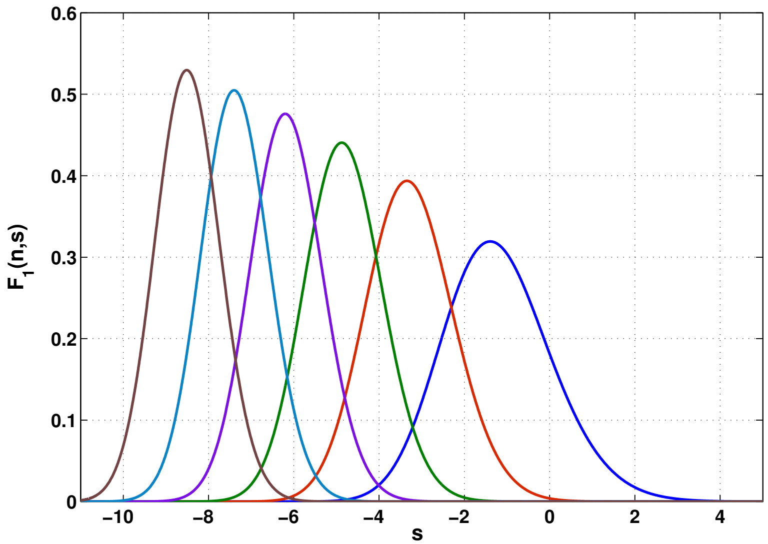
Accurate Numerical Evaluation of Distribution Functions

for Orthogonal and Symplectic Matrix Ensembles



pdf of n-th largest level in edge scaled GOE

Folkmar Bornemann

Two tools used in integrable systems

Ivar Fredholm (1866–1927)

determinant of integral operator (1899)

$$Ku(x) = \int_a^b K(x,y)u(y)\,dy$$

$$\rightsquigarrow$$

$$\det(I + zK) = \sum_{n=0}^{\infty} \frac{z^n}{n!} \int_{[a,b]^n} \det_{i,j=1}^{n} K(t_i, t_j)\,dt$$

Paul Painlevé (1863–1933)

six families of irreducible transcendental functions (1895)

$$u_{xx} = 6u^2 + x$$

$$u_{xx} = 2u^3 + xu - \alpha$$

$$u_{xx} = u^{-1}u_x^2 - x^{-1}u_x + x^{-1}(\alpha u^2 + \beta) + \gamma u^3 + \delta u^{-1}$$

$$u_{xx} = \cdots$$

$$u_{xx} = \cdots$$

$$u_{xx} = \cdots$$

n-th largest level in edge scaled GUE

$$\mathbb{P}(\text{exactly } n \text{ levels in } (s,\infty)) = \frac{(-1)^n}{n!} \frac{\partial^n}{\partial z^n} F_2(s;z)\Big|_{z=1}$$

(Forrester '93)

$$F_2(s;z) = \det\left(I - z\,K_{\mathrm{Ai}}\restriction_{L^2(s,\infty)}\right)$$

with kernel

$$K_{\mathrm{Ai}}(x,y) = \frac{\mathrm{Ai}(x)\,\mathrm{Ai}'(y) - \mathrm{Ai}'(x)\,\mathrm{Ai}(y)}{x-y}$$

(Tracy/Widom '93)

$$F_2(s;z) = \exp\left(-\int_s^\infty (x-s)u(x;z)^2\,dx\right)$$

with Painlevé II

$$u_{xx} = 2u^3 + xu$$

$$u(x;z) \simeq \sqrt{z}\,\mathrm{Ai}(x) \qquad (x\to\infty)$$

Without the Painlevé representations, the numerical evaluation of the Fredholm determinants is quite involved.

— Tracy/Widom '00

m-point quadrature formula

$$\int_a^b dt\, f(t) \approx \sum_{k=1}^m w_k\, f(x_k)$$

The Idea (Hilbert 1904, B. '10)

$$\det(I+zK) \overset{\text{Fredholm}}{\underset{1903}{=}} \sum_{n=0}^\infty \frac{z^n}{n!} \int_a^b dt_1 \cdots \int_a^b dt_n \det_{i,j=1}^n K(t_i,t_j)$$

$$\approx \sum_{n=0}^\infty \frac{z^n}{n!} \sum_{k_1=1}^m w_{k_1} \cdots \sum_{k_n=1}^m w_{k_n} \det_{i,j=1}^n K(x_{k_i},x_{k_j})$$

$$\overset{\text{v. Koch}}{\underset{1892}{=}} \det(I+zK_m)$$

with the $m\times m$-matrix

$$K_m = \left(w_i^{1/2} K(x_i,x_j) w_j^{1/2}\right)_{i,j=1}^m$$

Matlab-Code

```
[w,x] = QuadratureRule(a,b,m);
w = sqrt(w); [xi,xj] = ndgrid(x,x);
d = det(eye(m)+z*(w'*w).*K(xi,xj));
```

Theorem (B. '10)

For quadrature formula of order ν with positive weights:

- if kernel is $C^{k-1,1}([a,b]^2)$,

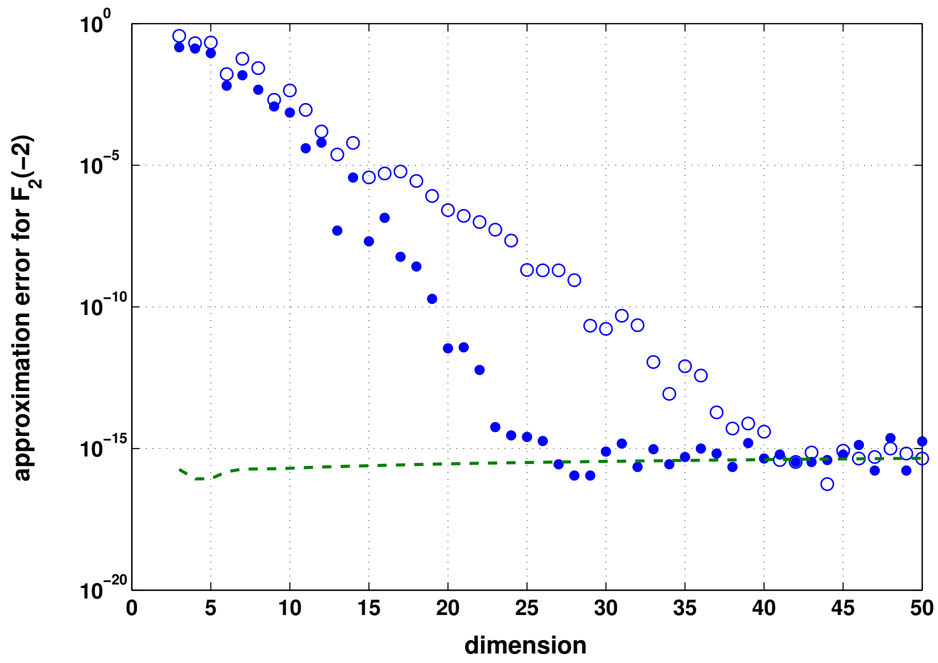
$$\text{error} = O(\nu^{-k}) \qquad (\nu \to \infty);$$

- if kernel is bounded analytic, there is $\rho > 1$ with

$$\text{error} = O(\rho^{-\nu}) \qquad (\nu \to \infty).$$

Tracy–Widom distribution

$$F_2(s) = \det\left(I - K_{\mathrm{Ai}}\restriction_{L^2(s,\infty)}\right), \qquad K_{\mathrm{Ai}}(x,y) = \frac{\mathrm{Ai}(x)\,\mathrm{Ai}'(y) - \mathrm{Ai}'(x)\,\mathrm{Ai}(y)}{x-y}$$



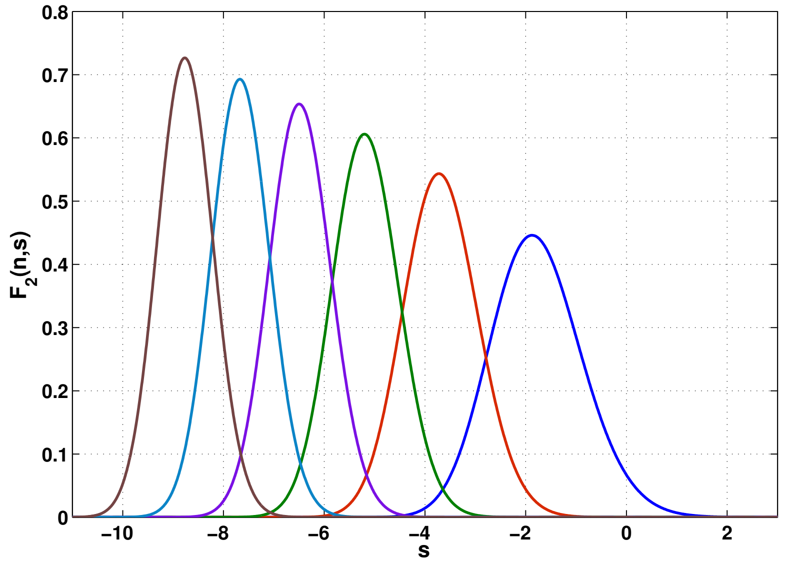
dots: Gauss–Legendre, circles: Clenshaw–Curtis

Perturbation bound for m-dimensional determinants: (B. '10)

$$\text{round-off error} \leqslant \sqrt{m}\,\|K_m\|_F \cdot u_{\text{machine}}$$

n-th largest level in edge scaled GUE

$$\mathbb{P}(\text{exactly } n \text{ levels in } (s,\infty)) = \frac{(-1)^n}{n!}\frac{\partial^n}{\partial z^n}\det\left(I - zK_{\text{Ai}}\restriction_{L^2(s,\infty)}\right)\Big|_{z=1}$$



probability density of n-th largest level in edge scaled GUE

Numerical method

$f(z) = \det(I + zK_{\text{Ai}})$ is *entire* of order 0

$$\boxed{\frac{f^{(n)}(z)}{n!} = \frac{1}{2\pi r^n}\int_0^{2\pi} e^{-in\theta} f(z + re^{i\theta})\, d\theta}$$

- trapezoidal rule *exponentially* convergent
- numerical stability: judicious choice of $r > 0$

  Example: (B. '11)

  f entire of order $\rho > 0$ and type $\tau > 0$

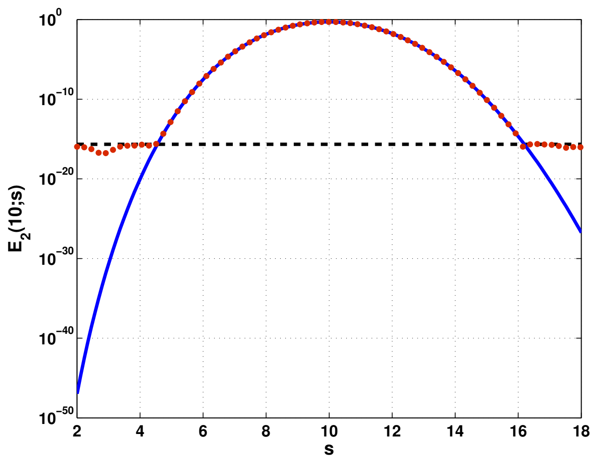
$$\boxed{r_{\text{opt}} = (n/\rho\tau)^{1/\rho}}$$

Generating functions of probabilities

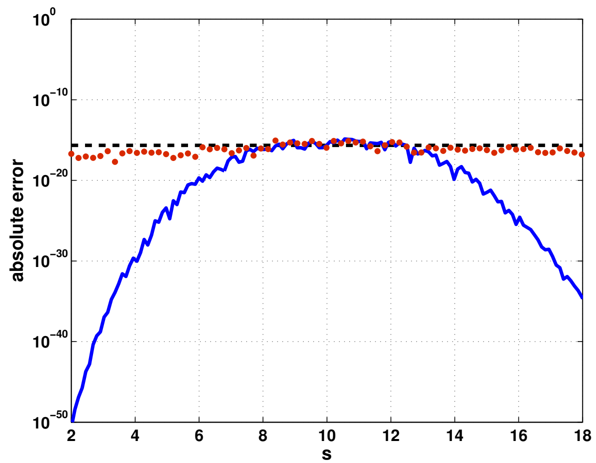
- absolute error: $r = 1$ reasonable (Lyness/Sande '71)
- relative error ($=$ accurate tails):

$$r_{\text{opt}} = \operatorname*{argmin}_{r>0} r^{-n} f(r)$$

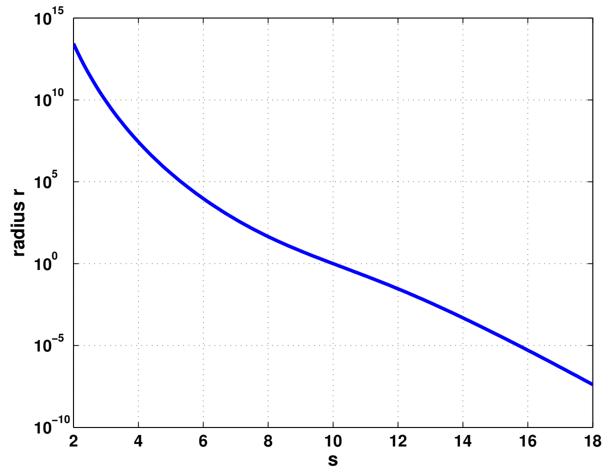
unique solution of *convex* optimization problem (B. '11)



gap probability $E_2(10; s)$ of GUE



absolute error



r_{opt} as a function of s

Combinatorial structure of determinantal processes

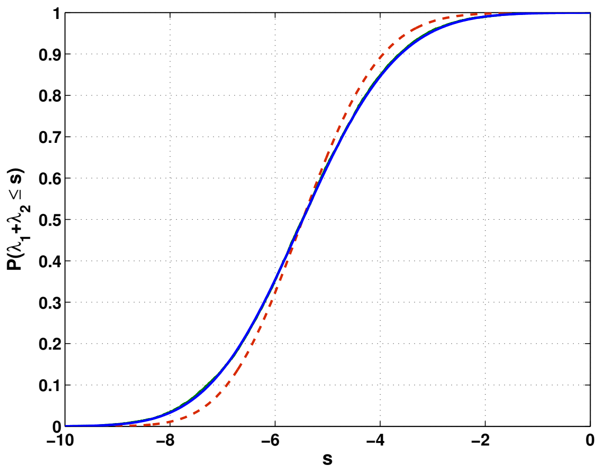
kernel K, disjoint intervals $J_1, \ldots, J_N$, multi-index $\alpha \in \mathbb{N}_0^N$

$$\mathbb{P}(\text{exactly } \alpha_j \text{ levels lie in } J_j,\ j = 1, \ldots, N)$$

$$= \frac{(-1)^{|\alpha|}}{\alpha!} \frac{\partial^\alpha}{\partial z^\alpha} \det\left(I - \begin{pmatrix} z_1 K & \cdots & z_N K \\ \vdots & & \vdots \\ z_1 K & \cdots & z_N K \end{pmatrix} \restriction_{L^2(J_1)\oplus\cdots\oplus L^2(J_N)} \right)\Bigg|_{z_1=\cdots=z_N=1}$$

Examples

- joint pdfs
- pdfs of sums, products, etc.



cdf of the sum of largest two levels in GUE

systems of integral operators = integral operator on coproduct

$$K = \begin{pmatrix} K_{11} & \cdots & K_{1N} \\ \vdots & & \vdots \\ K_{N1} & \cdots & K_{NN} \end{pmatrix} \quad \text{on} \quad \bigoplus_{k=1}^{N} L^2(I_k) \quad \text{with matrix kernel } K_{ij}(x,y)$$

representable as a *single* integral operator on

$$L^2\left(\coprod_{k=1}^{N} I_k\right) \cong \bigoplus_{k=1}^{N} L^2(I_k), \qquad \coprod_{k=1}^{N} I_k = \bigcup_{k=1}^{N} I_k \times \{k\},$$

with *scalar* kernel (Fredholm 1903)

$$K(x,y) = \sum_{i,j=1}^{N} \mathbb{1}_{I_i}(x) K_{ij}(x,y) \mathbb{1}_{I_j}(y)$$

$\rightsquigarrow$ straightforward extension of the quadrature method

n-th largest level in edge scaled GSE

$$\mathbb{P}(\text{exactly } n \text{ levels lie in } (s,\infty)) = E_4(n;s) = \frac{(-1)^n}{n!}\,\frac{\partial^n}{\partial z^n} F_4(s;z)\Big|_{z=1}$$

(Forrester/Nagao/Honner '99, Tracy/Widom '05)

$$\boxed{F_4(s;z) = \det\left(I - \frac{z}{2}\begin{pmatrix} S(x,y) & SD(x,y) \\ IS(x,y) & S(y,x)\end{pmatrix}\restriction_{L^2(s,\infty)\oplus L^2(s,\infty)}\right)^{1/2}}$$

$$S(x,y) = K_{\mathrm{Ai}}(x,y) - \tfrac{1}{2}\,\mathrm{Ai}(x)\int_y^\infty \mathrm{Ai}(\eta)\,d\eta$$

$$SD(x,y) = -\partial_y K_{\mathrm{Ai}}(x,y) - \tfrac{1}{2}\,\mathrm{Ai}(x)\,\mathrm{Ai}(y)$$

$$IS(x,y) = -\int_x^\infty K_{\mathrm{Ai}}(\xi,y)\,d\xi + \tfrac{1}{2}\int_x^\infty \mathrm{Ai}(\xi)\,d\xi \int_y^\infty \mathrm{Ai}(\eta)\,d\eta$$

$$K_{\mathrm{Ai}}(x,y) = \frac{\mathrm{Ai}(x)\,\mathrm{Ai}'(y) - \mathrm{Ai}'(x)\,\mathrm{Ai}(y)}{x-y}$$

factorization

$$K_{\mathrm{Ai}}(x,y) = \int_0^\infty K(x,\xi)K(\xi,y)\,d\xi, \qquad K(x,y) = \mathrm{Ai}(x+y),$$

yields

$$F_2(s;z) = F_+(s;z)\cdot F_-(s;z), \qquad F_\pm(s;z) = \det\left(I \mp \sqrt{z}\,K\restriction_{L^2(s/2,\infty)}\right)$$

and (Ferrari/Spohn '05)

$$F_4(s;1) = \tfrac{1}{2}\left(F_+(s;1) + F_-(s;1)\right)$$

A new formula (B. '10)

How about, in general, $\boxed{F_4(s;z) = \tfrac{1}{2}\left(F_+(s;z) + F_-(s;z)\right)}$?

- *first*, numerical tests with random s and z indicated the formula to be *true*
- *later*, proof via Painlevé II representation (B. '10, Forrester '06)

n-th largest level in edge scaled GOE

$$\mathbb{P}(\text{exactly } n \text{ levels lie in } (s,\infty)) = E_1(n;s) = \frac{(-1)^n}{n!}\,\frac{\partial^n}{\partial z^n} F_1(s;z)\Big|_{z=1}$$

(Forrester/Nagao/Honner '99, Tracy/Widom '05)

$$\boxed{F_1(s;z) = \det\left(I - z\begin{pmatrix} S(x,y) & SD(x,y) \\ IS(x,y) & S(y,x)\end{pmatrix}\upharpoonright_{X_1(s,\infty)\oplus X_2(s,\infty)}\right)^{1/2}}$$

$$S(x,y) = K_{\mathrm{Ai}}(x,y) + \tfrac{1}{2}\left(1 - \tfrac{1}{2}\,\mathrm{Ai}(x)\int_y^\infty \mathrm{Ai}(\eta)\,d\eta\right)$$

$$SD(x,y) = -\partial_y K_{\mathrm{Ai}}(x,y) - \tfrac{1}{2}\,\mathrm{Ai}(x)\,\mathrm{Ai}(y)$$

$$IS(x,y) = -\tfrac{1}{2}\operatorname{sgn}(x-y) - \int_x^\infty K_{\mathrm{Ai}}(\xi,y)\,d\xi + \tfrac{1}{2}\left(\int_y^x \mathrm{Ai}(\xi)\,d\xi + \int_x^\infty \mathrm{Ai}(\xi)\,d\xi\,\int_y^\infty \mathrm{Ai}(\eta)\,d\eta\right)$$

$$K_{\mathrm{Ai}}(x,y) = \frac{\mathrm{Ai}(x)\,\mathrm{Ai}'(y) - \mathrm{Ai}'(x)\,\mathrm{Ai}(y)}{x-y}$$

Hilbert–Schmidt operator with trace class diagonal $\rightsquigarrow$ *Hilbert–Carleman determinant* superposition/decimation relations (Forrester/Rains '01)

$$\mathrm{GSE}_m = \mathrm{even}(\mathrm{GOE}_{2m+1}), \qquad \mathrm{GUE}_m = \mathrm{even}(\mathrm{GOE}_m \cup \mathrm{GOE}_{m+1})$$

combinatorical implications

- $F_4(n;s) = F_1(2n;s)$
- $E_2(n;s) = \sum_{j=0}^{2n} E_1(j;s)E_1(2n-j;s) + \sum_{j=0}^{2n+1} E_1(j;s)E_1(2n+1-j;s)$

other relations

- $E_2(n;s) = \sum_{j=0}^{n} E_+(j;s)E_-(n-j;s)$
- $E_4(n;s) = \frac{1}{2}\left(E_+(n;s) + E_-(n;s)\right)$
- $E_1(0;s) = E_+(0;s)$ (Ferrari/Spohn '05)

generating functions

$$f_0(x) = \sum_{n=0}^{\infty} E_1(2n;s)x^{2n}$$

$$f_1(x) = \sum_{n=0}^{\infty} E_1(2n+1;s)x^{2n+1}$$

$$g_{\pm}(x) = \sum_{n=0}^{\infty} E_{\pm}(n;s)x^{2n} = F_{\pm}(1-x^2;s)$$

equations

$$f_0(x)^2 + 2f_0(x)f_1(x) + x^2 f_1(x)^2 = g_+(x)\cdot g_-(x)$$

$$f_0(x) + f_1(x) = \frac{1}{2}\left(g_+(x) + g_-(x)\right)$$

$$f_0(0) = g_+(0)$$

solution

$$f_0(x) = g_+(x) - (1 - \sqrt{1 - x^2}) f_0(x)$$

$$f_1(x) = \frac{1}{2}\left(g_+(x) + g_-(x)\right) - f_0(x)$$

transforms with

$$1 - \sqrt{1 - x^2} = \sum_{k=0}^{\infty} \frac{\binom{2k}{k}}{2^{2k+1}(k+1)} x^{2k+2}$$

into the recursion

$$E_1(2n; s) = E_+(n; s) - \sum_{k=0}^{n-1} \frac{\binom{2k}{k} E_1(2n - 2k - 1; s)}{2^{2k+1}(k+1)}$$

$$E_1(2n + 1; s) = \frac{E_+(n; s) + E_-(n; s)}{2} - E_1(2n; s)$$

A reformulation of the GOE recursion

(B. '10, Forrester '06)

$$F_1(s;z) = \sum_{n=0}^{\infty} E_1(n;s)(1-z)^n$$

$$= \frac{1}{2}\sum_{\pm}\left(1 \pm \sqrt{\frac{z}{2-z}}\right) \det\left(I \mp \sqrt{z(2-z)}\, K\restriction_{L^2(s/2,\infty)}\right)$$

with $K(x,y) = \mathrm{Ai}(x+y)$

- amenable to exponential convergence in the quadrature method
- branch-cut singularities at $z=0$ and $z=2$

  $\rightsquigarrow$ derivatives *not* along circles (new: automatic contours B./Wechslberger '11)

---

similar structure for

- bulk scaling limits
- hard-edge scaling limits (LOE/LUE/LSE)

an e-mail from a geneticist @ Broad Institute (MIT/Harvard)

> `I'm ... interested in the distribution of eigenvalues of very large (mostly Wishart) matrices. I recently found you ArXiv paper. A tremendous amount of information. ... Some things I would like: A table of the mean of F1(k,s) for as large k as is practical. I'd certainly like to get this for k = 1, ..., 50.`

---

mean and variance of the k-largest level in edge-scaled GOE *by using the recursion*

1	-1.2065335745	1.6077810345
2	-3.2624279028	1.0354474415
3	-4.8216302757	0.8223901151
4	-6.1620399636	0.7031581054
5	-7.3701147042	0.6242523679
6	-8.4862183723	0.5670071487
7	-9.5331810321	0.5229902526
.	.	.
.	.	.
.	.	.
38	-31.3497235299	0.2159078706
39	-31.9083474476	0.2129575824
40	-32.4621235403	0.2101204063
41	-33.011211545_	0.207380416_
42	-33.5557807___	0.2047596___
43	-34.09586_____	0.20201_____
44	-34.631_______	0.198_______
45	-35.14________	0.18________
46	-35.5_________	0.1_________
47	-3_.__________	0.__________

RMFredholmToolbox for Matlab (B. '10)

function	command
$E_2^{(n)}(k;J)$	`E(2,k,J,n)`
$E_2^{(n)}((k,0);J_1,J_2)$	`E(2,[k,0],{J1,J2},n)`
$E_2^{(\text{bulk})}(k;J)$	`E(2,k,J,'bulk')`
$E_2^{(\text{bulk})}((k,0);J_1,J_2)$	`E(2,[k,0],{J1,J2},'bulk')`
$E_2^{(\text{soft})}(k;J)$	`E(2,k,J,'soft')`
$E_2^{(\text{soft})}((k,0);J_1,J_2)$	`E(2,[k,0],{J1,J2},'soft')`
$F(x,y)$	`F2Joint(x,y)`
$E_4^{(\text{soft})}(k;J)$	`E(4,k,J,'soft','MatrixKernel')`
$E_{\text{LUE}}^{(n)}(k;J,\alpha)$	`E('LUE',k,J,n,alpha)`
$E_{\text{LUE}}^{(n)}((k,0);J_1,J_2,\alpha)$	`E('LUE',[k,0],{J1,J2},n,alpha)`
$E_2^{(\text{hard})}(k;J,\alpha)$	`E(2,k,J,'hard',alpha)`
$E_2^{(\text{hard})}((k,0);J_1,J_2,\alpha)$	`E(2,[k,0],{J1,J2},'hard',alpha)`

function	command
$E_2^{(\text{hard})}((k,0);J_1,J_2,\alpha)$	`E(2,[k,0],{J1,J2},'hard',alpha)`
$E_+(k;s)$	`E('+',k,s)`
$E_-(k;s)$	`E('-',k,s)`
$E_\beta(k;s)$	`E(beta,k,s)`
$\tilde{E}_+(k;s)$	`E('+',k,s,'soft')`
$\tilde{E}_-(k;s)$	`E('-',k,s,'soft')`
$\tilde{E}_\beta(k;s)$	`E(beta,k,s,'soft')`
$F_\beta(k;s)$	`F(beta,k,s)`
$F_\beta(s)$	`F(beta,s)`
$E_{+,\alpha}(k;s)$	`E('+',k,s,'hard',alpha)`
$E_{-,\alpha}(k;s)$	`E('-',k,s,'hard',alpha)`
$E_{\beta,\alpha}(k;s)$	`E(beta,k,s,'hard',alpha)`
$F_{\beta,\alpha}(k;s)$	`F(beta,k,s,alpha)`

send e-mail to: `bornemann@tum.de`

References

– F. Bornemann
On the numerical evaluation of Fredholm determinants
Math. Comp. 79 (2010) 871-915

– F. Bornemann
Asymptotic independence of the extreme eigenvalues of GUE
J. Math. Phys. 51 (2010) 023514, 8pp.

– F. Bornemann
On the numerical evaluation of distributions in random matrix theory
Markov Processes Relat. Fields 16, 803–866 (2010)

– F. Bornemann
Accuracy and stability of computing high-order derivatives of analytic functions by Cauchy integrals
Found. Comput. Math. 11, 1–63, 2011

– F. Bornemann, G. Wechslberger
Optimal contours for high-order derivatives
8pp., `arXiv:1107.0498v1`, 2011