Learning Fine-Grained Spatial Models for Dynamic Sports Play Prediction

Yisong Yue[*†1], Patrick Lucey[†], Peter Carr[†], Alina Bialkowski[‡†1], Iain Matthews[†]

[*]California Institute of Technology
yyue@caltech.edu
[†]Disney Research
{patrick.lucey,peter.carr,iainm}@disneyresearch.com
[‡]Queensland University of Technology
a.bialkowski@connect.qut.edu.au

Abstract—We consider the problem of learning predictive models for in-game sports play prediction. Focusing on basketball, we develop models for anticipating near-future events given the current game state. We employ a latent factor modeling approach, which leads to a compact data representation that enables efficient prediction given raw spatiotemporal tracking data. We validate our approach using tracking data from the 2012-2013 NBA season, and show that our model can make accurate in-game predictions. We provide a detailed inspection of our learned factors, and show that our model is interpretable and corresponds to known intuitions of basketball gameplay.

I. INTRODUCTION

Under what circumstances is a ball handler likely to take a shot or pass to a teammate? What defensive formations best deter a player's preferred actions? Such questions are of central importance in the study of decision making in team sports, or "sports analytics". In particular, we are interested in developing interpretable predictive models that can efficiently predict (or forecast) the outcomes of various game situations.

A grand challenge in predictive modeling is to capture the dynamics of adversarial gameplay in order to predict a complete sequence of events of how a game situation may unfold.[2] Within this challenge, one important first step is to predict the *immediate future*, i.e., the very next event. For example, we might be interested in predicting the probability that the ball handler will take a shot within the next second – in a sense, this prediction task corresponds to the "one-step lookahead" version of the general task of predicting complete sequences of gameplay. We consider developing models to predict near-term events (e.g., shots and passes) using fine-grained spatiotemporal tracking data, which has become a rapidly growing data source across many sports domains [5].

Developing predictive models using spatiotemporal data is challenging due to the lack of readily available semantically rich representations for encoding the game state. Although one could manually craft such a representation, it would be much more insightful to automatically infer one from data. For example, the propensity of the ball handler to take a shot likely depends on his/her position on the court, as well as that of the defenders. Similarly, the likelihood of passing to a teammate
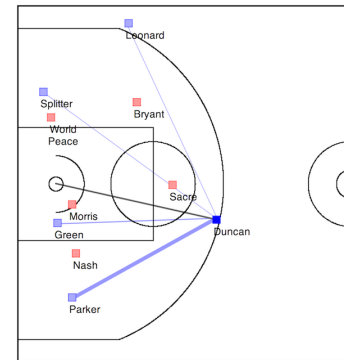

Fig. 1. Depicting probabilities that the ball handler (Duncan) will pass to his teammates (blue lines) or shoot (black line). Thicker lines indicates the higher probability. These predictions can be computed in real-time.[3]

likely depends on aspects such as the teammate's location and the opposition's coverage of the passing lane. Accurately modeling such in-game events requires a representation that can characterize the spatially and temporally varying behavior of the ball handler, the teammates, and the defenders.

In this paper, we take a discriminative learning approach to automatically train models to predict near-term game events given current game conditions. Building upon and combining discriminative learning techniques (such as Conditional Random Fields [6]) along with techniques for spatial regularization and non-negative matrix factorization [7], [8], we show how to leverage basic high-level domain knowledge of sports gameplay to train accurate predictive models while automatically inferring an interpretable feature representation. Our model learns and incorporates representations for several aspects of team gameplay, such as spatially varying shooting and passing tendencies in the presence of opposition, and how the duration of possessing the ball influences these tendencies.

An example of our near-term predictions for professional basketball is shown in Figure 1. This snapshot depicts the in-game predictions of the likelihood that the ball handler (Duncan) will pass to each teammate, as well as the probability that he will shoot the ball (thicker lines indicate higher likelihood). A good model should be able to make accurate predictions that vary smoothly (or sometimes sharply) depending on the locations of the ball handler, his teammates, and the opponents.

We validate our approach using spatiotemporal tracking data recorded from over 600 games of the 2012-2013 NBA

---

[1]This work was done while Yue and Bialkowski were at Disney Research.
[2]This challenge, or simpler variants of it, arises in many application domains such as Go [1], and is also of fundamental theoretical interest [2], [3], [4].

season. We verify that our model can learn to accurately predict near-term game events. We also analyze our learned factors, and show that they are interpretable and correspond to known intuitions of basketball gameplay. Due to the efficiency of our approach, we can make predictions in real-time, allowing for an interactive inspection of how shooting and passing probabilities change under varying game states.[3]

II. RELATED WORK

Leveraging spatiotemporal data in sports analytics has become increasingly popular in recent years due to the deployment of commercial tracking systems [9], [10], [11]. Compared to more typical settings for analyzing spatiotemporal data such crowd and surveillance domains [12], [13], analyzing behaviors in team sports is typically more complex as it is adversarial compared to cooperative, and is contingent on the interaction between and within groups compared to individuals.

Due to these challenges, most previous work in sports analytics have focused on relatively small datasets [14], [15], [16], did not build predictive models [17], [8], [18], used coarse aggregate statistics that do not model specific in-game scenarios [19], did not focus on adversarial team environments [20], do not model spatial information [21], [22], or require extensive manual annotation [5]. In contrast, we are interested in making in-game predictions of near-term events over a large selection of adversarial in-game scenarios.

A particularly illuminating paper by Miller et al. [8] showed how to incorporate spatial priors and non-negative matrix factorization to build a generative latent factor model in order to summarize basketball shot selection. They showed that their learned factors correspond to an interpretable representation that summarizes shot selection charts of NBA players. Indeed, non-negative matrix factorization techniques have been shown to yield interpretable data representations in a wide range of applications, including learning parts of objects [23], molecular pattern discovery [24], and overlapping community detection [25]. We will also build upon this technique in developing our approach.

Our approach also bears affinity to Bayesian techniques for spatial regularization such as Gaussian processes [26], [27], and their extensions [8]. In fact, our spatial regularization approach can be viewed as a frequentist analog to theirs.

From a dynamics perspective, our approach can be thought of as a one-step lookahead filtering method, whereby we estimate the probability of observing the next in-game event given the current game state. Recent work by Fewell et al. [16] estimated a transition graph of passes between players; however, their model was estimated at a rather coarse level of granularity and on a relatively small dataset, and did not take into account spatial properties such as the location of the players. As mentioned above, it the larger challenge would be to model entire sequences of gameplay.

The topics of greatest relevance to sports analysts are developing metrics to quantify the quality of a possession [28] or defensive prowess [18], and suggesting more optimal gameplay strategies using analytical approaches [22]. Our approach can be naturally integrated with such methods.

III. PREDICTING GAME EVENTS

We consider the problem of predicting future game event y given the current game state x. We develop our approach using basketball as the motivating application, although our approach can be applied to other group activities as well as modeling richer event classes within basketball. We assume the following information is provided in every game state $\mathbf{x}$:

- The identity b and location ℓ_b of the ball handler

- The identities and locations $\{(p_m, \ell_m)\}_{m=1}^4$ of the ball handler's four teammates

- The locations $\{\tilde{p}_m, \tilde{\ell}_m)\}_{m=1}^5$ of the five opponents

- The amount of time τ_b the ball handler has possessed the ball.

We are interested in predicting near-term events y from the ball handler such as:

- Passing to teammate p_i within the next t seconds

- Shooting the ball within the next t seconds

- Holding on to the ball (i.e., none of the above)[4]

When multiple events occur within the next t seconds (e.g., the ball handler passes to a teammate who then immediately shoots), we only consider the first such event (i.e., the pass). Given a training set $S = \{(y_i, \mathbf{x}_i)\}_{i=1}^N$, the goal is to learn a model that can accurately predict the correct y given $\mathbf{x}$.

A. Model

A natural approach is to model the response variable y using a multi-class conditional random field [6]. Let

$$Y(\mathbf{x}) = \{s, p_1, p_2, p_3, p_4, \perp\}$$

denote the space of possible predictions given $\mathbf{x}$, which corresponds to taking a shot, passing to one of four teammates, or none of the above, respectively. We can model the conditional probability of each event $y \in Y(\mathbf{x})$ as being log-linear w.r.t. a response function $F(y|\mathbf{x})$:

$$P(y|\mathbf{x}) = \frac{1}{Z(\mathbf{x}|F)} \exp\{F(y|\mathbf{x})\},$$

where $Z(\mathbf{x}|F)$ denotes the standard partition function:

$$Z(\mathbf{x}|F) = \sum_{y' \in Y(\mathbf{x})} \exp\{F(y'|\mathbf{x})\}.$$

Intuitively, each $F(y'|\mathbf{x})$ can be interpreted as the log-odds of event y' happening given game state $\mathbf{x}$.

Perhaps the most obvious aspects are the distances between the ball handler and the basket and his/her teammates. Shot difficulty increases with distance, so players are intuitively more likely to shoot closer to the basket. Similarly, players are more likely to pass the ball to teammates that are reasonably close to them. So one can instantiate a basic version of F as:

$$F(y|\mathbf{x}) = \begin{cases} F_s(\mathbf{x}) & \equiv w_s^\top \phi_s(b, \ell_b) & \text{if } y = s \\ F_p(i, \mathbf{x}) & \equiv w_p^\top \phi_p(p_i, \ell_i, b, \ell_b) & \text{if } y = p_i \\ F_\perp(\mathbf{x}) & \equiv w_\perp & \text{if } y = \perp \end{cases} , \quad (1)$$

[4]The ball handler is often dribbling and/or surveying the court in this case.

where ϕ_s and ϕ_p are feature mappings that characterize the distance between the ball handler and the basket and teammates, respectively (e.g., whether b is between 3-5 feet from the basket), and w_s, w_p and $w_\perp$ are parameters to be learned.

The simple model above ignores significant relevant information such as precise locations on the court with respect to specific players (certain players tend to shoot or receive passes at certain locations), and the locations of the defenders (blocking a passing lane could suppress the probability that the ball handler will pass to a teammate). However, crafting an appropriate feature representation to capture these semantics is a daunting and labor-intensive task. On the other hand, naively learning a linear model directly from raw spatial data leads to unreliable models that ignore known intuitions regarding basketball gameplay. Furthermore, shooting and passing data for individual players and locations is extremely sparse relative to the spatial granularity of the data.

Inspired by latent factor approaches to collaborative filtering and other problems with missing data [29], [21], [25] and in particular [8], we develop a spatially coherent latent factor model that yields a compact and interpretable data representation while offering good predictive performance. Using basic high-level domain knowledge of basketball gameplay, we define several intuitively useful views (or coordinate systems) of player locations and learn a factor model for each view.

B. Modeling Player Locations

1) Modeling Shooting: We begin with the simplest case of modeling where certain players tend to shoot the ball. Similar to [8], we first discretize the court into cells fine-grained enough to capture all significant spatial variations. We represent each player b using a K_s-dimensional latent factor B_b, and we represent each location ℓ also using a K_s-dimensional factor L_ℓ. Henceforth, we abuse notation and use L_ℓ to refer to the factor of the cell that position ℓ belongs to.

We can score the tendency of player b to shoot at location ℓ as $B_b^\top L_\ell$. Combining with the simple model in (1), our scoring function for shooting can be written as

$$F_s^\ell(\mathbf{x}) = w_s^\top \phi_s(b, \ell_b) + B_b^\top L_{\ell_b}, \qquad (2)$$

Intuitively, our use of latent factors follows from the assumption that the variability in how players shoot can be well captured by a low-rank projection (which is similar to [8]).[5] Since shot attempts for individual players is sparse, this modeling approach allows us to "share" information across players while still personalizing to each players' individual shooting tendencies. We will learn B and L from data.

2) Modeling Passing: When modeling where players are likely to receive passes, we consider two views of the data. First, similar to how we model shooting, we represent each player p_i using a K_p-dimensional latent factor P_i, and each location ℓ using a K_p-dimensional factor M_ℓ. The score for player p_i receiving a pass at location ℓ would then be $P_i^\top M_\ell$.

We next model pairwise location relationships independent of player identity. Let Q_{1,ℓ_1} and Q_{2,ℓ_2} denote K_q-dimensional factors representing passer and receiver locations, respectively.

---

[5]As shown in Figure 5, each of the K_s components in L correspond to a shooting archetype, and each player is represented as a linear combination.
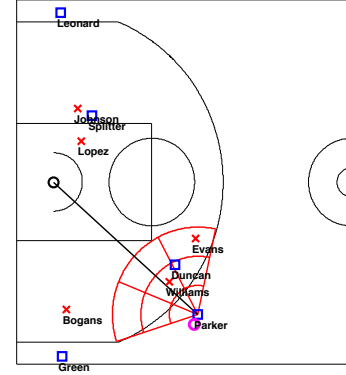

Fig. 2. Depicting how coordinate systems of defenders are transformed to be relative w.r.t. the ball handler. Here, Williams is calculated as being medium distance away and off right of Parker's path towards the basket.

We then model the score of a pass from ℓ_1 to ℓ_2 as $Q_{1,\ell_1}^\top Q_{2,\ell_2}$. Combining with the simple model in (1), our scoring function for passing can be written as

$$F_p^\ell(i, \mathbf{x}) = w_p^\top \phi_p(p_i, \ell_i, b, \ell_b) + P_i^\top M_{\ell_i} + Q_{1,\ell_b}^\top Q_{2,\ell_i}. \quad (3)$$

We will learn P, M, Q_1 and Q_2 from data.

C. Modeling Opponents

Beyond considering the locations of the ball handler and his teammates, it is also important to model the positions of the opponents that are defending. In order to accurately model opponents, it is important to model their *relative* positions with respect to the ball handler or a potential pass recipient.

Let $R(\mathbf{x}, \ell_1, \ell_2)$ denote a function that computes the relative positions of the defenders with respect to locations ℓ_1 and ℓ_2. Specifically, $R(\mathbf{x}, \ell_1, \ell_2)$ performs the following operations:

- rotates the coordinate system around position ℓ_1 until the direction $\ell_2 - \ell_1$ is the upward direction

- re-centers the coordinate system around ℓ_1

- applies this transformation to the locations of each defender and outputs their updated locations

Consider Figure 2, where the ball handler Parker is being defended by Williams. To compute the relative position of Williams w.r.t. to Parker and the basket, we can compute $R(\mathbf{x}, \ell_b, \ell_0)$, where ℓ_0 denotes the position of the basket. Williams's resulting relative position is computed as being a medium distance in front of Parker and slightly off-right.

1) Modeling Shooting: We first consider modeling the positions of defenders relative to the ball handler, which can suppress the probability of taking a shot. Let $\tilde{D}_{\tilde{\ell}}$ denote a K_d-dimensional latent factor corresponding to the relative location $\tilde{\ell}$ w.r.t. the ball handler, and let $\tilde{L}_{\ell_b}$ denote a factor representing the position of the ball handler. We can write the scoring function for how defenders impact shooting as:

$$F_s^d(\mathbf{x}) = \sum_{\tilde{\ell} \in R_b(\mathbf{x})} \left[v_{\tilde{\ell}} + \tilde{D}_{\tilde{\ell}}^\top \tilde{L}_{\ell_b} \right] + w_{d,s}^\top \phi_{d,s}(\mathbf{x}), \quad (4)$$

where $R_b(x) \equiv R(\mathbf{x}, \ell_b, \ell_0)$, $v_{\tilde{\ell}}$ is a global parameter independent of the location of the ball handler, and $\phi_{d,s}(\mathbf{x})$

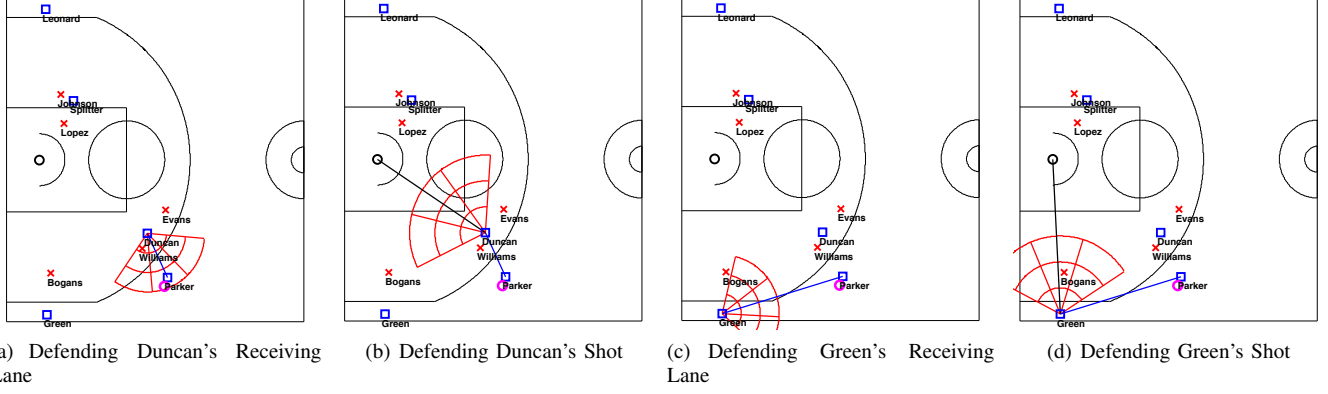| (a) Defending Duncan's Receiving Lane | (b) Defending Duncan's Shot | (c) Defending Green's Receiving Lane | (d) Defending Green's Shot |

Fig. 3. Depicting how coordinate systems of defenders are transformed to be relative w.r.t. a potential pass recipient. Figure (a) considers the passing direction from Parker to Duncan, and Williams is considered to be near Duncan and to his right. Figure (b) considers the shooting direction of Duncan, and no defenders are considered to be in front of him. Figure (c) considers the passing direction from Parker to Green, and Bogans is considered to be somewhat near Green but significantly to the left. Figure (d) considers the shooting direction of Green, and Bogans is considred to be relatively close and directly in front of Green.

computes a set of distance features, such as the distance of the closest defender to the ball handler. Intuitively, the way which opponents will defend a player will vary depending on location. For example, opponents tend to defend much more closely when the ball-handler is closer to the basket. Our final scoring function for shooting events can now be written as

$$F_s(\mathbf{x}) = F_s^\ell(\mathbf{x}) - F_s^d(\mathbf{x}), \tag{5}$$

for F_s^ℓ defined in (2).

2) Modeling Passing: We next consider modeling the positions of defenders relative a potential pass recipient teammate, which can suppress the probability of the ball handler passing that teammate the ball. We consider two sets of views or coordinate systems here. The first view models defenders relative to the direction of the pass, and the second view models defenders relative to the direction of the basket (which is similar to the defender model for shooting).

Figure 3 depicts examples of these two views. Figures 3(a) and 3(c) show the first view with the defenders relative to the direction of the pass. Figures 3(b) and 3(d) show the second view with the defenders relative to the recipient's path towards the basket. Having defenders blocking the foward direction in either view can contribute to reducing the probability that the ball handler will pass to that teammate.

Let $\tilde{C}_{1,\tilde{\ell}}$ and $\tilde{C}_{2,\tilde{\ell}}$ denote the latent factors representing the relative position $\tilde{\ell}$ of the defender w.r.t. the two views. Let $\tilde{M}_{1,\ell}$ and $\tilde{M}_{2,\ell}$ denote the latent factors representing the location of the pass recipient on the court. We can write the scoring function for how defenders impact a teammate from receiving a pass as:

$$F_p^d(i, \mathbf{x}) = \sum_{\tilde{\ell} \in R_1(i,\mathbf{x})} \left[\tilde{v}_{1,\tilde{\ell}} + \tilde{C}_{1,\tilde{\ell}}^\top \tilde{M}_{1,\ell} \right] \tag{6}$$

$$+ \sum_{\tilde{\ell} \in R_2(i,\mathbf{x})} \left[\tilde{v}_{2,\tilde{\ell}} + \tilde{C}_{2,\tilde{\ell}}^\top \tilde{M}_{2,\ell} \right] \tag{7}$$

$$+ w_{d,p}^\top \phi_d(i, \mathbf{x}), \tag{8}$$

where $\tilde{v}_{1,\tilde{\ell}}$ and $\tilde{v}_{2,\tilde{\ell}}$ are global parameters, and ϕ_d is a feature function that computes a set of distance features, such as the

distance between the defenders and the teammate and the ratio of the passing distance to the defender's distance. This results in the scoring function for passing events:

$$F_p(\mathbf{x}) = F_p^\ell(\mathbf{x}) - F_p^d(\mathbf{x}), \tag{9}$$

for F_p^ℓ defined as in (3).

D. Modeling Duration of Possession

We finally consider how player behavior varies depending on duration of possession. For example, at some locations, a player may often choose to immediately shoot or pass upon receiving the ball. Elsewhere, a player may often choose to keep the ball for a few seconds before shooting or passing.

Let $T_{s,\tau}$, $T_{s,\tau}$, and $T_{\perp,\tau}$ denote latent factors corresponding to the ball handler having had possession for τ time steps prior to the current state. Let V_ℓ denote a latent factor representing the location of the current ball handler. Then we can write the contribution of duration of possession to shooting, passing, or holding on to the ball, respectively, as:

$$
\begin{aligned}
F_s^\tau(\mathbf{x}) &= T_{s,\tau_b}^\top V_{\ell_b}, \\
F_p^\tau(\mathbf{x}) &= T_{p,\tau_b}^\top V_{\ell_b}, \\
F_\perp^\tau(\mathbf{x}) &= T_{\perp,\tau_b}^\top V_{\ell_b},
\end{aligned}
\tag{10}
$$

and add them to their corresponding scoring functions in (1).

E. Learning Objective

We now describe our learning criterion to be optimized during model fitting. Let Θ denote all of the parameters of our model (roughly 100-200 thousand, depending on the number of factors). We formulate a discriminative learning objective to minimize a trade-off between the negative conditional log likelihood of the training data and a regularization term controlling the complexity of our model:

$$\underset{\Theta \in \mathcal{C}}{\operatorname{argmin}} \, \mathcal{L}(\Theta) \equiv \sum_i -\log P(y_i|\mathbf{x}_i) + \lambda R(\Theta), \tag{11}$$

where $\mathcal{C}$ denotes the feasible space of model parameters, $R(\Theta)$ denotes the regularization function, and λ is a hyperparameter that trades off between the two.

We define $R(\Theta)$ using two components that decompose additively. The first component is the standard squared 2-norm that encourages parameter weights to be small. The second component is a spatial regularization term that encourages parameters of nearby cells to be similar to each other.

1) Spatial Regularization: As an example, consider the latent factors L representing the locations on the court that players tend to shoot from (see Section III-B). We can write the regularization term as

$$R_L(L) = \sum_{\ell \neq \ell'} \kappa_{\ell,\ell'} \|L_\ell - L_{\ell'}\|^2, \tag{12}$$

where $\kappa_{\ell,\ell'}$ controls for the degree of similarity between ℓ and ℓ'. (12) has the effect of encouraging the spatial parameters to vary smoothly. The similarity coefficients κ can be chosen using any kernel function. We use the RBF kernel:

$$\kappa_{\ell,\ell'} = \exp\left\{\|\ell - \ell'\|^2/\sigma\right\},$$

where σ is a tunable paramter controlling the degree of smoothness. We adopt an identical approach for regularizing all of our spatially coherent factors, L, M, Q_1, Q_2, $\tilde{L}$, $\tilde{M}_1$, $\tilde{M}_2$, and V. This approach bears affinity to generative models that employ a Gaussian process prior [26], [27], [21], [8].

2) Non-Negativity Constraints: Our feasible space of model parameters $\mathcal{C}$ requires that all of the latent factor components must be non-negative. This constraint makes our approach analogous to related approaches based on non-negative matrix factorization to derive interpretable models [23], [24], [25]. For instance, non-negative latent factor models have already been applied to basketball analysis for summarizing shot charts of NBA players [8].

Note that the form of our scoring functions in (5), (9), and (10) encode domain knowledge about which factors should contribute positively or negatively to various events. In particular, we model player locations as positively influencing the probability an event will happen, and defender locations as negatively influencing the probabilities. As we will see in Section VI, this leads to wide range of interpretable factors that correspond to known intuition of basketball game play.

IV. TRAINING

We train our model using gradient descent. For brevity, we only present derivations for two of our latent factors. The other partial derivatives can be derived analogously. We first derive the partial derivative for factor L_ℓ, which models whether the ball handler is likely to shoot at location ℓ:

$$\frac{\partial \mathcal{L}}{\partial L_\ell} = 2\lambda \left(L_\ell + \sum_{\ell' \neq \ell} \kappa_{\ell,\ell'}(L_\ell - L_{\ell'}) \right) - \sum_{i \in I_\ell^b} \frac{\partial \log P(y_i|\mathbf{x}_i)}{\partial L_\ell},$$

where I_ℓ^b denotes training data indices where the ball handler was in cell ℓ. For latent factors that do not have spatial regularization (e.g., B), there does not exist an analogous term for $\kappa_{\ell,\ell'}(L_\ell - L_{\ell'})$.

Let b_i denote the ball handler in the i-th training example. The partial derivative of conditional probabilities is:

$$\frac{\partial \log P(y_i|\mathbf{x}_i)}{\partial L_\ell} = \left(\mathbf{1}_{[y_i=b_i]} - P(b_i|\mathbf{x})\right) \frac{\partial F_s(\mathbf{x})}{\partial L_\ell},$$

where $\mathbf{1}_{[cond]}$ is an indicator function that takes value 1 if *cond* is true, and 0 otherwise. We finally observe that

$$\frac{\partial F_s(\mathbf{x}_i)}{\partial L_\ell} = B_{b_i}.$$

All other partial derivatives for parameters appearing in a shooting event can be derived analogously.

We next derive the partial derivatives for M_ℓ, which models whether a teammate at location ℓ is likely to receive a pass:

$$\frac{\partial \mathcal{L}}{\partial M_\ell} = 2\lambda \left(M_\ell + \sum_{\ell' \neq \ell} \kappa_{\ell,\ell'}(M_\ell - M_{\ell'}) \right) - \sum_{i \in I_\ell^p} \frac{\partial \log P(y_i|\mathbf{x}_i)}{\partial M_\ell},$$

where I_ℓ^p denotes training data indices where a teammate was in cell ℓ.

Let p_i denote the teammate in the i-th training example that was in cell ℓ. We can write the partial derivations w.r.t. each training example as:

$$\frac{\partial \log P(y_i|\mathbf{x}_i)}{\partial M_\ell} = \left(\mathbf{1}_{[y_i=p_i]} - P(p_i|\mathbf{x})\right) \frac{\partial F_p(p_i, \mathbf{x})}{\partial M_\ell}.$$

We finally observe that:

$$\frac{\partial F_p(p_i, \mathbf{x}_i)}{\partial M_\ell} = P_{p_i}.$$

All other partial derivatives for parameters appearing in a passing event can be derived analogously. Any gradient descent algorithm is then applicable. For our experiments, we used adaptive stochastic gradient descent [30].

A. Initialization

We find that our learning objective (which is non-convex) is particularly succeptible to local optima, and running gradient descent algorithms from random initializations leads to bad solutions. As such, finding a good initialization is key.

To find a good initialization, we first train our latent factor models as a full rank matrix of parameters. For example, instead of modeling the player by location shooting factors B and L as separate factors, we instead directly estimate the full player-by-location matrix of parameters $\bar{B} \equiv B^\top L$.[6]

Let b_i denote the ball handler in the i-th training instance, and let $\bar{I}_{ell}^b$ denote the training indices where ball handler was in location ℓ. We can write the partial derivative for each $\bar{B}_{b\ell}$:

$$\frac{\partial \mathcal{L}}{\partial \bar{B}_{b\ell}} = 2\lambda \left(\bar{B}_{b\ell} + \sum_{\ell' \neq \ell} \kappa_{\ell,\ell'}(\bar{B}_{b\ell} - \bar{B}_{b\ell'}) \right) - \sum_{i \in \bar{I}_\ell^b} \frac{\partial \log P(y_i|\mathbf{x}_i)}{\partial \bar{B}_{b\ell}},$$

and

$$\frac{\partial \log P(y_i|\mathbf{x}_i)}{\partial \bar{B}_{b\ell}} = \left(\mathbf{1}_{[y_i=b_i]} - P(b_i|\mathbf{x})\right) \frac{\partial F_s(\mathbf{x})}{\partial \bar{B}_{b\ell}},$$

and finally

$$\frac{\partial F_s(\mathbf{x})}{\partial \bar{B}_{b\ell}} = 1.$$

All other partial derivatives for combined factors can be derived analogously. This modification essentially decouples

---

[6]Each row of $\bar{B}$ corresponds to a player, and each column to a location.

(a) Predictive Performance
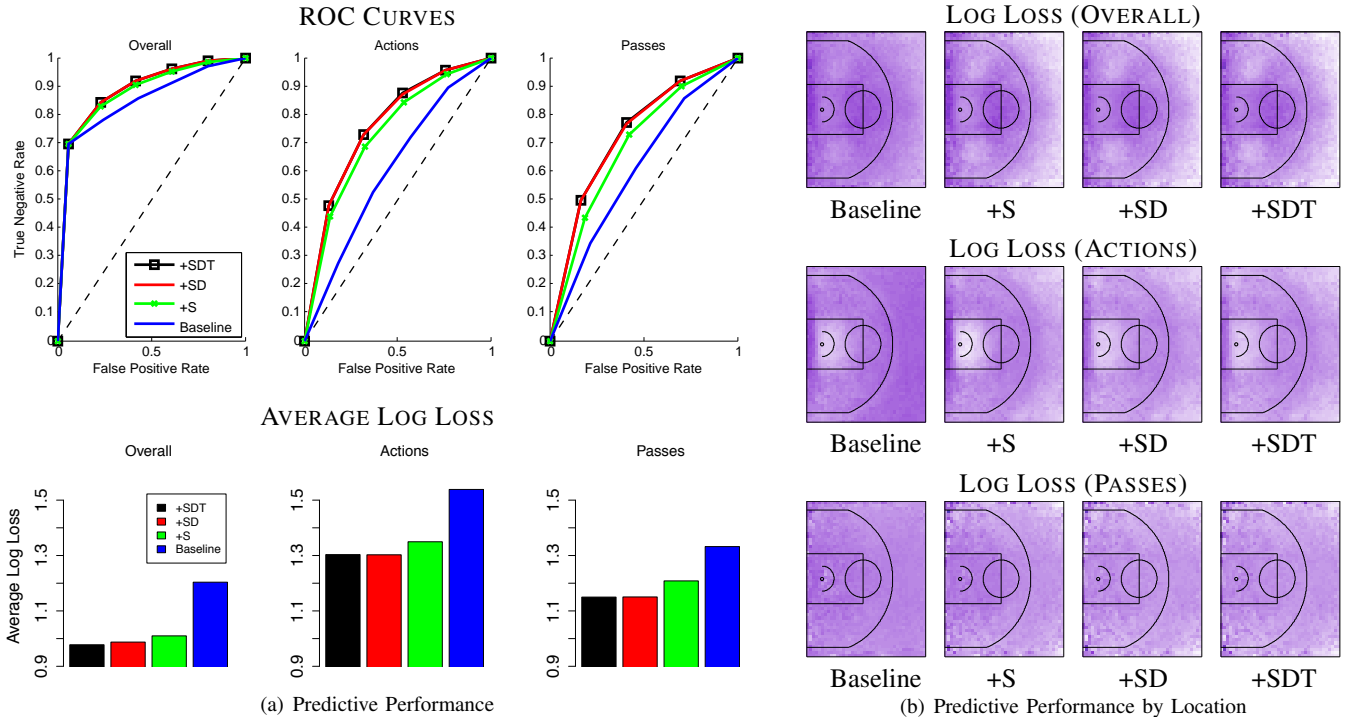
(b) Predictive Performance by Location

Fig. 4. *Left*: The left plots depict the aggregate results (ROC curves and log loss) for predicting all events (Overall), conditioned on actions (Actions), and conditioned on passes (Passes). For ROC curves, random prediction corresponds to the diagonal dotted line. For log loss, random prediction corresponds to $-\log(1/|Y|)$, where $|Y|$ is the number of possible outcomes (6 for Overall, 5 for Actions, 4 for Passes). Our spatial factor models consistently improve performance, although the gain of the temporal model (**+SDT**) is small. The ROC curves in the Overall task are similar, which is due to the majority of events being the null event of no action taking place; the Actions and Passes tasks offer a more nuanced comparison. *Right*: The right plots depict the spatially varying predictive performances (darker corresponds to higher error). Section VI offers insight on which components of our model leads to the various error reductions.

the interactions between players, since each row of $\bar{B}$ now does not interact with other rows of $\bar{B}$. As a consequence, the resulting optimization problem is convex and can be solved optimally,[7] at the expense of the learned model being statistically unreliable due to each individual player/location pair having very few data points.

After learning $\bar{B}$, we employ non-negative matrix factorization to recover a good initialization of B and L. Following the approach in [7], [8], our goal is find a B and L in order to minimize the matrix divergence:

$$\sum_{b,\ell} \bar{B}_{b\ell} \log \frac{\bar{B}_{b\ell}}{B_b L_\ell} - \bar{B}_{b\ell} + B_b L_\ell,$$

which discourages large ratios between $\bar{B}$ and the recovered factors B and L. We use the multiplicative updates approach from [7] to compute the factorization. Afterwards, we use the recovered B and L to initialize (11). We apply the same technique to initalize all of our latent factors. Since we use this approach only for initialization, we can learn each pair of factors independently while holding the other parameters fixed, which leads to substantial computational savings.

V. EXPERIMENTS

We evaluate our model based on STATS SportsVU tracking data from the 2012-2013 NBA season. The tracking data

includes coordinates of all players recorded at 25 frames per second. Also included are in-game events such as passes and shots. In total, our dataset comprises approximately 630 games.

We extracted every half court possession from each game, where a half court possession is defined to be when at least 4 players from each team have crossed into half court for at least four seconds. We focus on half court possessions since those are most likely to reveal in-game dynamics.[8] In total we extracted about 80 thousand possessions, with possessions averaging about 380 frames or 15 seconds in duration.

We selected 80% of the possessions at random (from each team) for training and validation, and used the remaining 20% for testing. Given the size of the dataset, we kept the regularization parameter λ at a fixed value and considered a few different kernel bandwidths during validation. We considered four variants of our model for comparison.

- **Baseline**: the simple model (1) based on distance features.
- **+S**: adds on top the spatial tendencies described in Section III-B.
- **+SD**: adds on top the defensive spatial tendencies described in Section III-C.
- **+SDT**: adds on top the temporal spatial tendencies described in Section III-D.

A. Results

Figure 4(a) shows our aggregate quantitative results. The results of predicting over all time frames are shown in the

---

[7]Since this intermediate objective is strongly convex, we employ accelerated gradient descent [31] to quickly converge to the optimal initialization point.

[8]The other major type of possession is a fast break.

Shooting Factors (L)

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |

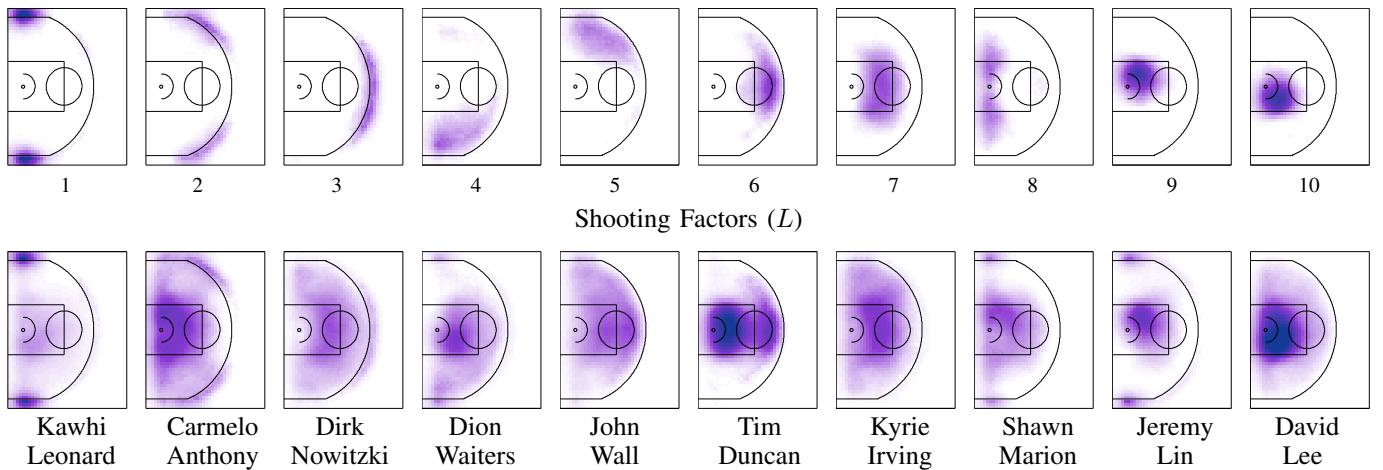| Kawhi Leonard | Carmelo Anthony | Dirk Nowitzki | Dion Waiters | John Wall | Tim Duncan | Kyrie Irving | Shawn Marion | Jeremy Lin | David Lee |

Fig. 5. *Top Row*: Depicting the spatial coefficents of the latent factors corresponding to the spatial coefficients contributing to the ball handler taking a shot at each location (L). Each player's shooting model can be represented as a combination of these factors. *Bottom Row*: Depicting the spatial coefficients of a player with high affinity to each of the latent factors.
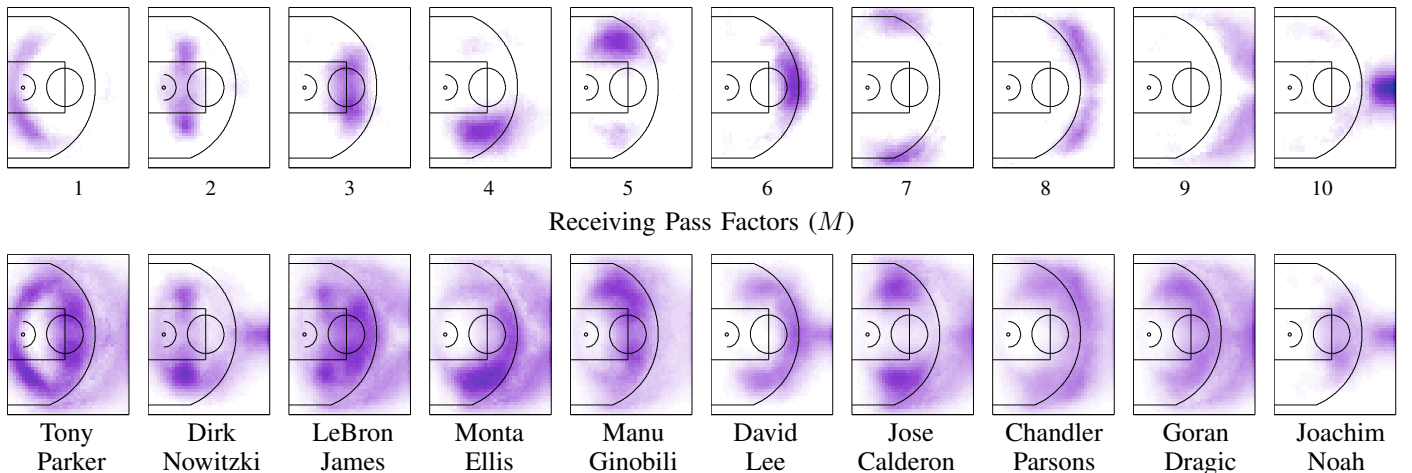

Receiving Pass Factors (M)

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |

| Tony Parker | Dirk Nowitzki | LeBron James | Monta Ellis | Manu Ginobili | David Lee | Jose Calderon | Chandler Parsons | Goran Dragic | Joachim Noah |

Fig. 6. *Top Row*: Depicting the spatial coefficents of the latent factors corresponding to spatial coefficients contributing to a player receiving a pass at each location (M). Each player's conditional shooting probabilities can be represented as a combination of these factors. *Bottom Row*: Depicting the spatial coefficents of a player with high affinity to each of the latent factors.

left column. We evaluated using both ROC curves as well as average log loss, and we observe that our latent factor models offer significant improvement. The gain from adding the temporal spatial components is small (i.e., **+SDT** offers about a 0.1% relative reduction in average log loss compared to **+SD**), but this gain is still statistically significant given the approximately 6 million examples/frames in our test set.

Our dataset exhibits a notable amount of class imbalance – about 70% of frames have no action events forthcoming in the next 1 second. To further tease out the interesting effects learned by our model, we also evaluate predictive performance when conditioned on an action event (shot or pass), and also on a pass event. These results are depicted in the second and third columns of Figure 4(a), respectively. We observe a more significant performance gain from our spatial latent factor models compared to the baseline, although the performance of **+SDT** is no longer distinguishable from **+SD**.

Figure 4(b) depicts how predictive performance varies across the basketball court. We see that the spatial latent factor model improves performance throughout the court. When

comparing all events (top row), we observe a significant error reduction in the two "low post" locations.[9] When comparing action events (middle row), we observe a significant error reduction along the baseline behind the basket (for **+SD** and **+SDT**). We comparing only passing events (bottom row), we observe a significant error reduction around the basket area (for **+SD** and **+SDT**). Our inspection of the learned factors in Section VI will shed light on how components of our model contribute to the various reductions in error.

VI. MODEL INSPECTION

We now provide a detailed inspection our learned model in order to identify patterns that match known intuitions of basketball game play. We will also identify aspects of our model that contribute to the various gains in performance depicted in Figure 4(b). Although our focus is on basketball, this type of analysis is applicable to a wide range of domains with fine-grained spatial patterns.

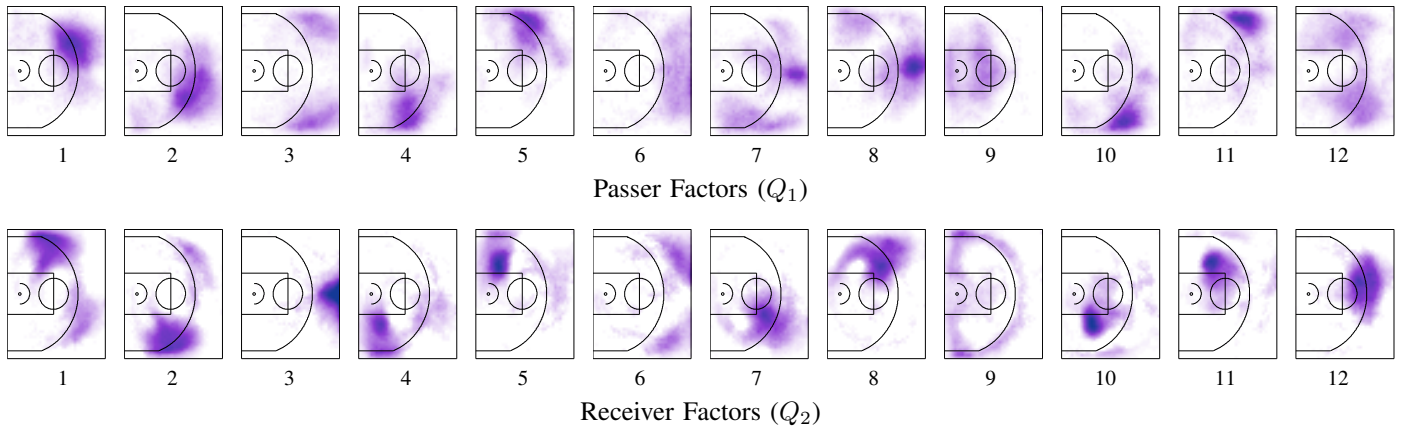[9]The two low post locations are situated about 10 feet and directed at 45 degree angles away from the basket.

Fig. 7. *Top Row*: Depicting the spatial coefficents of the latent factors corresponding to passing from each location. *Bottom Row*: Depicting the associated spatial coefficents of the latent corresponding to receiving a pass at each location. Figure 8 shows examples of aggregate passing and receiving coefficients.

A. Where do Players Tend to Shoot?

We begin by inspecting the player-by-location shooting factors B and L. We estimated a 10-dimensional representation, which visually correspond to known shooting patterns. Our results are shown in Figure 5. The top row depicts the ten learned factors (i.e., the rows of L), and the bottom row depicts the spatial coefficients of players which have high affinity for each of the factors. The first three factors correspond to players who shoot at the 3 point line. The next three factor correspond to players who shoot midrange shots. The seventh factor corresponds to players who shoot just inside the lane. The eight factor corresponds to players who shoot while attacking the basket from the baseline. The final two factors correspond to players who shoot near the basket.

We note that the first factor and the last two factors have the strongest coefficients, which is unsurprising since those are the most valuable shooting locations. Note that the first factor also corresponds to the large error reduction in the "corner three" locations in the middle row of Figure 4(b). We further note that none of the factors have strong coefficients along the baseline behind the basket, which is also a region of large error reduction in the middle row of Figure 4(b).

A similar analysis was conducted by [8], but with two key differences. First, [8] did not consider the in-game state but only modeled the shot selection chart. In contrast, we are interested in modeling the conditional probability of a player shooting given the current game state. For example, a player may often occupy a certain location without shooting the ball. Second, as described below, we also model and analyze many more spatial factors than just player shooting tendencies.

B. Where do Players Tend to Receive Passes?

We now inspect the player-by-location passing factors P and M. We estimated a 10-dimensional representation, which visually correspond to known passing patterns. The top row in Figure 6 shows the ten learned factors (i.e., the rows of M), and the bottom row depicts the spatial coefficients of players which have high affinity for each of the factors.

The first factor corresponds to players that receive passes while cutting across the baseline. In fact, many of the factors give off a visual effect of being in motion (i.e., spatial blur),
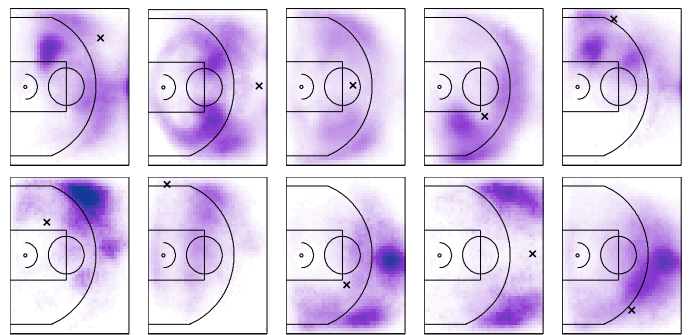

Fig. 8. *Top Row*: Depicting spatial coefficients of passing to different locations on the court. The "X" denotes the location of the passer and corresponds to a row in $Q_1^\top Q_2$. *Bottom Row*: Depicting the spatial coefficients of receiving a pass from different locations on the court. The "X" denotes the location of the receiver, and corresponds to a column in $Q_1^\top Q_2$.

which suggests that players are often moving right before receiving a pass.[10] The second factor corresponds to players that receive passes in the low post, which also corresponds to areas of high error reduction in the top row of Figure 4(b).[11] The next three factors depict various midrange regions. Typically, right handed players associate more with the fourth factor, and left handed players associate more with the fifth factor.[12] The next three factors depict regions at the three-point line, and the final two factors depict regions in the back court.

The players with the strongest coefficients tend to be from playoff teams that have tracking equipment deployed on their home courts during the 2012-2013 NBA season, since those players have the most training data available for them. For example, we see that Figure 5 and Figure 6 contain several players from the San Antonio Spurs, which had the most games in our dataset.

---

[10]This effect is most likely an artifact of how we formulated our task to predict events up to 1 second into the future. It may be beneficial to model multiple time intervals into the future (e.g., half a second, one second).

[11]Note that although players often receive passes in the low post, they typically do not shoot from that location, as evidenced in the shooting factors in Figure 5. Capturing this sharp spatial contrast between receiving passes and shooting contributes to the error reduction in those regions in Figure 4(b).

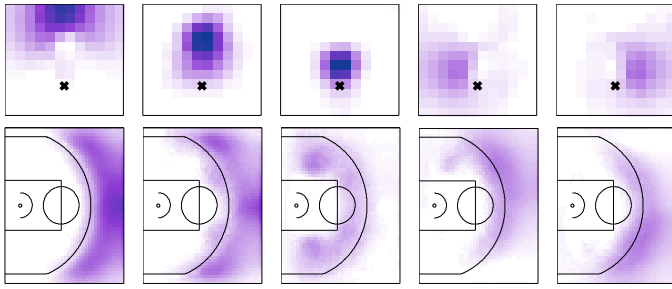[12]This effect is due to players preferring to drive with their strong hand.

Fig. 9. *Top Row*: Depicting the spatial coefficients of the latent factors $\tilde{D}$ corresponding to how defender positions relative to the ball handler suppresses the conditional probability that the ball handler will shoot. The "X" denotes the ball handler, and the up direction corresponds to toward the basket. *Bottom Row*: Depicting the corresponding spatial coefficents on the court $\tilde{L}$.

C. Where Does the Ball Tend to Get Passed To and From?

We now inspect the twelve Q_1 and Q_2 factors which correspond to pairwise passer and teammate locations. The top row of Figure 7 depicts the factors Q_1 representing locations of the passer, and the bottom row depicts the associated factors Q_2 representing the locations of the pass recipient.[13] We observe that the passing factors tend to exhibit more local coherence, which fits with our intuition that players tend to be stationary when passing. In contrast, we see that the receiving factors exhibit more complex spatial structure. For example, the seventh and eighth factors depict players who receive passes after cutting across the baseline from behind the basket. These results suggest a direction for future work: modeling not only the current game state but also the recent trajectories between the current state and the previous game states.

Figure 8 depicts the aggregate spatial coefficients $Q_1^\top Q_2$ of passing and receiving at a few example locations on the court. Interestingly, one could interpret $Q_1^\top Q_2$ as analogous to the transition matrix of a Markov chain random walk in terms of passing the ball. It would be interesting to combine this approach with the approach in [16] to model and analyze the flow of basketball passes during a possession.

D. What Defensive Positions Suppress Shot Probability?

We now inspect the defender factors, starting with the shot suppression factors $\tilde{D}$ and $\tilde{L}$. Figure 9 depicts the five factors we estimated here. The top row corresponds to the relative positioning of defenders around the ball handler ($\tilde{D}$), and the bottom row corresponds to the locations on the court that are most affected by this kind of defensive positioning ($\tilde{L}$).

We see that the $\tilde{D}$ factors naturally correspond to different defensive positionings, which is unsurprising since defenders are known the defend more closely the closer the ball handler is to the basket. We also note a complete lack of spatial coefficients in the region around the basket, indicating that defensive positioning does not suppress shot probability near the basket[14] – this effect also contributes indirectly to the error reduction in the bottom row of Figure 4(b), since it allows for estimating stronger coefficients elsewhere in the court and then suppressing those scores through the defender factors.

---

[13]Players located in areas where the passing factor is strong tend to pass to players located where the associated receiving factor is strong.

[14]Although it certainly affects the probability of making the shot.

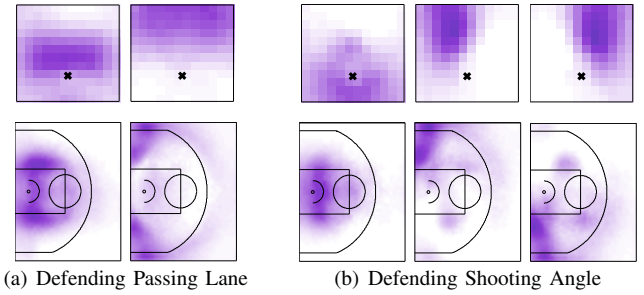(a) Defending Passing Lane   (b) Defending Shooting Angle

Fig. 10. Depicting spatial coefficients corresponding to how the positions of defenders relative to a player can suppress the probability of that player receiving a pass. The "X" denotes the ball handler. In the left plots, the up direction is towards the passer. In the right plots, the up direction is towards the basket.
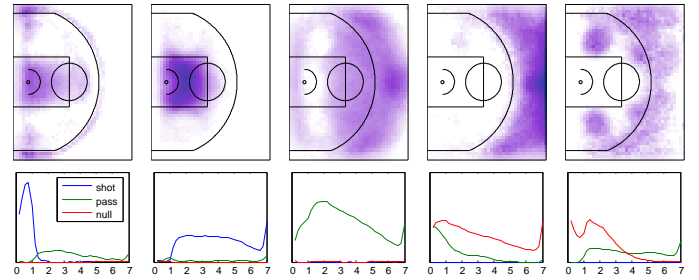

Fig. 11. *Top Row*: Depicting the spatial coefficients (five factors of V) of the tendency to shoot, pass, and keep the ball (null). *Bottom Row*: Depicting the time-varying coefficients for shooting T_s, passing T_p and keeping the ball $T_\perp$ over seven seconds.

E. What Defensive Positions Suppress Pass Probability?

We now inspect the factors $\tilde{C}_1$, $\tilde{M}_1$, $\tilde{C}_2$, $\tilde{M}_2$ that correspond to pass suppression. These factors are depicted in Figure 10. The left plots in the figure correspond to defending the direction towards the ball handler, and the right plots correspond to defending in the direction towards the basket.

In Figure 10(a), we see that the defenders tend to defend more closely when the player is closer to the basket in order to suppress a pass, which agrees with our intuition. In Figure 10(b), we see that defenders tend to hedge towards the middle of the court, while still defending the path to the basket, which also agrees with our intuition.[15] We also see in the first factor in Figure 4(b) that suppressing passes into the region near the basket is best accomplished by "fronting" the offensive player which leads to playing behind the player's path to the basket. Both sets of defender factors devote significant model capacity to the region near the basket, which also contributes to the error reduction in the bottom row of Figure 4(b).

F. How Does Behavior Vary with Duration of Possession?

We finally inspect the temporally varying factors T_s, T_p, $T_\perp$, and their corresponding spatial factors V. The top row of Figure 11 depicts the five dimensions of V that we estimated, and the bottom row depicts the corresponding temporally varying coefficients of the five factors of T_s, T_p, and $T_\perp$. The first factor corresponds to locations where the ball handler

---

[15]Intuitively, pass to the corner often comes from the middle of the court, so hedging against the middle should lead to suppressing passing probability.

tends to immediately shoot upon receiving a pass, as evidenced by the high T_s weight within the first second.[16] The second factor corresponds to locations where the ball handler tends to shoot after possessing the ball for at least one second, and is focused on areas near the basket. The third factor correspond to locations where the ball handler tends to pass, especially after possessing the ball for two seconds. The fourth factor corresponds to regions where the ball handler tends to pass immediately upon receiving the ball.[17] When the player does not immediately pass the ball, then the probability of passing decreases and the probability of the player holding on to the ball increases. The last factor correspond to regions where the ball handler tends to keep the ball for several seconds, such as in the low post where players typically either spin into the lane for a shot (the second factor) or pass the ball back to the perimeter (as evidenced by the rise of T_p in the last factor).

VII. DISCUSSION & CONCLUSIONS

We have presented a fine-grained spatial model for predicting dynamic in-game events in team sports, such as shooting and passing. Our approach learns semantically meaningful and compact representations across several views of spatial and temporal variations, which enables efficient predicting at test time (such as for interactive visualization). We validated our model using spatiotemporal tracking data from the 2012-2013 NBA season. We find that our model does indeed learn to make good predictions. We further conducted an extensive inspection the learned factors in our model, and find that they correspond to known intuitions of basketball gameplay.

While our model did show good prediction performance, it can still be substantially improved. For example, we found several instances where our learned factors exhibited "motion blur". Another interesting direction is to model multiple recent time steps (e.g., trajectories of players), as opposed to just the current game state. Other avenues for future work include predicting other types of in-game events such as driving the ball, cutting by teammates, pick and rolls [15]. Finally, we addressed only the one-step lookahead version of the more general problem of predicting complete sequences of sports gameplay, which is a much more difficult task.

REFERENCES

[1] S. Gelly, L. Kocsis, M. Schoenauer, M. Sebag, D. Silver, C. Szepesvári, and O. Teytaud, "The grand challenge of computer go: Monte carlo tree search and extensions," *Communications of the ACM*, vol. 55, no. 3, pp. 106–113, 2012.

[2] M. L. Littman, "Markov games as a framework for multi-agent reinforcement learning." in *International Conference on Machine Learning (ICML)*, 1994.

[3] Y. Shoham, R. Powers, and T. Grenager, "If multi-agent learning is the answer, what is the question?" *Artificial Intelligence*, vol. 171, no. 7, pp. 365–377, 2007.

[4] K. Waugh, B. D. Ziebart, and J. A. Bagnell, "Computational rationalization: The inverse equilibrium problem," in *International Conference on Machine Learning (ICML)*, 2013.

[5] Z. Lowe, "Lights, Cameras, Revolution," Grantland, 19 March, 2013.

[6] J. Lafferty, A. McCallum, and F. Pereira, "Conditional random fields: Probabilistic models for segmenting and labeling sequence data," in *International Conference on Machine Learning (ICML)*, 2001.

[7] D. Lee and H. S. Seung, "Algorithms for non-negative matrix factorization," in *Neural Information Processing Systems (NIPS)*, 2001.

[8] A. Miller, L. Bornn, R. Adams, and K. Goldsberry, "Factorized point process intensities: A spatial analysis of professional basketball," in *International Conference on Machine Learning (ICML)*, 2014.

[9] Prozone, www.prozonesports.com.

[10] STATS SportsVU, www.sportvu.com.

[11] Hawk-Eye, www.hawkeyeinnovations.co.uk.

[12] S. Pellegrini, A. Ess, K. Schindler, and L. van Gool, "You'll Never Walk Alone: Modeling Social Behavior for Multi-Target Tracking," in *CVPR*, 2009.

[13] M. Rodriguez, J. Sivic, I. Laptev, and J. Audibert, "Data-Driven Crowd Analysis in Video," in *ICCV*, 2011.

[14] K. Goldsberry and E. Weiss, "The Dwight effect: A new ensemble of interior defense analytics for the NBA," in *MIT Sloan Sports Analytics Conference (SSAC)*, 2012.

[15] A. McQueen, J. Wiens, and J. Guttag, "Automatically recognizing on-ball screens," in *MIT Sloan Sports Analytics Conference (SSAC)*, 2014.

[16] J. Fewell, D. Armbruster, J. Ingraham, A. Petersen, and J. Waters, "Basketball teams as strategic networks," *PLOS ONE*, vol. 7, no. 11, p. e47445, 2012.

[17] P. Maymin, "Acceleration in the NBA: Towards an algorithmic taxonomy of basketball plays," in *MIT Sloan Sports Analytics Conference (SSAC)*, 2013.

[18] A. Franks, A. Miller, L. Bornn, and K. Goldsberry, "Characterizing the spatial structure of defensive skill in professional basketball," http://arxiv.org/abs/1405.0231, 2014.

[19] A. Bialkowski, P. Lucey, P. Carr, Y. Yue, and I. Matthews, ""Win at home and draw away": Automatic formation analysis highlighting the differences in home and away team behaviors," in *MIT Sloan Sports Analytics Conference (SSAC)*, 2014.

[20] X. Wei, L. Sha, P. Lucey, S. Morgan, and S. Sridharan, "Large-Scale Analysis of Formations in Soccer," in *DICTA*, 2013.

[21] R. Adams, G. Dahl, and I. Murray, "Incorporating side information into probabilistic matrix factorization using gaussian processes," in *Conference on Uncertainty in Artificial Intelligence (UAI)*, 2010.

[22] M. Goldman and J. M. Rao, "Allocative and dynamic efficiency in NBA decision making," in *MIT Sloan Sports Analytics Conference (SSAC)*, 2011.

[23] D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol. 401, no. 6755, pp. 788–791, 1999.

[24] J.-P. Brunet, P. Tamayo, T. R. Golub, and J. P. Mesirov, "Metagenes and molecular pattern discovery using matrix factorization," *Proceedings of the National Academy of Sciences*, vol. 101, no. 12, pp. 4164–4169, 2004.

[25] J. Yang and J. Leskovec, "Overlapping community detection at scale: A nonnegative matrix factorization approach," in *ACM Conference on Web Search and Data Mining (WSDM)*, 2013.

[26] C. Rasmussen and C. Williams, *Gaussian processes for machine learning*. MIT press Cambridge, MA, 2006, vol. 1.

[27] B. Yu, J. Cunningham, G. Santhanam, S. Ryu, K. Shenoy, and M. Sahani, "Gaussian-process factor analysis for low-dimensional single-trial analysis of neural population activity." in *Neural Information Processing Systems (NIPS)*, 2008.

[28] D. Cervone, A. D'Amour, L. Bornn, and K. Goldsberry, "POINTWISE: Predicting points and valuing decisions in real time with nba optical tracking data," in *MIT Sloan Sports Analytics Conference (SSAC)*, 2014.

[29] Y. Koren and R. Bell, "Advances in collaborative filtering," in *Recommender Systems Handbook*. Springer, 2011, pp. 145–186.

[30] J. Duchi, E. Hazan, and Y. Singer, "Adaptive subgradient methods for online learning and stochastic optimization," *Journal of Machine Learning Research (JMLR)*, vol. 12, pp. 2121–2159, 2011.

[31] Y. Nesterov, "Gradient methods for minimizing composite objective function," 2007.

[16]When the ball handler does not shoot within the first second, then the probability of shooting decreases and the probability of passing increases.

[17]These regions correspond to locations where players rapidly pass, or "swing" the ball around three point line.