A Multiresolution Stochastic Process Model for Basketball Possession Outcomes

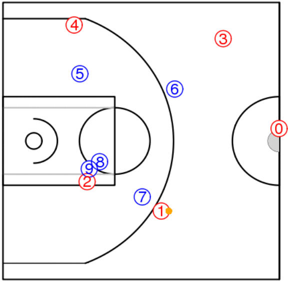
Dan Cervone, Alex D'Amour, Luke Bornn, Kirk Goldsberry

Harvard Statistics Department

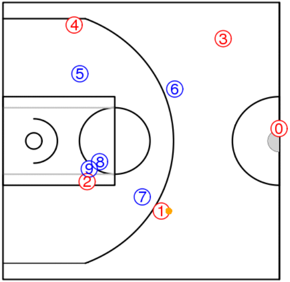
August 11, 2015

NBA optical tracking data

NBA optical tracking data

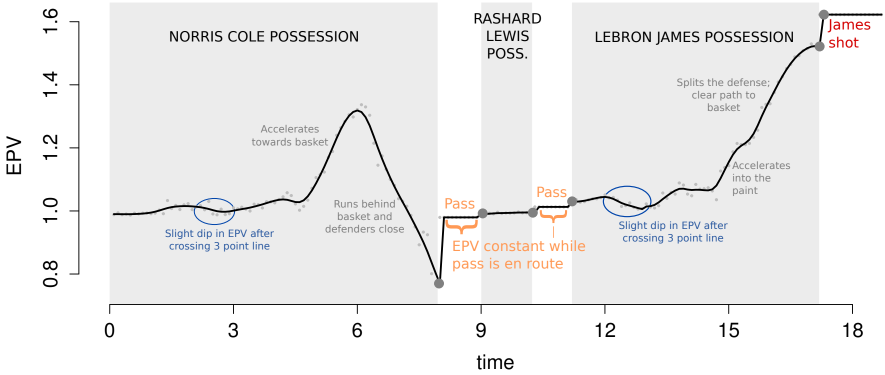


NBA optical tracking data



- (x, y) locations for all 10 players (5 on each team) at 25Hz.
- (x, y, z) locations for the ball at 25Hz.
- Event annotations (shots, passes, fouls, etc.).
- 1230 games from 2013-14 NBA, each 48 minutes, featuring 461 players in total.

Expected Possession Value (EPV)



EPV definition

Let Ω be the space of all possible basketball possessions. For $\omega \in \Omega$

EPV definition

Let Ω be the space of all possible basketball possessions. For $\omega \in \Omega$

- $X(\omega) \in \{0, 2, 3\}$: point value of possession ω.

EPV definition

Let Ω be the space of all possible basketball possessions. For $\omega \in \Omega$

- $X(\omega) \in \{0, 2, 3\}$: point value of possession ω.
- $T(\omega)$: possession length

EPV definition

Let Ω be the space of all possible basketball possessions. For $\omega \in \Omega$

- $X(\omega) \in \{0, 2, 3\}$: point value of possession ω.
- $T(\omega)$: possession length
- $Z_t(\omega), 0 \leq t \leq T(\omega)$: time series of optical tracking data.

EPV definition

Let Ω be the space of all possible basketball possessions. For $\omega \in \Omega$

- $X(\omega) \in \{0, 2, 3\}$: point value of possession ω.
- $T(\omega)$: possession length
- $Z_t(\omega), 0 \leq t \leq T(\omega)$: time series of optical tracking data.
- $\mathcal{F}_t^{(Z)} = \sigma(\{Z_s^{-1} : 0 \leq s \leq t\})$: natural filtration.

EPV definition

Let Ω be the space of all possible basketball possessions. For $\omega \in \Omega$

- $X(\omega) \in \{0, 2, 3\}$: point value of possession ω.
- $T(\omega)$: possession length
- $Z_t(\omega), 0 \leq t \leq T(\omega)$: time series of optical tracking data.
- $\mathcal{F}_t^{(Z)} = \sigma(\{Z_s^{-1} : 0 \leq s \leq t\})$: natural filtration.

Definition

The *expected possession value* (EPV) at time $t \geq 0$ during a possession is $\nu_t = \mathbb{E}[X|\mathcal{F}_t^{(Z)}]$.

EPV definition

Let Ω be the space of all possible basketball possessions. For $\omega \in \Omega$

- $X(\omega) \in \{0, 2, 3\}$: point value of possession ω.
- $T(\omega)$: possession length
- $Z_t(\omega), 0 \leq t \leq T(\omega)$: time series of optical tracking data.
- $\mathcal{F}_t^{(Z)} = \sigma(\{Z_s^{-1} : 0 \leq s \leq t\})$: natural filtration.

Definition

The *expected possession value* (EPV) at time $t \geq 0$ during a possession is $\nu_t = \mathbb{E}[X|\mathcal{F}_t^{(Z)}]$.

- EPV provides an instantaneous snapshot of the possession's value, given its full spatiotemporal history.
- ν_t is a Martingale: $\mathbb{E}[\nu_{t+h}|\mathcal{F}_t^{(Z)}] = \nu_t$ for all $h > 0$.

Calculating EPV

$\nu_t = \mathbb{E}[X|\mathcal{F}_t^{(Z)}]$

Regression-type prediction methods:

- Data are not traditional input/output pairs.
- No guarantee of stochastic consistency.

Calculating EPV

$\nu_t = \mathbb{E}[X|\mathcal{F}_t^{(Z)}]$

Regression-type prediction methods:

- Data are not traditional input/output pairs.
- No guarantee of stochastic consistency.

Markov chains:

+ Stochastically consistent.
- Information is lost through discretization.
- Many rare transitions.

Calculating EPV

$\nu_t = \mathbb{E}[X|\mathcal{F}_t^{(Z)}]$

Regression-type prediction methods:

- Data are not traditional input/output pairs.
- No guarantee of stochastic consistency.

Markov chains:

+ Stochastically consistent.
- Information is lost through discretization.
- Many rare transitions.

Brute force, "God model" for basketball.

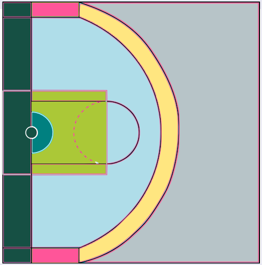
+ Allows Monte Carlo calculation of ν_t by simulating future possession paths.
- Z_t is high dimensional and includes discrete events (passes, shots, turnovers).

A coarsened process

Finite collection of states $\mathcal{C} = \mathcal{C}_{\text{poss}} \cup \mathcal{C}_{\text{end}} \cup \mathcal{C}_{\text{trans}}$.

$\mathcal{C}_{\text{poss}}$: Ball possession states

$\{\text{player}\} \times \{\text{region}\} \times \{\text{defender within 5 feet}\}$

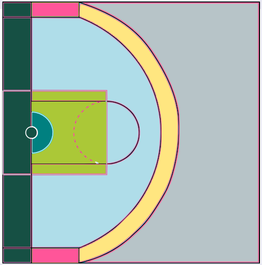


A coarsened process

Finite collection of states $\mathcal{C} = \mathcal{C}_{\text{poss}} \cup \mathcal{C}_{\text{end}} \cup \mathcal{C}_{\text{trans}}$.

$\mathcal{C}_{\text{poss}}$: Ball possession states

$\{\text{player}\} \times \{\text{region}\} \times \{\text{defender within 5 feet}\}$



$\mathcal{C}_{\text{end}}$: End states

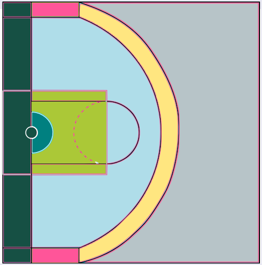
$\{\text{made 2, made 3, turnover}\}$

A coarsened process

Finite collection of states $\mathcal{C} = \mathcal{C}_{\text{poss}} \cup \mathcal{C}_{\text{end}} \cup \mathcal{C}_{\text{trans}}$.

$\mathcal{C}_{\text{poss}}$: Ball possession states

$\{\text{player}\} \times \{\text{region}\} \times \{\text{defender within 5 feet}\}$



$\mathcal{C}_{\text{end}}$: End states

$\{\text{made 2, made 3, turnover}\}$

$\mathcal{C}_{\text{trans}}$: Transition states

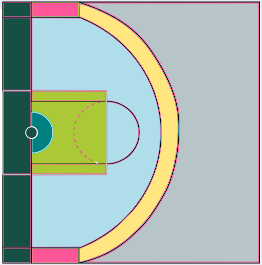
$\{\{\text{ pass linking } c, c' \in \mathcal{C}_{\text{poss}}\}$,
$\{\text{shot attempt from } c \in \mathcal{C}_{\text{poss}}\}$,
turnover in progress,
rebound in progress $\}$.

A coarsened process

Finite collection of states $\mathcal{C} = \mathcal{C}_{\text{poss}} \cup \mathcal{C}_{\text{end}} \cup \mathcal{C}_{\text{trans}}$.

$\mathcal{C}_{\text{poss}}$: Ball possession states

$\{\text{player}\} \times \{\text{region}\} \times \{\text{defender within 5 feet}\}$



$\mathcal{C}_{\text{end}}$: End states

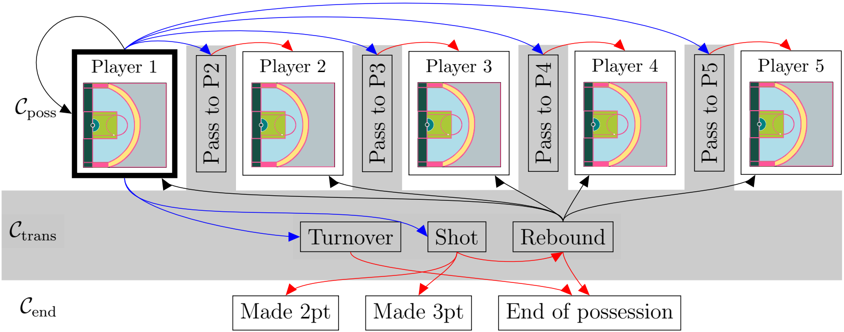
$\{\text{made 2, made 3, turnover}\}$

$\mathcal{C}_{\text{trans}}$: Transition states

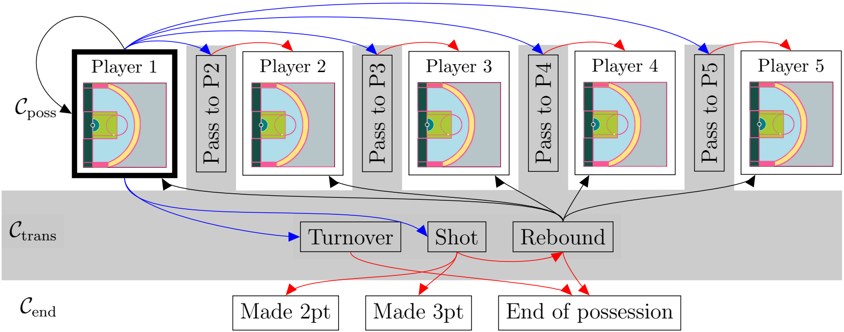
$\{\{$ pass linking $c, c' \in \mathcal{C}_{\text{poss}}\}$,
$\{$shot attempt from $c \in \mathcal{C}_{\text{poss}}\}$,
turnover in progress,
rebound in progress $\}$.

- $C_t \in \mathcal{C}$: state of the possession at time t.
- $C^{(0)}, C^{(1)}, \ldots, C^{(K)}$: discrete sequence of distinct states.

Possible paths for C_t



Possible paths for C_t



$$\tau_t = \begin{cases} \min\{s : s > t, C_s \in \mathcal{C}_{\text{trans}}\} & \text{if } C_t \in \mathcal{C}_{\text{poss}} \\ t & \text{if } C_t \notin \mathcal{C}_{\text{poss}} \end{cases}$$

$$\delta_t = \min\{s : s \geq \tau_t, C_s \notin \mathcal{C}_{\text{trans}}\}.$$

Stopping times for switching resolutions

$$\tau_t = \begin{cases} \min\{s : s > t, C_s \in \mathcal{C}_{\text{trans}}\} & \text{if } C_t \in \mathcal{C}_{\text{poss}} \\ t & \text{if } C_t \notin \mathcal{C}_{\text{poss}} \end{cases}$$

$$\delta_t = \min\{s : s \geq \tau_t, C_s \notin \mathcal{C}_{\text{trans}}\}.$$

Stopping times for switching resolutions

$$\tau_t = \begin{cases} \min\{s : s > t, C_s \in \mathcal{C}_{\text{trans}}\} & \text{if } C_t \in \mathcal{C}_{\text{poss}} \\ t & \text{if } C_t \notin \mathcal{C}_{\text{poss}} \end{cases}$$

$$\delta_t = \min\{s : s \geq \tau_t, C_s \notin \mathcal{C}_{\text{trans}}\}.$$

Key assumptions:

A1 C is marginally semi-Markov.

Stopping times for switching resolutions

$$\tau_t = \begin{cases} \min\{s : s > t, C_s \in \mathcal{C}_{\text{trans}}\} & \text{if } C_t \in \mathcal{C}_{\text{poss}} \\ t & \text{if } C_t \notin \mathcal{C}_{\text{poss}} \end{cases}$$

$$\delta_t = \min\{s : s \geq \tau_t, C_s \notin \mathcal{C}_{\text{trans}}\}.$$

Key assumptions:

A1 C is marginally semi-Markov.

A2 For all $s > \delta_t$, $\mathbb{P}(C_s | C_{\delta_t}, \mathcal{F}_t^{(Z)}) = \mathbb{P}(C_s | C_{\delta_t})$.

Stopping times for switching resolutions

$$\tau_t = \begin{cases} \min\{s : s > t, C_s \in \mathcal{C}_{\text{trans}}\} & \text{if } C_t \in \mathcal{C}_{\text{poss}} \\ t & \text{if } C_t \notin \mathcal{C}_{\text{poss}} \end{cases}$$

$$\delta_t = \min\{s : s \geq \tau_t, C_s \notin \mathcal{C}_{\text{trans}}\}.$$

Key assumptions:

- A1 C is marginally semi-Markov.
- A2 For all $s > \delta_t$, $\mathbb{P}(C_s | C_{\delta_t}, \mathcal{F}_t^{(Z)}) = \mathbb{P}(C_s | C_{\delta_t})$.

Theorem

Assume (A1)–(A2), then for all $0 \leq t < T$,

$$\nu_t = \sum_{c \in \{\mathcal{C}_{trans} \cup \mathcal{C}_{end}\}} \mathbb{E}[X | C_{\delta_t} = c] \mathbb{P}(C_{\delta_t} = c | \mathcal{F}_t^{(Z)}).$$

Multiresolution models

EPV:

$$\nu_t = \sum_{c \in \{\mathcal{C}_{\text{trans}} \cup \mathcal{C}_{\text{end}}\}} \mathbb{E}[X | C_{\delta_t} = c] \mathbb{P}(C_{\delta_t} = c | \mathcal{F}_t^{(Z)}).$$

Multiresolution models

EPV:

$$\nu_t = \sum_{c \in \{\mathcal{C}_{\text{trans}} \cup \mathcal{C}_{\text{end}}\}} \mathbb{E}[X | C_{\delta_t} = c] \mathbb{P}(C_{\delta_t} = c | \mathcal{F}_t^{(Z)}).$$

Let $M(t)$ be the event $\{\tau_t \leq t + \epsilon\}$.

M1 $\mathbb{P}(Z_{t+\epsilon} | M(t)^c, \mathcal{F}_t^{(Z)})$: the *microtransition model*.

M2 $\mathbb{P}(M(t) | \mathcal{F}_t^{(Z)})$: the *macrotransition entry model*.

M3 $\mathbb{P}(C_{\delta_t} | M(t), \mathcal{F}_t^{(Z)})$: the *macrotransition exit model*.

M4 $\mathbf{P}$, with $P_{qr} = \mathbb{P}(C^{(n+1)} = c_r | C^{(n)} = c_q)$: the *Markov transition probability matrix*.

Multiresolution models

EPV:

$$\nu_t = \sum_{c \in \{\mathcal{C}_{\text{trans}} \cup \mathcal{C}_{\text{end}}\}} \mathbb{E}[X | C_{\delta_t} = c] \mathbb{P}(C_{\delta_t} = c | \mathcal{F}_t^{(Z)}).$$

Let $M(t)$ be the event $\{\tau_t \leq t + \epsilon\}$.

M1 $\mathbb{P}(Z_{t+\epsilon} | M(t)^c, \mathcal{F}_t^{(Z)})$: the *microtransition model*.

M2 $\mathbb{P}(M(t) | \mathcal{F}_t^{(Z)})$: the *macrotransition entry model*.

M3 $\mathbb{P}(C_{\delta_t} | M(t), \mathcal{F}_t^{(Z)})$: the *macrotransition exit model*.

M4 $\mathbf{P}$, with $P_{qr} = \mathbb{P}(C^{(n+1)} = c_r | C^{(n)} = c_q)$: the *Markov transition probability matrix*.

Monte Carlo computation of ν_t:

- Draw $C_{\delta_t} | \mathcal{F}_t^{(Z)}$ using (M1)–(M3).
- Calculate $\mathbb{E}[X | C_{\delta_t}]$ using (M4).

Microtransition model

Player ℓ's position at time t is $\mathbf{z}^\ell(t) = (x^\ell(t), y^\ell(t))$.

$$x^\ell(t+\epsilon) = x^\ell(t) + \alpha_x^\ell[x^\ell(t) - x^\ell(t-\epsilon)] + \eta_x^\ell(t)$$

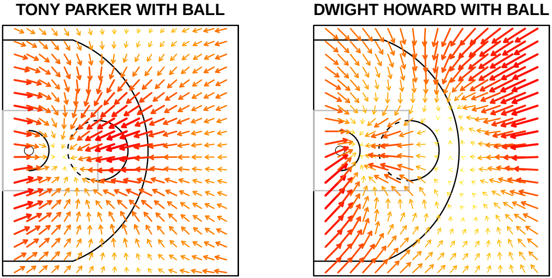
- $\eta_x^\ell(t) \sim \mathcal{N}(\mu_x^\ell(\mathbf{z}^\ell(t)), (\sigma_x^\ell)^2)$.
- μ_x has Gaussian Process prior.
- $y^\ell(t)$ modeled analogously (and independently).

Microtransition model

Player ℓ's position at time t is $\mathbf{z}^\ell(t) = (x^\ell(t), y^\ell(t))$.

$$x^\ell(t+\epsilon) = x^\ell(t) + \alpha_x^\ell[x^\ell(t) - x^\ell(t-\epsilon)] + \eta_x^\ell(t)$$

- $\eta_x^\ell(t) \sim \mathcal{N}(\mu_x^\ell(\mathbf{z}^\ell(t)), (\sigma_x^\ell)^2)$.
- μ_x has Gaussian Process prior.
- $y^\ell(t)$ modeled analogously (and independently).

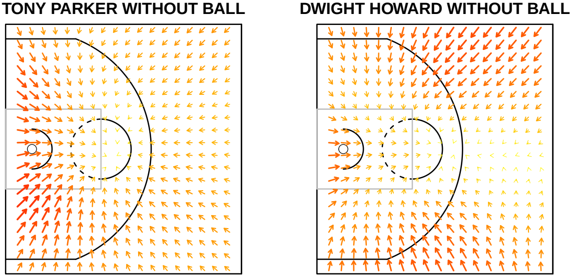


Microtransition model

Player ℓ's position at time t is $\mathbf{z}^\ell(t) = (x^\ell(t), y^\ell(t))$.

$$x^\ell(t+\epsilon) = x^\ell(t) + \alpha_x^\ell[x^\ell(t) - x^\ell(t-\epsilon)] + \eta_x^\ell(t)$$

- $\eta_x^\ell(t) \sim \mathcal{N}(\mu_x^\ell(\mathbf{z}^\ell(t)), (\sigma_x^\ell)^2)$.
- μ_x has Gaussian Process prior.
- $y^\ell(t)$ modeled analogously (and independently).



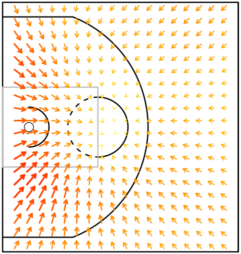
Microtransition model

Player ℓ's position at time t is $\mathbf{z}^\ell(t) = (x^\ell(t), y^\ell(t))$.

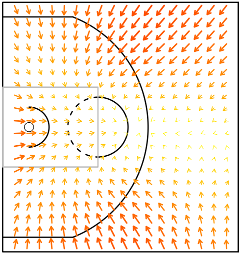
$$x^\ell(t+\epsilon) = x^\ell(t) + \alpha_x^\ell[x^\ell(t) - x^\ell(t-\epsilon)] + \eta_x^\ell(t)$$

- $\eta_x^\ell(t) \sim \mathcal{N}(\mu_x^\ell(\mathbf{z}^\ell(t)), (\sigma_x^\ell)^2)$.
- μ_x has Gaussian Process prior.
- $y^\ell(t)$ modeled analogously (and independently).

TONY PARKER WITHOUT BALL



DWIGHT HOWARD WITHOUT BALL



- Defensive microtransition model based on defensive matchups [Franks et al., 2015].

Macrotransition entry model

Recall $M(t) = \{\tau_t \leq t + \epsilon\}$:

- Six different "types", based on entry state $C_{\tau_t}, \cup_{j=1}^{6} M_j(t) = M(t)$.
- Hazards: $\lambda_j(t) = \lim_{\epsilon \to 0} \frac{\mathbb{P}(M_j(t) | \mathcal{F}_t^{(Z)})}{\epsilon}$.

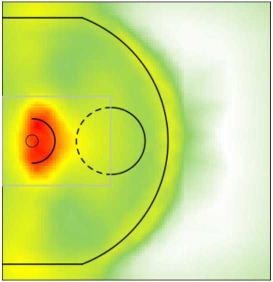
$$\log(\lambda_j(t)) = [\mathbf{W}_j^{\ell}(t)]' \boldsymbol{\beta}_j^{\ell} + \xi_j^{\ell}\left(\mathbf{z}^{\ell}(t)\right)$$

Macrotransition entry model

Recall $M(t) = \{\tau_t \leq t + \epsilon\}$:

- Six different "types", based on entry state $C_{\tau_t}, \cup_{j=1}^{6} M_j(t) = M(t)$.
- Hazards: $\lambda_j(t) = \lim_{\epsilon \to 0} \frac{\mathbb{P}(M_j(t)|\mathcal{F}_t^{(Z)})}{\epsilon}$.

$$\log(\lambda_j(t)) = [\mathbf{W}_j^{\ell}(t)]'\beta_j^{\ell} + \xi_j^{\ell}\left(\mathbf{z}^{\ell}(t)\right)$$

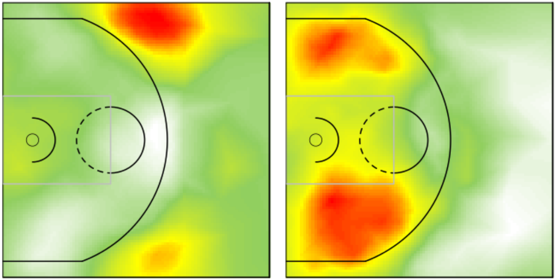


Macrotransition entry model

Recall $M(t) = \{\tau_t \leq t + \epsilon\}$:

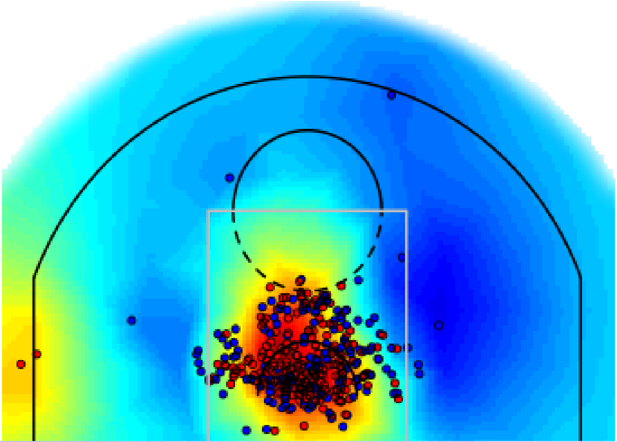
- Six different "types", based on entry state $C_{\tau_t}, \cup_{j=1}^{6} M_j(t) = M(t)$.
- Hazards: $\lambda_j(t) = \lim_{\epsilon \to 0} \frac{\mathbb{P}(M_j(t)|\mathcal{F}_t^{(Z)})}{\epsilon}$.

$$\log(\lambda_j(t)) = [\mathbf{W}_j^{\ell}(t)]'\beta_j^{\ell} + \xi_j^{\ell}\left(\mathbf{z}^{\ell}(t)\right) + \tilde{\xi}_j^{\ell}\left(\mathbf{z}_j(t)\right)\mathbf{1}[j \leq 4]$$



Hierarchical modeling

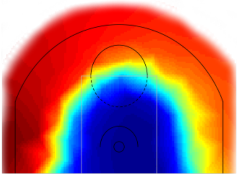
Dwight Howard's shot chart:

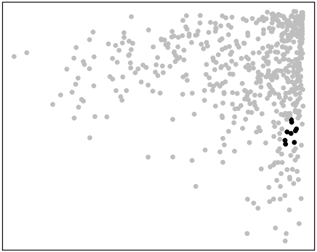


Hierarchical modeling

Shrinkage needed:

- Across space.
- Across different players.





Basis representation of spatial effects

Spatial effects ξ_j^ℓ

- ℓ: ballcarrier identity.
- j: macrotransition type (pass, shot, turnover).

Basis representation of spatial effects

Spatial effects ξ_j^ℓ

- ℓ: ballcarrier identity.
- j: macrotransition type (pass, shot, turnover).

Functional basis representation

$$\xi_j^\ell(\mathbf{z}) = [\mathbf{w}_j^\ell]' \boldsymbol{\phi}_j(\mathbf{z}).$$

- $\boldsymbol{\phi}_j = (\phi_{ji} \ \ \dots \phi_{jd})'$: d spatial basis functions.
- $\mathbf{w}_j^\ell$: weights/loadings.

Basis representation of spatial effects

Spatial effects ξ_j^ℓ

- ℓ: ballcarrier identity.
- j: macrotransition type (pass, shot, turnover).

Functional basis representation

$$\xi_j^\ell(\mathbf{z}) = [\mathbf{w}_j^\ell]'\phi_j(\mathbf{z}).$$

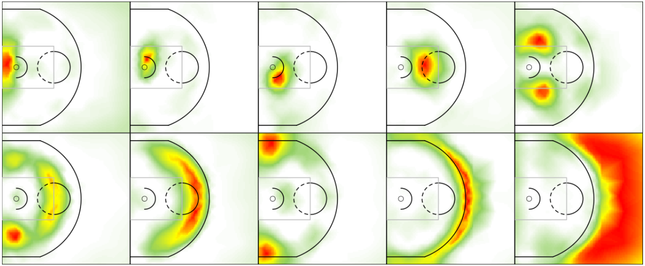
- $\boldsymbol{\phi}_j = (\phi_{ji} \ \ \dots \phi_{jd})'$: d spatial basis functions.
- $\mathbf{w}_j^\ell$: weights/loadings.

Information sharing

- $\boldsymbol{\phi}_j$ allows for non-stationarity, correlations between disjoint regions [Higdon, 2002].
- $\mathbf{w}_j^\ell$: weights across players follow a CAR model [Besag, 1974] based on player similarity graph $\mathbf{H}$.

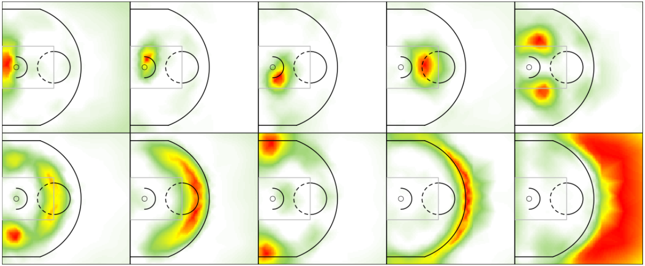
Basis representation of spatial effects

Basis functions ϕ_j learned in pre-processing step:

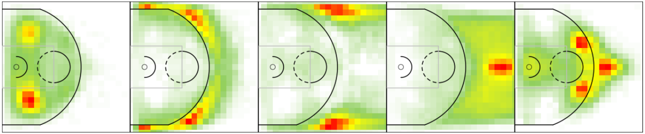


Basis representation of spatial effects

Basis functions ϕ_j learned in pre-processing step:



Graph $\mathbf{H}$ learned from players' court occupancy distribution:



Inference

"Partially Bayes" estimation of all model parameters:

Inference

"Partially Bayes" estimation of all model parameters:

- Multiresolution transition models provide partial likelihood factorization [Cox, 1975].
- All model parameters estimated using R-INLA software [Rue et al., 2009, Lindgren et al., 2011].

Inference

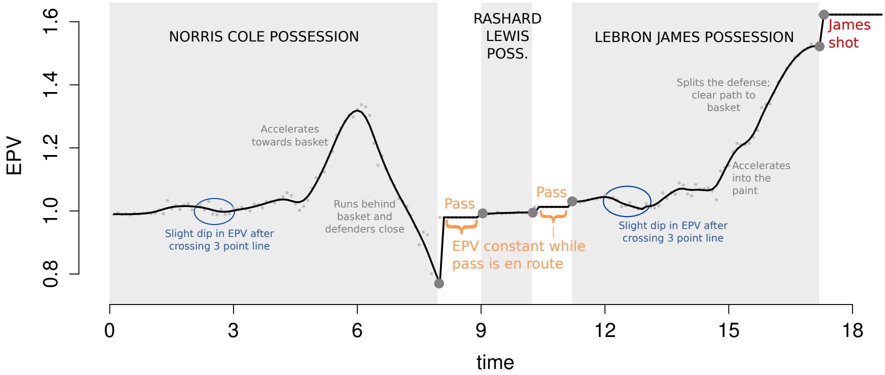
"Partially Bayes" estimation of all model parameters:

- Multiresolution transition models provide partial likelihood factorization [Cox, 1975].
- All model parameters estimated using R-INLA software [Rue et al., 2009, Lindgren et al., 2011].

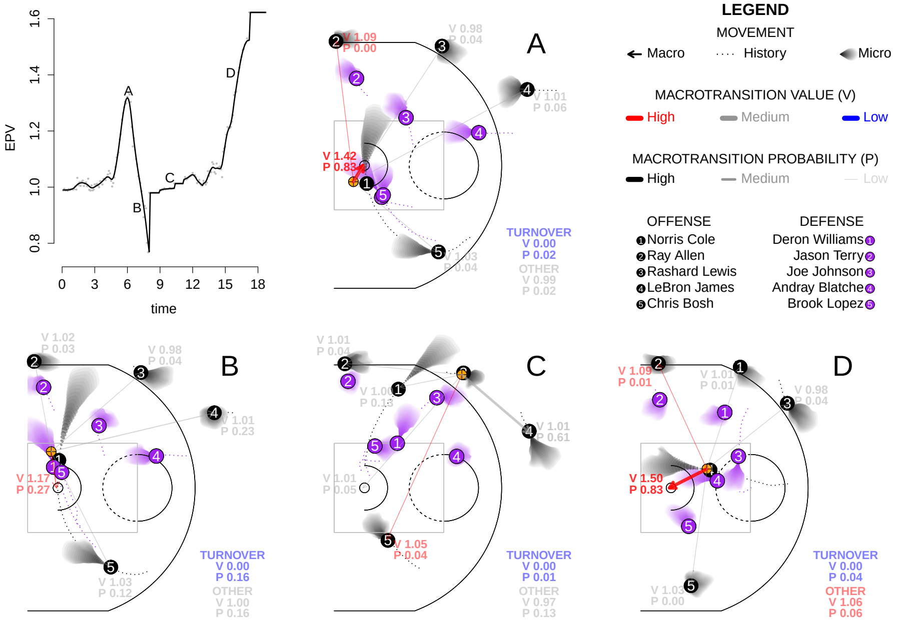
Distributed computing implementation:

- Preprocessing involves low-resource, highly parallelizable tasks.
- Parameter estimation involves several CPU- and memory-intensive tasks.
- Calculating EPV from parameter estimates involves low-resource, highly parallelizable tasks.

New insights from basketball possessions



New insights from basketball possessions



New metrics for player performance

EPV-added:

Rank	Player	EPVA
1	Kevin Durant	3.26
2	LeBron James	2.96
3	Jose Calderon	2.79
4	Dirk Nowitzki	2.69
5	Stephen Curry	2.50
6	Kyle Korver	2.01
7	Serge Ibaka	1.70
8	Channing Frye	1.65
9	Al Horford	1.55
10	Goran Dragic	1.54

Rank	Player	EPVA
277	Zaza Pachulia	-1.55
278	DeMarcus Cousins	-1.59
279	Gordon Hayward	-1.61
280	Jimmy Butler	-1.61
281	Rodney Stuckey	-1.63
282	Ersan Ilyasova	-1.89
283	DeMar DeRozan	-2.03
284	Rajon Rondo	-2.27
285	Ricky Rubio	-2.36
286	Rudy Gay	-2.59

Table : Top 10 and bottom 10 players by EPV-added (EPVA) per game in 2013-14, minimum 500 touches during season.

New metrics for player performance

Shot satisfaction:

Rank	Player	Satis.
1	Mason Plumlee	0.35
2	Pablo Prigioni	0.31
3	Mike Miller	0.27
4	Andre Drummond	0.26
5	Brandan Wright	0.24
6	DeAndre Jordan	0.24
7	Kyle Korver	0.24
8	Jose Calderon	0.22
9	Jodie Meeks	0.22
10	Anthony Tolliver	0.22

Rank	Player	Satis.
277	Garrett Temple	-0.02
278	Kevin Garnett	-0.02
279	Shane Larkin	-0.02
280	Tayshaun Prince	-0.03
281	Dennis Schroder	-0.04
282	LaMarcus Aldridge	-0.04
283	Ricky Rubio	-0.04
284	Roy Hibbert	-0.05
285	Will Bynum	-0.05
286	Darrell Arthur	-0.05

Table : Top 10 and bottom 10 players by shot satisfaction in 2013-14, minimum 500 touches during season.

Acknowledgements and future work

Our EPV framework can be extended to better incorporate unique basketball strategies:

- Additional macrotransitions can be defined, such as pick and rolls, screens, and other set plays.
- Use more information in defensive matchups (only defender locations, not identities, are currently used).
- Summarize and aggregate EPV estimates into useful player- or team-specific metrics.

Acknowledgements and future work

Our EPV framework can be extended to better incorporate unique basketball strategies:

- Additional macrotransitions can be defined, such as pick and rolls, screens, and other set plays.
- Use more information in defensive matchups (only defender locations, not identities, are currently used).
- Summarize and aggregate EPV estimates into useful player- or team-specific metrics.

Thanks to:

- Co-authors: Alex D'Amour, Luke Bornn, Kirk Goldsberry.
- Colleagues: Alex Franks, Andrew Miller.

References I

[0] Besag, J. (1974).
Spatial interaction and the statistical analysis of lattice systems.
Journal of the Royal Statistical Society: Series B (Methodological), 36(2):192–236.

[0] Cox, D. R. (1975).
A note on partially bayes inference and the linear model.
Biometrika, 62(3):651–654.

[0] Franks, A., Miller, A., Bornn, L., and Goldsberry, K. (2015).
Characterizing the spatial structure of defensive skill in professional basketball.
Annals of Applied Statistics.

[0] Higdon, D. (2002).
Space and space-time modeling using process convolutions.
In *Quantitative Methods for Current Environmental Issues*, pages 37–56. Springer, New York, NY.

[0] Lindgren, F., Rue, H., and Lindström, J. (2011).
An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach.
Journal of the Royal Statistical Society: Series B (Methodological), 73(4):423–498.

[0] Rue, H., Martino, S., and Chopin, N. (2009).
Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations.
Journal of the Royal Statistical Society: Series B (Methodological), 71(2):319–392.