From infinite random matrices over finite fields to square ice

Leonid Petrov
University of Virginia

March 13, 2019
• Infinite binary sequences

• Infinite triangular random matrices over a finite field

11011111110111010001010010110011110011110010001010...
Exchangeable random binary sequences

Exchangeability
A random sequence $X_1, X_2, \ldots$, where $X_i \in \{0, 1\}$, is called exchangeable if its distribution does not change under (finitary) permutations of indices.

Invariance under uniform resampling given the boundary conditions

Exchangeable distributions form a convex set

$$\mu = \alpha \mu_1 + (1 - \alpha) \mu_2, \quad \alpha \in [0, 1]$$

Extreme exchangeable distributions are the $\mu$’s which cannot be decomposed as above with $\mu_{1,2} \neq \mu$ and $\alpha \neq 0, 1$

Classification of extreme exchangeable distributions

Extreme exchangeable distributions are precisely the Bernoulli product measures $\mu_p$ indexed by $p \in [0, 1]$

Under $\mu_p$, the $X_i$’s are independent with $\mathbb{P}(X_i = 1) = p$. 

[de Finetti 1930s, Hewitt-Savage 1955]
Classification of extreme exchangeable distributions

Extreme exchangeable distributions are precisely the Bernoulli product measures $\mu_p$ indexed by $p \in [0, 1]$

Under $\mu_p$, the $X_i$'s are independent with $\mathbb{P}(X_i = 1) = p$.

How to sample

To sample an exchangeable sequence, first pick random $p$ from a mixing distribution $\nu$ on $[0, 1]$

Then, given $p$, sample independent $X_i$'s according to $\mu_p$

Example

The uniform mixing distribution $\nu$ on $[0, 1]$ corresponds to the “Polya urn”:

for each $n$, $X_1, \ldots, X_n$ has a uniformly random number $k \in \{0, \ldots, n\}$ of zeroes and ones (homework problem: how to pass from $n$ to $n+1$?)

Parameter recovery: Law of Large Numbers

$$\lim_{n \to +\infty} \frac{X_1 + \ldots + X_n}{n} = \nu$$

in distribution and a.s.
Ergodic approach for describing “boundaries”

1. Want to classify probability distributions with certain symmetry and sequential structure

2. Distributions form a convex set

3. Classify extreme distributions using the sequential structure (each infinite-level extreme is a limit of finite-level ones)

4. Each distribution is a convex combination of extremes

5. Law of Large Numbers for parameter recovery (the first focus of the talk)

6. Select non-extreme distributions are very interesting (won’t discuss this in the talk)

The ergodic approach was employed by Vershik and Kerov in 1970-80s to apply to representation theory of “big” groups:

- the infinite symmetric group \( S(\infty) \) [Edrei 1950s, Thoma 1964]

- the infinite-dimensional unitary group \( U(\infty) \) [Edrei 1950s, Voiculescu 1976], Vadim’s talk on Monday

Related applications include the study of ergodic central measures on matrices, both Hermitian/\( \mathbb{C} \), and over finite fields
• Infinite binary sequences

\[ (X, Y) \equiv (X + Y, Y) \mod 2 \]

• Infinite triangular random matrices over a finite field

Another symmetry of the \( \frac{1}{2} \) i.i.d. coin flip sequence!
Random triangular matrices over finite fields

\( \mathbb{F}_q \) — finite field (\( q = 2 \) when \( \mathbb{F}_2 = \{0, 1\} \) suffices)

\( \mathbb{U} \) — group of infinite uni upper triangular matrices over \( \mathbb{F}_q \)

Each \( n \times n \) triangular matrix is conjugate to a Jordan form by an element of \( GL_n(\mathbb{F}_q) \)

Jordan forms are encoded by Young diagrams \( \lambda = (\lambda_1 \geq \lambda_2 \geq \ldots \geq 0) \), \( \lambda_i \in \mathbb{Z} \)

\[
\begin{pmatrix}
1 & 1 \\
1 & 1 \\
\vdots & \ddots & \\
0 & \ldots & 0
\end{pmatrix}
\sim
\begin{pmatrix}
\lambda_1 & 1 \\
1 & 1 \\
\vdots & \ddots & \ddots \\
0 & \ldots & \lambda_2
\end{pmatrix}
\]

\( GL_\infty(\mathbb{F}_q) \) — group of infinite matrices which finitely differ from the identity

Exchangeability analogue (symmetry of measures)

A probability Borel measure \( \mu \) on \( \mathbb{U} \) is called **central** if \( \mu(M) = \mu(gMg^{-1}) \) for all measurable \( M \subset \mathbb{U} \) and \( g \in GL_\infty(\mathbb{F}_q) \) such that \( gMg^{-1} \subset \mathbb{U} \)
Exchangeability analogue

A probability Borel measure $\mu$ on $\mathbb{U}$ is called central if $\mu(M) = \mu(gMg^{-1})$ for all measurable $M \subset \mathbb{U}$ and $g \in GL_\infty(\mathbb{F}_q)$ such that $gMg^{-1} \subset \mathbb{U}$.

Example: uniform product measure on $\mathbb{U}$ for which $X_{ij}$, $i < j$, are independent $\in \mathbb{F}_q$

$g \in GL_\infty(\mathbb{F}_q)$

$$g = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ \vdots \end{pmatrix} \quad \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$g^{-1} \quad \overset{d}{=} \quad \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$

Exercise: for $X, Y \ iid\ from \ \mathbb{F}_2$, we have $(X, Y) \overset{d}{=} (X + Y, Y)$
Central measures form a convex set. **Goal:** classify extreme central measures

- Related to representation theory of $GL_n(\mathbb{F}_q)$ as $n \to \infty$  
  [Vershik-Kerov 90s+]
  [Gorin-Kerov-Vershik 2012]

- At a level of (some) tools, is a one-parameter deformation of the representation theory of $S(\infty)$ (the latter corresponds to $q \to \infty$)

- The answer was conjectured by Kerov in 1992 and proven by Matveev in 2017 (together with a Macdonald generalization which adds yet one more parameter)

**Theorem**

Extreme central measures are in one to one correspondence with tuples

\[ \alpha_1 \geq \alpha_2 \geq \ldots \geq 0, \quad \beta_1 \geq \beta_2 \geq \ldots \geq 0, \quad \gamma \geq 0 \]

such that

\[ \frac{\gamma}{1 - t} + \sum_{i \geq 1} \left( \alpha_i + \frac{\beta_i}{1 - t} \right) = 1 \]

\[ t \coloneqq \frac{1}{q} \]

($t=0$ - infinite symmetric group)
Realization of extreme central measures

"coin flips"

Central measures are determined by a sequence of random Jordan block structures $\lambda(n)$ of $n \times n$ corners, with

$$|\lambda(n)| = \lambda_1(n) + \lambda_2(n) + \ldots + \lambda_n(n), \ n = 1, 2, \ldots$$
Let $\omega = (\alpha; \beta; \gamma)$ where $\alpha = (\alpha_1 \geq \alpha_2 \geq \ldots \geq 0)$, 
$\beta = (\beta_1 \geq \beta_2 \geq \ldots \geq 0)$, $\gamma \geq 0$, and 
$\frac{\gamma}{1-t} + \sum_{i \geq 0} \left( \alpha_i + \frac{\beta_i}{1-t} \right) = 1$

$\omega_0$ be $\alpha_i = \beta_i = 0$, $\gamma = 1 - t$

Central measures are determined by a sequence of random Jordan block structures $\lambda(n)$ of $n \times n$ corners, with $|\lambda(n)| = \lambda_1(n) + \lambda_2(n) + \ldots + \lambda_n(n)$, $n = 1, 2, \ldots$

**Realization of extreme central measures**

$$\text{Prob}(\lambda(n) = \nu) = \frac{1}{Z} P_{\nu}(\omega_0)Q_{\nu}(\omega)$$

$P_{\nu}, Q_{\nu} — \text{ Hall-Littlewood symmetric polynomials}$

Uniform measure is extreme and corresponds to $\alpha_i = (1 - q^{-1})q^{1-i}$, $i = 1, 2, \ldots$; $\beta_j = \gamma = 0$
Example of a Hall-Littlewood polynomial

\[ P_{(4,0)}(x_1, x_2) = x_1^4 + x_2^4 + (1 - t)(x_1^3 x_2 + x_1 x_2^3) + (1 - t)x_1^2 x_2^2 \]

\[ Q_\nu = b_\nu(t) P_\nu \]

\[ \text{Prob} (\lambda(n) = \nu) = \frac{1}{Z} P_\nu(\omega_0)Q_\nu(\omega) \]

Couple of useful facts about Hall-Littlewood polynomials

\[ \sum_\nu P_\nu(x_1, \ldots, x_N)Q_\nu(y_1, \ldots, y_M) = \prod_{i=1}^{N} \prod_{j=1}^{M} \frac{1 - tx_iy_j}{1 - x_iy_j} \]

\[ t = 0 \quad \text{— Schur polynomials} \quad s_\lambda(x_1, \ldots, x_N) = \frac{\det[x_i^{\lambda_j+N-j}]_{i,j=1}^{N}}{\prod_{1\leq i<j\leq N}(x_i - x_j)} \]
<table>
<thead>
<tr>
<th>De Finetti</th>
<th>Matrices over finite fields</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random infinite binary sequences</td>
<td>Random infinite uni-uppertriangular matrices</td>
</tr>
<tr>
<td>Exchangeability (invariance under permutations)</td>
<td>Centrality (invariance under conjugations)</td>
</tr>
<tr>
<td>Extremes are parametrized by 1-d space $p \in [0, 1]$</td>
<td>Extremes are parametrized by $\infty$-d space $\alpha = (\alpha_1 \geq \alpha_2 \geq \ldots \geq 0)$, $\beta = (\beta_1 \geq \beta_2 \geq \ldots \geq 0)$, $\gamma \geq 0$ such that $\frac{\gamma}{1-t} + \sum_{i \geq 0} \left( \alpha_i + \frac{\beta_i}{1-t} \right) = 1$</td>
</tr>
<tr>
<td>Realization of extremes: iid Bernoulli (coin tossing)</td>
<td>Realization of extremes through Hall-Littlewood polynomials (example: uniform Bernoulli product measure on uni-uppertriangular matrices)</td>
</tr>
</tbody>
</table>
Law of Large Numbers

Theorem [Bufetov-P. 2014]

Take an extreme measure $\mu$ corresponding to $\omega = (\alpha; \beta; \gamma)$. Let $\lambda(n)$ be the Jordan block structure of the $n \times n$ corner.

Then as $n \to +\infty$, 
\[
\frac{\lambda_i(n)}{n} \to \alpha_i, \quad \frac{\lambda'_i(n)}{n} \to \frac{\beta_i}{1-t}
\]

rows columns

Theorem describes asymptotic sizes of large Jordan blocks & asymptotic frequencies of small Jordan blocks for matrices from extreme measures

Earlier results

- Uniform upper triangular matrices [Borodin 1995], answering a question of A.A.Kirillov
- $t = 0$, asymptotic character theory of $S(\infty)$ [Vershik-Kerov 1980s]
Law of Large Numbers: idea of proof

1. Construct a (randomized) algorithm for exact sampling of $\lambda(n)$ coming from the extreme measure $\mu_\omega$

2. Analyze the algorithm probabilistically to get limiting frequencies of rows and columns

The Young diagrams $\lambda(n)$ are sampled by constructing random Young tableaux

$$T(k + 1) = T(k) \leftarrow a$$

Insertions are randomized; for $t = 0$ reduce to the classical Robinson-Schensted-Knuth ones (with Vershik–Kerov modifications)

New letters appear independently using $\alpha_j, \beta_j, \gamma$ parameters

$\alpha\beta\gamma$-tableaux (generalize semistandard Young tableaux)

[Vershik-Kerov 1986]
Take random words with independent letters (the sum of probabilities is 1):

\[ P(k) = \alpha_k, \quad P(\hat{k}) = \frac{\beta_k}{1 - t}, \quad \frac{\gamma}{1 - t} \quad \text{— continuous part} \]

After \( n \) steps, the Hall-Littlewood RSK sampling produces a random Young diagram \( \lambda(n) \), the shape of the random tableau

**Theorem** [Borodin-P. 2013], [Bufetov-P. 2014]

The distribution of \( \lambda(n) \) coincides with the Jordan block structure of the \( n \times n \) corner of the random matrix coming from the extreme measure \( \mu_\omega \).
Probabilistic consequences

Typical tableaux:
Mostly letters $k$ in the $k$-th row, mostly letters $\hat{k}$ in the $k$-th column, plus some “Plancherel dust” and other lower order errors.

This leads to the Law of Large Numbers for the Jordan blocks structure of extreme central measures.

Remark
Under additional restrictions (all parameters are distinct and $\gamma = 0$), one should also get a Central Limit Theorem with Gaussian fluctuations of order $\sqrt{n}$.

The fluctuations of each row and column are almost independent, modulo that the number of boxes is fixed.
Towards the six vertex model
For about 20 years now, **Robinson-Schensted-Knuth** type combinatorial algorithms are providing exact observables of **1-dimensional interacting particle systems** and related models through Schur functions.

Less than a decade ago, a new wave has started involving deformations of Schur functions. The Robinson-Schensted-Knuth type constructions are also deformed *(randomized)*.

**Higher spin vertex models, dynamical vertex models, ...**

- **Spin q-Whittaker** \((q,s)\)
- **Spin Hall-Littlewood** \((t,s)\)
- **Macdonald** \((q,t)\)
- **q-Whittaker** \((q)\)
- **Jack** \((\beta)\)
- **Hall-Littlewood** \((t)\)
- **Random polymers, geometric RSK**

**TASEPs, longest increasing subsequences**

- **q-TASEPs**
  - **q-Whittaker** \((q)\)
  - **Jack** \((\beta)\)
  - **Schur**

**ASEP, six vertex model**

- **TASEPs**, longest increasing subsequences

**References**

- **[Baik-Deift-Johansson, Johansson, ...]**
- **[A.N. Kirillov, Noumi-Yamada, O'Connell, Corwin-O'Connell-Seppalainen-Zygouras]**
- **[Bufetov-P. 2017, Mucciconi-Bufetov-P. in prep]**
- **[Bufetov-P. 2014]**
- **[Bufetov-Matveev, Borodin-Bufetov-Wheeler]**
Six vertex (square ice) model

- The system is a **Markov process** (= stochastic interacting particle system)
- Partition functions are products, not determinants
- This and many other stochastic vertex models are exactly solvable to the point of asymptotics

**Stochastic six vertex model**

\[ \sum w(\rightarrow_i) = 1 \]

- \( a_1 = a_2 = 1 \)
- \( b_i + c_i = 1 \)

\[ a_1 = a_2 = b_2 = c_1 = c_2 = 1, \quad b_1 = 3, \]

domain wall boundary conditions

\[ a_1 = a_2 = 1, \quad b_i + c_i = 1 \]  
(easier to analyze)

stochastic six vertex model in a quadrant
Marginally Markovian 1d projection in the Hall-Littlewood Robinson-Schensted insertion

Keep only parameters $\alpha_1, \alpha_2, \ldots$, and let $\beta_j = \gamma = 0$
The Young tableau is semistandard

Defects in the first column of the tableau = particles on $\mathbb{Z}$

- RSK-insert $k \Leftrightarrow$ clock rings at site $k$ (rate $\alpha_k$)
- if there is a particle at $k$, it wakes up and jumps up by one
- if the destination is occupied, the next particle is pushed by one and wakes up, the pusher stops
- the active particle moves through the empty space with probability $t$ per step; stops with probability $1 - t$

Called the half-continuous stochastic six vertex model

Asymptotics (homogeneous $\alpha$) - [Ghosal 2017]
Get a continuous-time Markov chain on particle configurations. Plot trajectories of particles:
This is the same as...

$\beta_j = \gamma = 0$

$\frac{b_2}{1} \rightarrow 1$, Poisson type limit in the horizontal direction

$b_1 = t$ fixed
Half-continuous stochastic six vertex model

Theorem which follows from the constructions:

\[ N - \lambda'_1(n), \] where \( \lambda \) comes from \( \omega = ((\alpha_1, \ldots, \alpha_N); 0; 0) \), is the height function of the half-continuous stochastic six vertex model at \((n, N)\), where \( n \) is the number of (independent) jumps occurred.

The distribution of \( \lambda(n) \) is expressed through the Hall-Littlewood symmetric polynomials, which allows to write down explicit distributional formulas for \( \lambda'_1(n) \), and obtain asymptotics.

Theorem

The height function at \((x, y)\) is distributed as \(y - \lambda_1\), where \(\lambda\) has the Hall-Littlewood distribution \(\propto P_\lambda(a_1, \ldots, a_x)Q_\lambda(b_1, \ldots, b_y)\)

Theorem (t-moment formula) \(\forall \ell = 1, 2, \ldots\)

\[
\mathbb{E} t^{\ell} h(x, y) = t^{\ell/2} \int \cdots \int \prod_{1 \leq i < j \leq \ell} \frac{w_i - w_j}{w_i - tw_j} \prod_{i=1}^{\ell} \left( \frac{dw_i}{2\pi i w_i} \prod_{r=1}^{x} \frac{a_r - w_i}{a_r - tw_i} \prod_{r=1}^{y} \frac{tw_i - b_r}{w_i - b_r} \right)
\]

Contours are around \(\{b_i\}\) and \(0\) (in a certain order)

Has many proofs...

- A la [Tracy-Widom 2007+] for ASEP based on coordinate Bethe Ansatz
- Yang-Baxter equation and Cauchy identities via \(q\)-correlations
- Randomized Robinson-Schensted-Knuth plus Macdonald difference operators for Hall-Littlewood polynomials
- Randomization of the Yang-Baxter equation + HL polynomials
Asymptotics in the homogeneous stochastic six vertex model

Height function has a limit shape \( \frac{1}{L} h(Lx, Ly) \rightarrow \mathcal{H}(x, y) \)

The nontrivial part of the limit shape is \( \mathcal{H}(x, y) = \frac{\left( \sqrt{x(1-b_1)} - \sqrt{y(1-b_2)} \right)^2}{b_2 - b_1} \)

Fluctuations are governed by the GUE Tracy–Widom distribution (originated about 25 years ago in random matrix theory)

\[
\lim_{L \to +\infty} \mathbb{P} \left( \frac{h(Lx, Ly) - L\mathcal{H}(x, y)}{\sigma_{x,y} L^{1/3}} \geq -s \right) = F_{\text{GUE}}(s)
\]

- Higher spin versions + spin Hall-Littlewood polynomials
- Multilayer systems
- Degenerates to ASEP
- Limits to Kardar-Parisi-Zhang equation
- There is also a stochastic telegraph equation
- ...
Conclusion: It is worthwhile to connect particle systems to random partitions associated with symmetric functions...

Higher spin vertex models, dynamical vertex models, ...
Thank you!