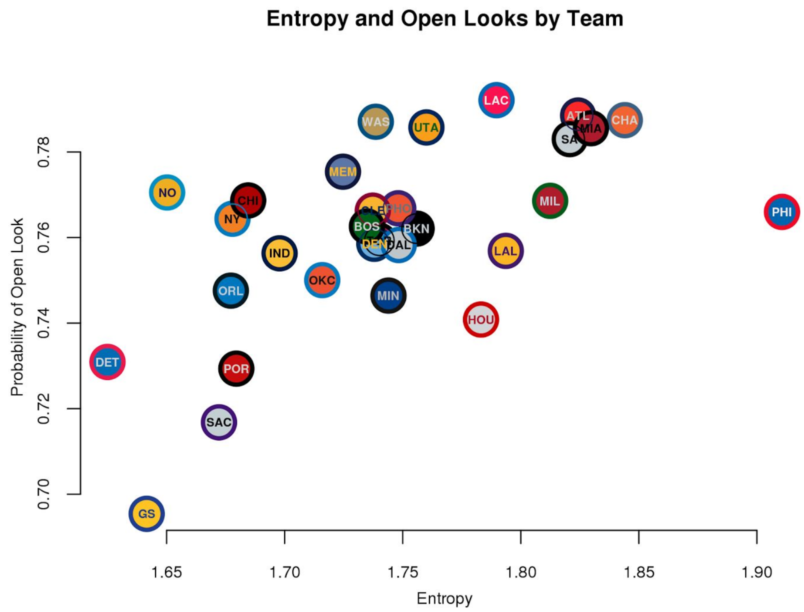



**Figure 5** Entropy versus open look probability by team. When correctly calibrated, high entropy teams can force defenses to mount suboptimal responses more often. Notably, the Philadelphia 76ers are one team that appears to have an offense that is too chaotic and does not induce the desired response from the defense.

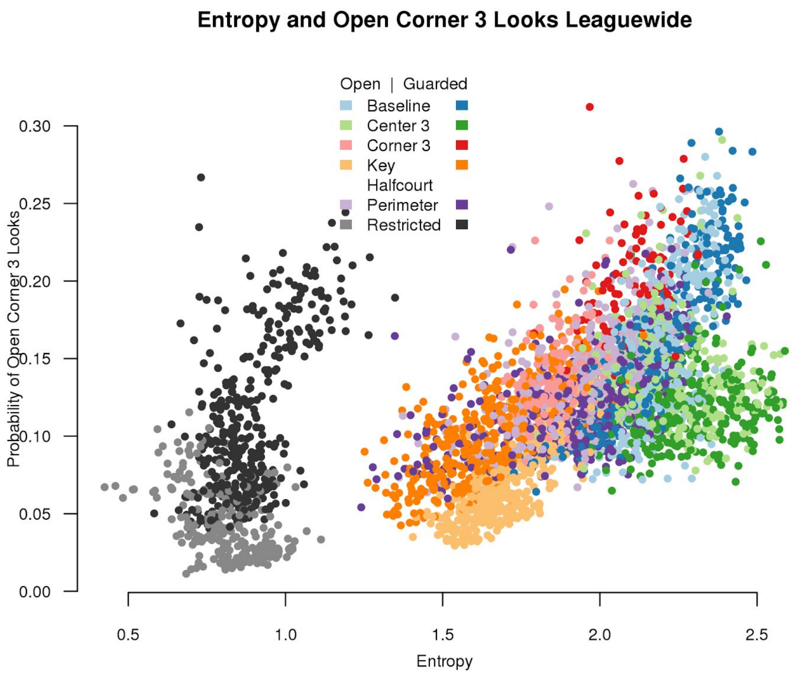



**Figure 6.** Relationship of entropy with a particularly appealing offensive situation, an open look at a corner 3. The positive correlation of entropy and this particular shot opportunity highlights the game theoretic value of unpredictability.

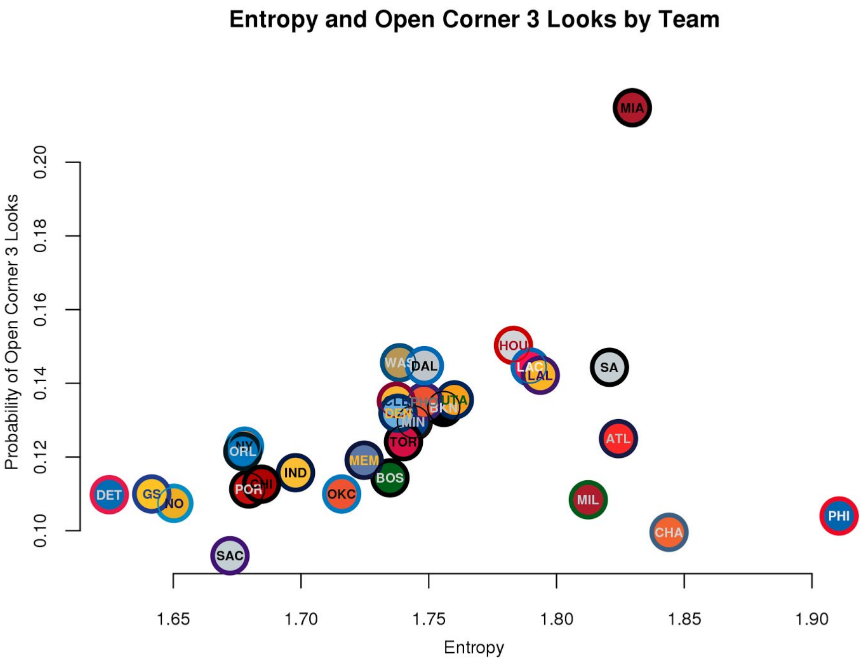



**Figure 7.** Entropy versus open corner 3 look probability by team. The Miami Heat are particularly adept at creating this situation.

.

# 6 Discussion

Ball movement has long been a mantra for basketball coaches at all levels, but has escaped effective quantification as a predictor for success, even in the era of high-resolution player tracking. In this paper, we present one reason that this may be the case -- ball movement is important because it introduces unpredictability into offenses, and effective unpredictability has a complex relationship with scoring-based outcomes. By focusing instead on metrics for opportunity creation independent of shot-making, we have presented a methodology that effectively quantifies the critical role of unpredictable ball-movement in basketball offense.

The complex relationship of ball movement with success is compounded by its role in the game theoretic tactics that underlie basketball. As we attempt more fine-grained analysis of the mechanics of basketball offense, these game theoretic issues, beyond simple observed outcomes, will become increasingly important. We hope that this work has shed light on this fact and that more detailed analyses of these "games within the game" can appear in future work.

# 5 References

[1] Bukiet, Bruce, Elliotte Rusty Harold, and José Luis Palacios. "A Markov chain approach to baseball." Operations Research 45.1:14-23 (1997)

[2] Cervone, Dan, et al. "POINTWISE: Predicting Points and Valuing Decisions in Real Time with NBA Optical Tracking Data." MIT Sloan Sports Analytics Conference 2014 (2014)

[3] Fewell, Jennifer H., et al. "Basketball teams as strategic networks." *PloS one* 7.11 (2012)

[4] Franks, Alexander, et al. "Characterizing the Spatial Structure of Defensive Skill in Professional Basketball." To Appear in the Annals of Applied Statistics (2014)

[5] Goldner, Keith. "A Markov model of football: Using stochastic processes to model a football drive." Journal of Quantitative Analysis in Sports 8.1 (2012)

[6] Kubatko, Justin, et al. "A starting point for analyzing basketball statistics."*Journal of Quantitative Analysis in Sports* 3.3 (2007)

[7] Miller, Andrew, et al. "Factorized Point Process Intensities: A Spatial Analysis of Professional Basketball." Proceedings of The 31st International Conference on Machine Learning. (2014)

[8] Omidiran, Dapo. "A new look at adjusted plus/minus for basketball analysis." MIT Sloan Sports Analytics Conference 2011 (2011)

[9] Piette, James, Lisa Pham, and Sathyanarayan Anand. "Evaluating basketball player performance via statistical network modeling." MIT Sloan Sports Analytics Conference, 2011.