Universality of the Stochastic Bessel Operator

Brian Rider
Patrick Waters $\leftarrow$ me
5/9/16

Laguerre-type β-ensembles

Distribution on $0 < \lambda_1 < \lambda_2 < \ldots < \lambda_n$:

$$dP_n(\lambda) = C|\mathrm{Vand}(\lambda)|^\beta \prod_{k=1}^{n} \lambda_k^{\frac{\beta}{2}(a+1)-1} e^{-n\beta V(\lambda_k)} d\lambda_k$$

- $a > -1$ and $\beta \geq 1$.
- V is a polynomial such that $V(\lambda^2)$ is uniformly convex.
- In the literature typically $V(\lambda) = \lambda/2$ "Pure Laguerre case"

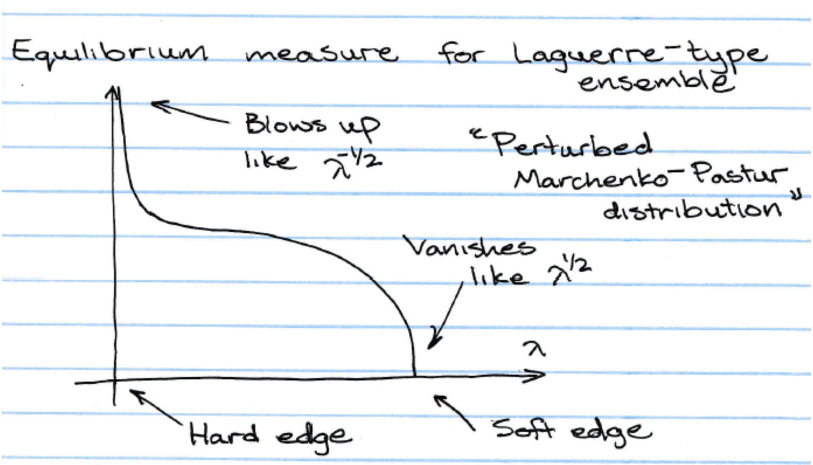
We prove for any fixed $k = 1, 2, 3, \ldots$ that $\lambda_1, \ldots, \lambda_k$ converge in distribution to the squares of the smallest k singular values of

$$\sqrt{x}\frac{d}{dx} + \frac{a+1}{2\sqrt{x}} + \beta^{-1/2} b'_x$$

with Dirichlet boundary condition at $x = 0$.

"Universal" – the limit does not depend on V.

Equilibrium measure for Laguerre-type ensembles



Starting point of our project:

- Propose a random bi-diagonal matrix with e-val's of BB^t distributed $dP_n(\lambda)$
- In the cases $\beta = 1, 2, 4$ the bi-diagonal model should be similar to the classical Laguerre ensembles by Householder transformations

Bidiagonal matrix model

Bidiagonal and tridiagonal matrices:

$$B_n = \begin{pmatrix} x_1 & & & \\ -y_1 & x_2 & & \\ & \ddots & \ddots & \\ & & -y_{n-1} & x_n \end{pmatrix}.$$

Bidiagonal matrix model

Bidiagonal and tridiagonal matrices:

$$B_n = \begin{pmatrix} x_1 & & & \\ -y_1 & x_2 & & \\ & \ddots & \ddots & \\ & & -y_{n-1} & x_n \end{pmatrix}.$$

Spectral map:

$$e_n^t (B_n B_n^t - zI)^{-1} e_n = \int_{\mathbb{R}} \frac{d\mu(w)}{w - z}, \qquad \mu = \sum_{k=1}^{n} q_k^2 \delta_{\lambda_k}$$

$$(x, y) \mapsto (\lambda, q)$$

Bidiagonal matrix model (cont'd)

- If B_n has this distribution:

$$
\begin{aligned}
dP_n(x,y) =& C \ \exp(-n\beta H(x,y)) \prod_k dx_k \, dy_k \\
H(x,y) =& \mathrm{tr}\, V(BB^t) - \sum_k \frac{k+a-\beta^{-1}}{n} \log x_k \\
& - \sum_k \frac{k-\beta^{-1}}{n} \log y_k
\end{aligned}
$$

then the eigenvalues of $B_n B_n^t$ have distribution $dP_n(\lambda)$.

Why this model?

Relation to the classical models
$\beta = 1, 2, 4$

Case $\beta = 1, 2, 4$

- Let L be a random $n \times (n+a)$ matrix with entries in $\mathbb{R}, \mathbb{C}$ or $\mathbb{H}$ and distribution

$$dP_n(L) = Ce^{-n\beta \operatorname{tr} V(LL^\dagger)}\, dL,$$

then $dP_n(\lambda)$ is the distribution of squares of the singular values of L.

Case $\beta = 1, 2, 4$

- Let M be a random positive definite $n \times n$ symmetric/Hermitian/self-dual matrix (entries in $\mathbb{R}/\mathbb{C}/\mathbb{H}$) and distribution

$$dP_n(M) = Ce^{-n\beta \operatorname{tr} V(M)} \det(M)^{\frac{\beta}{2}(a+1)-1}\, dM,$$

then $dP_n(\lambda)$ is the distribution of the eigenvalues of M.

Random matrix $\rightarrow$ Random diff'l operator

Matrix $\to$ finite rank operator on $L^2[0,1]$

Embed $\mathbb{R}^n$ into $L^2(\mathbb{R})$:

$$e_k(x) = n^{1/2}\mathbb{1}_{[\frac{k-1}{n},\frac{k}{n}]}(x)$$

Matrix $\rightarrow$ finite rank operator on $L^2[0,1]$

Embed $\mathbb{R}^n$ into $L^2(\mathbb{R})$:

$$e_k(x) = n^{1/2}\mathbb{1}_{[\frac{k-1}{n},\frac{k}{n}]}(x)$$

Induces mapping from matrices to linear operators:

$$M \cdot f(x) = \begin{pmatrix} e_1(x) \\ \vdots \\ e_n(x) \end{pmatrix}^t M \begin{pmatrix} < f, e_1(x) > \\ \vdots \\ < f, e_n(x) > \end{pmatrix}$$

Edelman and Sutton 2006: "From random matrices to stochastic operators"

Main idea: bidiagonal and tridiagonal models for β-ensembles look like random differential operators.

Example: pure Laguerre case $V(\lambda) = \lambda/2$

$$nB_n \text{ ``} \to \text{''} \sqrt{x}\frac{d}{dx} + \frac{a+1}{2\sqrt{x}} + \beta^{-1/2} b'_x$$

(Stochasic Bessel Operator)

Why random differential operators?

Differential operator:

$$n \begin{pmatrix} 1 & & & \\ -1 & 1 & & \\ & \ddots & \ddots & \\ & & -1 & 1 \end{pmatrix} \begin{pmatrix} f(1/n) \\ f(2/n) \\ \vdots \\ f(1) \end{pmatrix} \approx \begin{pmatrix} f'(1/n) \\ f'(2/n) \\ \vdots \\ f'(1) \end{pmatrix}$$

$$n \begin{pmatrix} 1 & & & \\ -1 & 1 & & \\ & \ddots & \ddots & \\ & & -1 & 1 \end{pmatrix} \text{“} \rightarrow \text{”} \ \frac{d}{dx} \ \text{ with b.c.'s } f(0) = 0$$

Why random differential operators? (cont'd)

White noise:

$$n^{-1/2}\begin{pmatrix} N[0,1] & & \\ & \ddots & \\ & & N[0,1] \end{pmatrix}\begin{pmatrix} f(1/n) \\ \vdots \\ f(1) \end{pmatrix} \approx \begin{pmatrix} (b_{1/n} - b_0)f(1/n) \\ \vdots \\ (b_{(n-1)/n} - b_1)f(1) \end{pmatrix}$$

$$n^{1/2}\begin{pmatrix} N[0,1] & & \\ & \ddots & \\ & & N[0,1] \end{pmatrix} \text{ ``}\rightarrow\text{'' } b'_x$$

Literature on Stochastic Bessel/Airy operators

Ramirez and Rider 2008: "Diffusion at the random matrix hard edge"

Laguerre case $V(\lambda) = \lambda/2$.

Rigorous proof that

$$\begin{matrix} k \text{ smallest} \\ \text{singular val's of } nB_n \end{matrix} \Longrightarrow \begin{matrix} k \text{ smallest singular val's of} \\ \sqrt{x}\frac{d}{dx} + \frac{a+1}{2\sqrt{x}} + \beta^{-1/2} b'_x \end{matrix}$$

(Stochasic Bessel Operator)

Krishnapur, Rider and Virag 2013: "Universality of the Stochastic Airy Operator"

Hermite-type β-ensemble:

$$dP_n(\lambda) = C|\mathrm{Vand}(\lambda)|^\beta \prod_{k=1}^{n} e^{-n\beta V(\lambda_k)}\, d\lambda_k,$$

V is a convex polynomial.

Krishnapur, Rider and Virag 2013: “Universality of the Stochastic Airy Operator”

Hermite-type β-ensemble:

$$dP_n(\lambda) = C|\mathrm{Vand}(\lambda)|^\beta \prod_{k=1}^{n} e^{-n\beta V(\lambda_k)}\, d\lambda_k,$$

V is a convex polynomial.

Rigorous proof that

k largest e-val's of scaled tridiagonal model $\Longrightarrow$ k largest e-val's of $-\frac{d^2}{dx^2} + x + \frac{2}{\sqrt{\beta}} b'_x$

(Stochasic Airy Operator)

Hard vs. soft edge

Soft edge

- SAO acts on functions with domain $[0, \infty)$
- $e_k(x) = n^{1/6} \mathbb{1}_{[(j-1)n^{-1/3}, jn^{-1/3}]}(x)$
- $n^{-1/3}$ proportion of rows determine behavior on $[0, n^{1/3}]$

Hard vs. soft edge

Soft edge

- SAO acts on functions with domain $[0, \infty)$
- $e_k(x) = n^{1/6} \mathbb{1}_{[(j-1)n^{-1/3}, jn^{-1/3}]}(x)$
- $n^{-1/3}$ proportion of rows determine behavior on $[0, n^{1/3}]$

Hard edge

- SBO acts on functions with domain $[0, 1]$
- $e_k(x) = n^{1/2} \mathbb{1}_{[(j-1)n^{-1}, jn^{-1}]}(x)$
- Any fraction of the rows determine behavior on same fraction of domain.

Statement of our theorem

Universality of the Stochastic Bessel Operator

Polynomial potential: $V(\lambda) = \sum g_m \lambda^m$, with $V(\lambda^2)$ unif. convex

Universality of the Stochastic Bessel Operator

Polynomial potential: $V(\lambda) = \sum g_m \lambda^m$, with $V(\lambda^2)$ unif. convex
Define two auxiliary functions:

$$s = \sum g_m m \binom{2m}{m} \phi(s)^{2m} \qquad (\phi = \text{unique pos. real sol'n})$$

$$\theta(s) = \left(\frac{1}{c} \int_0^s \frac{du}{\phi(u)} \right)^2 \qquad (c \text{ s.t. } \theta(1) = 1).$$

Universality of the Stochastic Bessel Operator (cont'd)

With the following embedding of $\mathbb{R}^n$ into $L^2[0,1]$:

$$e_k(s) = \left(\theta\left(\frac{k}{n}\right) - \theta\left(\frac{k-1}{n}\right)\right)^{-1/2} \mathbb{1}_{[\theta(\frac{k}{n}),\theta(\frac{k-1}{n})]}(s),$$

then

$$(nB_n)^{-1} \to \overset{\text{Integral operator with kernel}}{\mathbb{1}_{t<s}\frac{1}{\sqrt{s}}\left(\frac{t}{s}\right)^{a/2}\exp\int_t^s \frac{db_w}{\sqrt{\beta w}}}$$

- Convergence in distribution w.r.t. Hilbert-Schmidt norm + extra domination condition.

Comments on universality of SBO statement

- This implies smallest k singular values of nB_n converge in distribution to those of SBO.

Outline of our proof

Formula for the integral kernel (finite n)

Forget about θ, and use embedding with mesh size $1/n$.
Integral kernel for B_n^{-1}:

$$(B_n^{-1} f)(s) = \int_0^1 K_n(s,t) f(t)\, dt$$

$$K_n(s,t) = \mathbb{1}_{\lfloor tn \rfloor \le \lfloor sn \rfloor} \frac{1}{X_{\lfloor sn \rfloor}} \exp \sum_{k=\lfloor tn \rfloor}^{\lfloor sn \rfloor - 1} \log \frac{Y_k}{X_k}$$

Goal: Compute limit of $nK_n(s,t)$ as a functional of Br. Mo.

Random integral kernel:

$$K_n(s,t) = \mathbb{1}_{\lfloor tn \rfloor \le \lfloor sn \rfloor} \frac{1}{X_{\lfloor sn \rfloor}} \exp \sum_{k=\lfloor tn \rfloor}^{\lfloor sn \rfloor - 1} \log \frac{Y_k}{X_k}$$

Components of our proof

1. Second order asymptotics as $n \to \infty$ for mode x^o, y^o of distribution on $\{X_k, Y_k : \; k = 1, \ldots, n\}$
2. CLT for $\displaystyle\sum_{k=\lfloor tn \rfloor}^{n} \log \frac{Y_k/y_k^o}{X_k/x_k^o}$
3. Additional tightness estimate
4. With the above 3 things, an argument of Rider and Ramirez '08 finishes the proof.

Second order asymptotics for x^o, y^o

LLN for X_{sn}, Y_{sn} (First order asymptotics)

Measure on X_k, Y_k:

$$dP_n(x,y) = C\exp(-n\beta H(x,y))\,dx\,dy$$
$$H(x,y) = \mathrm{tr}\ V(B_n B_n^t) - \sum\left(\frac{i+a-\beta^{-1}}{n}\log x_k + \frac{i-\beta^{-1}}{n}\log y_k\right)$$

LLN for X_{sn}, Y_{sn} (First order asymptotics)

Measure on X_k, Y_k:

$$dP_n(x,y) = C \exp(-n\beta H(x,y))\, dx\, dy$$
$$H(x,y) = \mathrm{tr}\ V(B_n B_n^t) - \sum \left(\frac{i+a-\beta^{-1}}{n} \log x_k + \frac{i-\beta^{-1}}{n} \log y_k \right)$$

Continuum limit approximation, find "coarse minimizer" of H:

$$X_{sn}, Y_{sn} \to \phi(s), \qquad s = \sum g_m m \binom{2m}{m} \phi(s)^{2m}.$$

For context, in the Laguerre case $\phi(s) = \sqrt{s}$.

How to get second order asymptotics of the mode?

Candidate for fine approximation of mode:

$$\begin{aligned} x_k^o =& \phi(k/n) + n^{-1} x^{(1)}(k/n) + O(n^{-2}) \\ =& x^\dagger(k/n) \qquad\qquad \text{(and same for } y_k^o\text{)} \end{aligned}$$

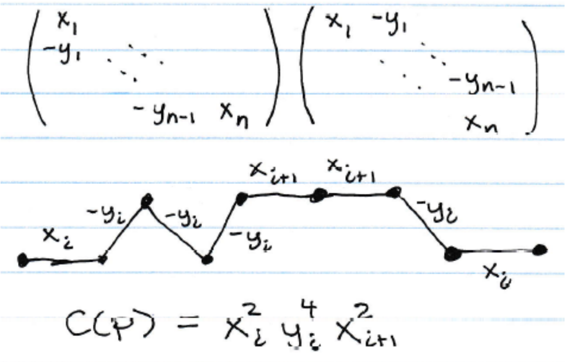
Since H is uniformly convex:

$$\left\| (x^o, y^o) - (x^\dagger, y^\dagger) \right\|_2 \leq c_u^{-1} \left\| \nabla H(x^\dagger, y^\dagger) \right\|_2$$

Method: compute n^0 and n^{-1} terms of ∇H at the candidate minimizer; define $x^{(1)}(s), y^{(1)}(s)$ so that these terms vanish.

Lattice paths

$$(B_nB_n^t)^m_{ii} = \sum_{\text{paths of length } 2m} \text{contribution of path}$$



Main ingredient for 2nd order asymptotics of x^o, y^o

$$\frac{\partial H}{\partial x_i}(x^\dagger, y^\dagger) = \sum_{m=1}^{d} g_m \left[A_m \phi^{2m-1} + \frac{1}{n}\phi^{2m-2}\left(B_m x^{(1)} + C_m y^{(1)} + D_m \phi'\right)\right] - \frac{i + a - \beta^{-1}}{n\phi} + \frac{i x^{(1)}}{n^2\phi^2} + O\left(n^{-2}\right).$$

The functions $\phi, \phi', x^{(1)}$ and $y^{(1)}$ are all evaluated at i/n, and

$$A_m = m\binom{2m}{m}, \qquad B_m = \frac{2m^2 - 2m + 1}{2m - 1} m \binom{2m}{m},$$

$$C_m = \frac{2m^2 - 2m}{2m - 1} m \binom{2m}{m}, \qquad D_m = -\frac{m^2 - m}{2m - 1} m \binom{2m}{m}.$$

Result of drift calculation

$$\lim_{n\to\infty}\sum_{i=nt}^{ns}\log\frac{x_i^o}{y_i^o} = \int_t^s \frac{x^{(1)}(\tau)-y^{(1)}(\tau)}{\phi(\tau)}\,d\tau$$
$$= \dots$$
$$= \left(\frac{a}{2}+\frac{1}{4}\right)\log\frac{\theta(t)}{\theta(s)} - \frac{1}{2}\log\frac{\phi(t)}{\phi(s)}.$$

Recall that $\phi(k/n)$ is the LLN behavior of x_k, y_k and

$$\theta(s) = \left(\frac{1}{c}\int_0^s \frac{du}{\phi(u)}\right)^2$$

CLT for noise term in integral kernel

Analysis of the integral kernels

$$\sum_{k=tn}^{sn} \log \frac{X_k}{Y_k} \approx \overbrace{-\left(a+\frac{1}{2}\right)\log\frac{\theta(s)}{\theta(t)} + \frac{1}{2}\log\frac{\phi(s)}{\phi(t)}}^{\text{From } 2^{\text{nd}} \text{ order asympt of } x^o, y^o} + \sum_{k=tn}^{sn} \underbrace{\frac{X_k - x_k^o - Y_k + y_k^o}{\phi(k/n)}}_{\text{Approximately centered RV}}$$

(x_k^o, y_k^o is the minimizer of H)

$$K_n(s,t) = \mathbb{1}_{\lfloor tn \rfloor \leq \lfloor sn \rfloor} \frac{1}{X_{\lfloor sn \rfloor}} \exp \sum_{k=\lfloor tn \rfloor}^{\lfloor sn \rfloor - 1} \log \frac{Y_k}{X_k}$$

For all $\delta \in (0,1]$ we prove the following convergence in distribution with respect to the Skorokhod topology on $D[\delta, 1]$:

$$\sum_{k=tn}^{n} \log \frac{X_k/x_k^o}{Y_k/y_k^o} \implies \frac{1}{\sqrt{\beta}} \int_{\theta(t)}^{1} \frac{db_\tau}{\sqrt{\tau}}.$$

CLT for fluctuation term

$$\mathrm{Var}(X_k - Y_k) \approx \frac{2\phi(k/n)}{n\beta \int_0^{k/n} \frac{du}{\phi(u)}}$$

$$\sum_{k=tn}^{sn} \frac{X_k - x_k^o - Y_k + y_k^o}{\phi(k/n)} \Longrightarrow \left(\int_t^s \frac{2d\tau}{\beta\phi(\tau) \int_0^\tau \frac{du}{\phi(u)}} \right)^{1/2} N[0,1]$$

$$= \frac{1}{\beta} \int_{\theta(t)}^{\theta(s)} \frac{db_\tau}{\sqrt{\tau}}$$

Properties of the measure

$$
\begin{aligned}
dP_n(x, y) &= C \exp(-n\beta H(x, y))\, dx\, dy \\
H(x, y) &= \mathrm{tr}\ V(B_n B_n^t) - \sum \left(\frac{i + a - \beta^{-1}}{n} \log x_k + \frac{i - \beta^{-1}}{n} \log y_k \right)
\end{aligned}
$$

- Concentration of measure inequality
- Decay of covariances
- Approximate translation symmetry of measure

Local interactions

Notation:

- $Z_{2k-1} = X_k$, $Z_{2k} = Y_k$.
- $I = \{i_0, \ldots, i_1\}$ (consecutive set of index variables)
- $\partial I = \{i_0 - d, \ldots, i_0 - 1\} \cup \{i_1 + 1, \ldots, i_1 + d\}$
- $Z_j = q_j \qquad j \in \partial I$ (boundary values on ∂I)

Local interactions

Notation:

- $Z_{2k-1} = X_k$, $Z_{2k} = Y_k$.
- $I = \{i_0, \dots, i_1\}$ (consecutive set of index variables)
- $\partial I = \{i_0 - d, \dots, i_0 - 1\} \cup \{i_1 + 1, \dots, i_1 + d\}$
- $Z_j = q_j \qquad j \in \partial I$ (boundary values on ∂I)

Then the conditional measure $dP(\cdot|q)$ is independent of $\{Z_k : \; k \notin I \cup \partial I\}$

Decay of dependence of conditional minimizers on boundary values

Notation:

- $Z_{2k-1} = X_k$, $Z_{2k} = Y_k$.
- $I = \{i_0, \ldots, i_1\}$ (consecutive set of index variables)
- $\partial I = \{i_0 - d, \ldots, i_0 - 1\} \cup \{i_1 + 1, \ldots, i_1 + d\}$
- $Z_j = q_j \qquad j \in \partial I$ (boundary values on ∂I)
- $z^q =$ mode of conditional measure on $\{Z_k : \ k \in I\}$ given q

Decay of dependence of conditional minimizers on boundary values

Notation:

- $Z_{2k-1} = X_k$, $Z_{2k} = Y_k$.
- $I = \{i_0, \ldots, i_1\}$ (consecutive set of index variables)
- $\partial I = \{i_0 - d, \ldots, i_0 - 1\} \cup \{i_1 + 1, \ldots, i_1 + d\}$
- $Z_j = q_j \qquad j \in \partial I$ (boundary values on ∂I)
- z^q = mode of conditional measure on $\{Z_k : \; k \in I\}$ given q

$$|z_i^q - z_i^o| \leq C \max_{j \in \partial I} \left(|z_j^q - z_j^o| e^{-r|i-j|} \right)$$

Decay of covariances

Notation:

- $I = \{i_0, \ldots, i_1\}$, $J = \{j_0, \ldots, j_1\}$
- $F \in \sigma\{Z_k : \ k \in I\}$, $G \in \sigma\{Z_k : \ k \in J\}$

We believe the following to be true:

$$\mathrm{Cov}(F, G) \leq Ce^{-r \,\mathrm{distance}(I,J)}$$

Decay of covariances

Notation:

- $I = \{i_0, \ldots, i_1\}$, $J = \{j_0, \ldots, j_1\}$
- $F \in \sigma\{Z_k : k \in I\}$, $G \in \sigma\{Z_k : k \in J\}$

We believe the following to be true:

$$\mathrm{Cov}(F, G) \leq Ce^{-r\,\mathrm{distance}(I,J)}$$

We are able to prove the following:

Under reasonable assumptions on F, G and if $\mathrm{distance}(I, J) > C' \log n$, then

$$\mathrm{Cov}(F, G) \leq Cn^{-5/4} \max\left(|\mathbb{E}F|, |\mathbb{E}G|\right).$$

Concentration of measure

The following is surprisingly non-trivial to prove:

$$P\left\{\|X, Y - x^o, y^o\|_\infty > r\right\} < Ce^{-cnr^2}$$

The ingredients one has to work with are:

- Gaussian domination for a subset of the variables:

$$P\left\{\|X, Y - x^o, y^o\|_2 > r\right\} < P\{\|G\|_2 > r\}$$
$$G = \text{Vector of indep. } N[0, (n\beta c_u)^{-1/2}] \text{ RV's}$$

- Borel type inequality of Ledoux

$$P\{F > \mathbb{E}[F] + r\} \leq \exp\left(\frac{n\beta c_u r^2}{2\kappa^2}\right)$$

Summary of CLT proof

$$\sum_{k=tn}^{n} \log \frac{X_k/x_k^o}{Y_k/y_k^o} \Longrightarrow \frac{1}{\sqrt{\beta}} \int_{\theta(t)}^{1} \frac{db_\tau}{\sqrt{\tau}}.$$

- Use a Bernstein blocking argument
- Approximate characteristic functions for one increment with steepest descent calculation
- Use decay of covariances to prove finite distributions have approx. indep. increments
- Upgrade to FCLT using moment condition from Billingsley *Convergence of Probability Measures*

Additional domination/tightness property

The RV's κ_n defined by

$$\kappa_n = \sup_{C_0 \log(n)/n < t < 1} \frac{\displaystyle\sum_{k=nt}^{n} \log\left(\frac{X_k/x_k^o}{Y_k/y_k^o}\right)}{(-\log t)^{3/4}},$$

are tight.