Free Probability Theory

and

Random Matrices

Roland Speicher
Queen's University
Kingston, Canada

What is Operator-Valued Free Probability and Why Should Engineers Care About it?

Roland Speicher
Queen's University
Kingston, Canada

Many approximations for calculating eigenvalue distribution of matrices consist in replacing

independent Gaussian random variables	$\longrightarrow$	free (semi)circular variables

Reasons for doing so:

- in limit $N \to \infty$, we have this transition asymptotically
- but even for finite N, this approximation is usually quite close to original problem
- this is an approximation which is usually exactly calculable for each N

Example: selfadjoint Gaussian $N \times N$ random matrix

$$X_N = (x_{ij})_{i,j=1}^N$$

with

- x_{ij} $(1 \leq i \leq j \leq)$ are independent complex $(i \neq j)$ or real $(i = j)$ Gaussian random variables

- $x_{ij} = \overline{x_{ji}}$

- $\varphi(x_{ij}) = 0, \qquad \varphi(x_{ij}\overline{x_{ij}}) = 1/N$

replacing

independent Gaussian variables

by

free (semi)circular variables

gives

selfadjoint noncommutative $N \times N$ random matrix

$$S_N = (c_{ij})_{i,j=1}^N$$

with

- c_{ij} $(1 \leq i \leq j \leq)$ are free circular $(i \neq j)$ or semicircular $(i = j)$ random variables

- $c_{ij} = c_{ji}^*$

- $\varphi(c_{ij}) = 0, \qquad \varphi(c_{ij}c_{ij}^*) = 1/N$

$$X_N = (x_{ij})_{i,j=1}^N \qquad \longrightarrow \qquad S_N = (c_{ij})_{i,j=1}^N$$

- X_N has a complicated averaged eigenvalue distribution (i.e., with respect to $\text{tr} \otimes \varphi$)

- S_N has a very simple distribution with respect to $\text{tr} \otimes \varphi$: for each N, S_N is a semicircular variable

$$X_N = (x_{ij})_{i,j=1}^N \qquad \longrightarrow \qquad S_N = (c_{ij})_{i,j=1}^N$$

Morale: S_N is an approximation for X_N with

- the approximation gets better for large N
- distr(S_N) can be calculated exactly

Why is S_N better treatable than X_N?

- taking matrices does not fit well with independence and Gaussianity

- freeness and (semi)circular variables, on the other hand, go very nicely with matrices freeness and (semi)circular variables go very nicely with matrices

caveat: this is true for free (semi)circulars which are centered and all have same variance freeness and (semi)circular variables go very nicely with matrices

caveat: this is true for free (semi)circulars which are centered and all have same variance

however: we might be interested in more general situations:

- x_{ij} might have different variances for different ij
- x_{ij} might not be centered ($\rightarrow$ Ricean model)
- there might even be correlations between different x_{ij} and x_{kl}

Example: non-zero mean, independent Gaussian variables with constant variance

$$Y_N = A_N + X_N$$

$\uparrow$	$\uparrow$
deterministic matrix of means	independent centered Gaussians constant variance

We replace this by . . .

$$A_N \quad + \quad S_N$$

A_N		S_N
$\uparrow$		$\uparrow$
same deterministic matrix as before		free centered (semi)circulars of constant variance

We have:

- A_N is free from S_N , thus
- $\text{distr}(A_N + S_N) = \text{distr}(A_N) \boxplus \text{distr}(S_N)$

Example: non-zero mean, independent Gaussian variables with varying variance

$$Y_N \quad = \quad A_N \quad + \quad X_N$$

$\uparrow$	$\uparrow$
deterministic matrix of means	independent centered Gaussians $\varphi(x_{ij}\overline{x_{ij}}) = \sigma_{ij}/N$

We replace this by . . .

$$A_N \qquad + \qquad S_N$$

$$\uparrow \qquad\qquad\qquad \uparrow$$

same deterministic matrix as before	free centered (semi)circulars with $\varphi(c_{ij}c_{ij}^*) = \sigma_{ij}/N$

We replace this by . . .

$$A_N \qquad + \qquad S_N$$

$$\uparrow \qquad\qquad \uparrow$$

same deterministic matrix as before	free centered (semi)circulars with $\varphi(c_{ij}c_{ij}^*) = \sigma_{ij}/N$

Now we have a **problem**

- S_N is **not semicircular** in general
- A_N and S_N are **not free** in general

So what do we gain by replacing

independent Gaussians

by

free (semi)circulars

in such a case?

$$X = (x_{ij})_{i,j=1}^{N} \qquad Y = (y_{kl})_{k,l=1}^{N}$$

with

$\{x_{ij}\}$ and $\{y_{kl}\}$ independent

$$\Downarrow$$

X and Y are not independent

$$X = (x_{ij})_{i,j=1}^{N} \qquad Y = (y_{kl})_{k,l=1}^{N}$$

with

$\{x_{ij}\}$ and $\{y_{kl}\}$ independent

$$\Downarrow$$

X and Y are not independent

actually: relation between X and Y is untreatable

$$X = (x_{ij})_{i,j=1}^{N} \qquad Y = (y_{kl})_{k,l=1}^{N}$$

with

$\{x_{ij}\}$ and $\{y_{kl}\}$ free

$$\Downarrow$$

X and Y are not free

$$X = (x_{ij})_{i,j=1}^{N} \qquad Y = (y_{kl})_{k,l=1}^{N}$$

with

$\{x_{ij}\}$ and $\{y_{kl}\}$ free

$$\Downarrow$$

X and Y are not free

however: relation between X and Y is more complicated, but still treatable in terms of

operator-valued freeness

Let $(\mathcal{C}, \varphi)$ be non-commutative probability space.

Consider $N \times N$ matrices over $\mathcal{C}$:

$$M_N(\mathcal{C}) := \{(a_{ij})_{i,j=1}^N \mid a_{ij} \in \mathcal{C}\}$$

$$M_N(\mathcal{C}) = M_N(\mathbb{C}) \otimes \mathcal{C}$$

is a non-commutative probability space with respect to

$$\mathrm{tr}_N \otimes \varphi,$$

but there is also an intermediate level

Instead of

$$
\begin{array}{c}
M_N(\mathcal{C}) \\
\downarrow \mathsf{tr} \otimes \varphi \\
\mathbb{C}
\end{array}
$$

consider ...

$$
\begin{array}{ccc}
M_N(\mathcal{C}) = M_N(\mathbb{C}) \otimes \mathcal{C} & & \\
\downarrow \mathsf{id} \otimes \varphi & & \Big| \\
M_N(\mathbb{C}) & & \Big| \ \mathsf{tr} \otimes \varphi \\
\downarrow \mathsf{tr} & & \downarrow \\
\mathbb{C} & &
\end{array}
$$

$$
\begin{array}{ccl}
M_N(\mathcal{C}) = M_N(\mathbb{C}) \otimes \mathcal{C} =: \mathcal{A} & & \\
\downarrow \mathrm{id} \otimes \varphi =: E & & \Big| \\
M_N(\mathbb{C}) =: \mathcal{B} & & \Big| \ \mathrm{tr} \otimes \varphi \\
\downarrow \mathrm{tr} & & \downarrow \\
\mathbb{C} & &
\end{array}
$$

Let $\mathcal{B} \subset \mathcal{A}$. A linear map

$$E : \mathcal{A} \to \mathcal{B}$$

is a **conditional expectation** if

$$E[b] = b \qquad \forall b \in \mathcal{B}$$

and

$$E[b_1 a b_2] = b_1 E[a] b_2 \qquad \forall a \in \mathcal{A}, \quad \forall b_1, b_2 \in \mathcal{B}$$

An **operator-valued probability space** consists of $\mathcal{B} \subset \mathcal{A}$ and a conditional expectation $E : \mathcal{A} \to \mathcal{B}$

Consider an operator-valued probability space $(\mathcal{A}, E : \mathcal{A} \to \mathcal{B})$. The **operator-valued distribution** of $a \in \mathcal{A}$ is given by all operator-valued moments

$$E[ab_1ab_2 \cdots b_{n-1}a] \in \mathcal{B} \qquad (n \in \mathbb{N}, b_1, \dots, b_{n-1} \in \mathcal{B})$$

Consider an operator-valued probability space $(\mathcal{A}, E : \mathcal{A} \to \mathcal{B})$. The **operator-valued distribution** of $a \in \mathcal{A}$ is given by all operator-valued moments

$$E[ab_1ab_2 \cdots b_{n-1}a] \in \mathcal{B} \qquad (n \in \mathbb{N}, b_1, \dots, b_{n-1} \in \mathcal{B})$$

Random variables $x, y \in \mathcal{A}$ are **free with respect to E** (or **free with amalgamation over $\mathcal{B}$**) if

$$E[p_1(x)q_1(y)p_2(x)q_2(y) \cdots] = 0$$

whenever p_i, q_j are polynomials with coefficients from $\mathcal{B}$ and

$$E[p_i(x)] = 0 \quad \forall i \qquad \text{and} \qquad E[q_j(y)] = 0 \quad \forall j.$$

Note: polynomials in x with coefficients from $\mathcal{B}$ are of the form

- x^2
- $b_0 x^2$

Note: polynomials in x with coefficients from $\mathcal{B}$ are of the form

- x^2

- $b_0 x^2$

- $b_1 x b_2 x b_3$

b's and x do not commute in general!

Note: polynomials in x with coefficients from $\mathcal{B}$ are of the form

- x^2
- b_0x^2
- $b_1xb_2xb_3$
- $b_1xb_2xb_3 + b_4xb_5xb_6 + \cdots$
- etc.

b's and x do not commute in general!

Operator-valued freeness works mostly like ordinary freeness, one only has to take care of the order of the variables; in all expressions they have to appear in their original order!

Example: If x and $\{y_1, y_2\}$ are free, then one has

$$E[y_1 x y_2] = E\Big[y_1 E[x] y_2\Big];$$

and more general

$$E[y_1 b_1 x b_2 y_2] = E\Big[y_1 b_1 E[x] b_2 y_2\Big].$$

Consider $E : \mathcal{A} \to \mathcal{B}$.

Define **free cumulants**

$$\kappa_n^{\mathcal{B}} : \mathcal{A}^n \to \mathcal{B}$$

by

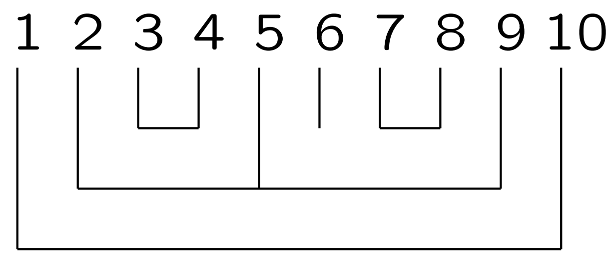
$$E[a_1 \cdots a_n] = \sum_{\pi \in NC(n)} \kappa_\pi^{\mathcal{B}}[a_1, \dots, a_n]$$

- arguments of $\kappa_\pi^{\mathcal{B}}$ are distributed according to blocks of π

- but now: cumulants are nested inside each other according to nesting of blocks of π

Example:

$$\pi = \Big\{\{1,10\},\{2,5,9\},\{3,4\},\{6\},\{7,8\}\Big\} \in NC(10),$$



$$\kappa_\pi^{\mathcal{B}}[a_1,\dots,a_{10}]$$

$$= \kappa_2^{\mathcal{B}}\Big(a_1 \cdot \kappa_3^{\mathcal{B}}\big(a_2 \cdot \kappa_2^{\mathcal{B}}(a_3,a_4), a_5 \cdot \kappa_1^{\mathcal{B}}(a_6) \cdot \kappa_2^{\mathcal{B}}(a_7,a_8), a_9\big), a_{10}\Big)$$

For $a \in \mathcal{A}$ define its **operator-valued Cauchy transform**

$$G_a(b) := E[\frac{1}{b-a}] = \sum_{n \geq 0} E[b^{-1}(ab^{-1})^n]$$

and **operator-valued R-transform**

$$\begin{aligned} R_a(b) &:= \sum_{n \geq 0} \kappa_{n+1}^{\mathcal{B}}(ab, ab, \ldots, ab, a) \\ &= \kappa_1^{\mathcal{B}}(a) + \kappa_2^{\mathcal{B}}(ab, a) + \kappa_3^{\mathcal{B}}(ab, ab, a) + \cdots \end{aligned}$$

Then

$$bG(b) = 1 + R(G(b)) \cdot G(b) \qquad \text{or} \qquad G(b) = \frac{1}{b - R(G(b))}$$

If x and y are free over $\mathcal{B}$, then

- mixed $\mathcal{B}$-valued cumulants in x and y vanish

- $R_{x+y}(b) = R_x(b) + R_y(b)$

- $G_{x+y}(b) = G_x\Big[b - R_y\Big(G_{x+y}(b)\Big)\Big]$ subordination

If s is a semicircle over $\mathcal{B}$ then

$$R_s(b) = \eta(b)$$

where $\eta : \mathcal{B} \to \mathcal{B}$ is a linear map given by

$$\eta(b) = E[sbs].$$

Back to random matrices

What can we say about

$$A_N \quad + \quad S_N$$

$$\uparrow \qquad\qquad \uparrow$$

deterministic matrix of means	free centered (semi)circulars with $\varphi(c_{ij}c_{ij}^*) = \sigma_{ij}/N$

Consider

$$T_N := A_N + S_N$$

We want Cauchy transform

$$g_T(z) = \mathrm{tr} \otimes \varphi[\frac{1}{z - T}]$$

Consider

$$T_N := A_N + S_N$$

We want Cauchy transform

$$g_T(z) = \mathrm{tr} \otimes \varphi[\frac{1}{z - T}]$$

$$\begin{array}{ccc} M_N(\mathcal{C}) = M_N(\mathbb{C}) \otimes \mathcal{C} & & \\ \downarrow \mathrm{id} \otimes \varphi & & | \\ & & | \\ M_N(\mathbb{C}) & & | \ \mathrm{tr} \otimes \varphi \\ & & | \\ \downarrow \mathrm{tr} & & \downarrow \\ \mathbb{C} & & \end{array}$$

Consider

$$T := A + S$$

We want Cauchy transform

$$g_T(z) = \mathrm{tr} \otimes \varphi[\frac{1}{z-T}]$$

$$\begin{aligned} g_T(z) &= \mathrm{tr} \otimes \varphi[\frac{1}{z-T}] \\ &= \mathrm{tr}\Big[\underbrace{\mathrm{id} \otimes \varphi(\frac{1}{z-T})}_{G_T(z)}\Big] = \mathrm{tr}[G_T(z)] \end{aligned}$$

We have nice behavior as $M_N(\mathbb{C})$-operator-valued objects

- A_N, S_N are free over $M_N(\mathbb{C})$, i.e.,

$$G_T(z) = G_A\Big[z - R_S\Big(G_T(z)\Big)\Big]$$

- S_N is semicircular over $M_N(\mathbb{C})$, i.e.,

$$R_S(B) = \eta(B)$$

  with $\eta : M_N(\mathbb{C}) \to M_N(\mathbb{C})$ linear, given by $\eta(B) = E[SBS]$.

Thus the distribution of $T_N = A_N + S_N$ is determined via its Cauchy-transform g_T according to

$$g_T(z) = \mathrm{tr}\Big[G_T(z)\Big]$$

and

$$G_T(z) = G_A\left[z - \eta\Big(G_T(z)\Big)\right] = E\left[\frac{1}{z - \eta\Big(G_T(z)\Big) - A}\right]$$

Thus the distribution of $T_N = A_N + S_N$ is determined via its Cauchy-transform g_T according to

$$g_T(z) = \text{tr}\Big[G_T(z)\Big]$$

and

$$G_T(z) = G_A\left[z - \eta\Big(G_T(z)\Big)\right] = E\left[\frac{1}{z - \eta\Big(G_T(z)\Big) - A}\right]$$

Note: by [Helton, Far, Speicher IMRN 2007], there exists exactly one solution of above fixed point equation with the right positivity property!

Note: the more "symmetries" we have in entries of S_N, the "better" is the freeness between A_N and S_N!

For considered situation where different c_{ij} are free, we have for

$$\eta : M_N(\mathbb{C}) \to M_N(\mathbb{C})$$

that actually

$$[\eta(B)]_{ij} = E[SBS]_{ij} = \sum_{k,l} \varphi(c_{ik} b_{kl} c_{lj}) = \sum_{k,l} \underbrace{\varphi(c_{ik} c_{jl}^*)}_{\delta_{kl}\delta_{ij}\sigma_{ik}} b_{kl} = \delta_{ij} \sum_k \sigma_{ik} b_{kk}$$

thus η is effectively a mapping on diagonal matrices

Consider in addition to

$$E_M : M_N(\mathcal{C}) \to M_N(\mathbb{C}); \begin{pmatrix} a_{11} & \dots & a_{1N} \\ \vdots & \ddots & \dots \\ a_{N1} & \dots & a_{NN} \end{pmatrix} \mapsto \begin{pmatrix} \varphi(a_{11}) & \dots & \varphi(a_{1N}) \\ \vdots & \ddots & \vdots \\ \varphi(a_{N1}) & \dots & \varphi(a_{NN}) \end{pmatrix}$$

also

$$D_N(\mathbb{C}) := \{\text{diagonal matrices}\} \subset M_N(\mathbb{C})$$

and corresponding conditional expectation

$$E_D : M_N(\mathcal{C}) \to D_N(\mathbb{C}); \begin{pmatrix} a_{11} & \dots & a_{1N} \\ \vdots & \ddots & \dots \\ a_{N1} & \dots & a_{NN} \end{pmatrix} \mapsto \begin{pmatrix} \varphi(a_{11}) & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \varphi(a_{NN}) \end{pmatrix}$$

Then we have in our situation

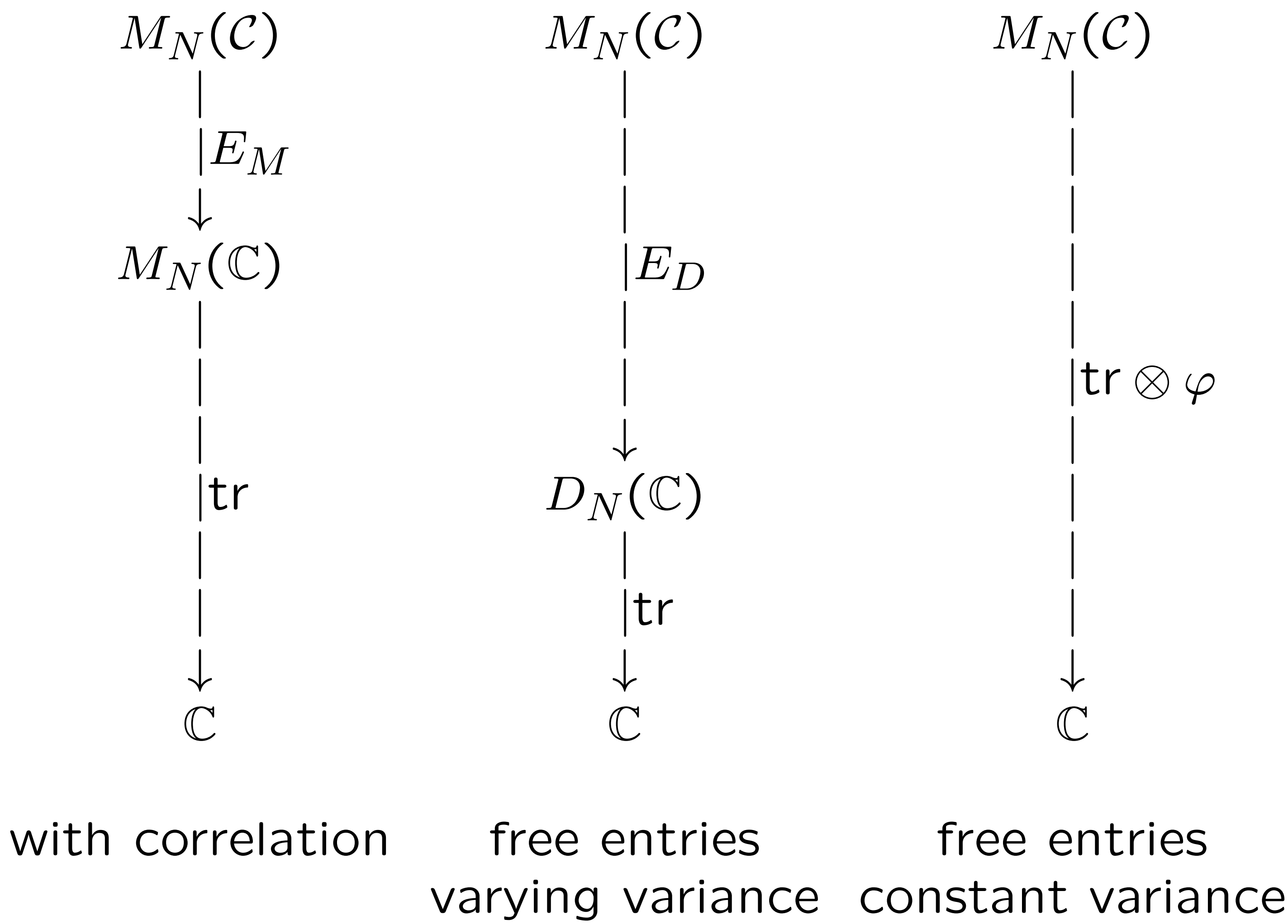
$$A_N \qquad + \qquad S_N$$

$$\uparrow \qquad\qquad \uparrow$$

deterministic matrix of means	free centered (semi)circulars with $\varphi(c_{ij}c_{ij}^*) = \sigma_{ij}/N$

that actually, by [Nica, Shlyakhtenko, Speicher, JFA 2002],

- A_N and S_N are free over $D_N(\mathbb{C})$
- S_N is semicircular over $D_N(\mathbb{C})$


$M_N(\mathcal{C})$
E_M
$M_N(\mathbb{C})$
tr
$\mathbb{C}$
with correlation
$M_N(\mathcal{C})$
E_D
$D_N(\mathbb{C})$
tr
$\mathbb{C}$
free entries
varying variance
$M_N(\mathcal{C})$
$\mathrm{tr} \otimes \varphi$
$\mathbb{C}$
free entries
constant variance

Let us now treat more relevant non-selfadjoint situation

$$H_N = A_N + C_N$$

A_N ↑ deterministic matrix of means

C_N ↑ free circulars, no symmetry condition, $\varphi(c_{ij}c_{ij}^*) = \sigma_{ij}/N$

Calculate the distribution of HH^*!

HH^* has the same distribution as square of

$$T := \begin{pmatrix} 0 & H \\ H^* & 0 \end{pmatrix} = \begin{pmatrix} 0 & A \\ A^* & 0 \end{pmatrix} + \begin{pmatrix} 0 & C \\ C^* & 0 \end{pmatrix}$$

These are $2N \times 2N$ selfadjoint matrices of the type considered before.

We have

- $\begin{pmatrix} 0 & A \\ A^* & 0 \end{pmatrix}$ and $\begin{pmatrix} 0 & C \\ C^* & 0 \end{pmatrix}$ are free over $D_{2N}(\mathbb{C})$

- $\begin{pmatrix} 0 & C \\ C^* & 0 \end{pmatrix}$ is semicircular over $D_{2N}(\mathbb{C})$ with

$$\eta \begin{pmatrix} B_1 & 0 \\ 0 & B_2 \end{pmatrix} = \begin{pmatrix} \eta_1(B_2) & 0 \\ 0 & \eta_2(B_1) \end{pmatrix},$$

  where

$$\eta_1(B_2) = E_{D_N}[CD_2C^*]$$

$$\eta_2(B_1) = E_{D_N}[C^*D_1C]$$

We have

$$G_T(z) = zG_{T^2}(z^2)$$

and

$$G_{T^2}(z) = \begin{pmatrix} G_1(z) & 0 \\ 0 & G_2(z) \end{pmatrix}$$

Thus

$$\begin{aligned} zG_{T^2}(z^2) = G_T(z) &= G_{\begin{pmatrix} 0 & A \\ A^* & 0 \end{pmatrix}} \left[z - R_{\begin{pmatrix} 0 & C \\ C^* & 0 \end{pmatrix}}(G_T(z)) \right] \\ &= E_{D_{2N}} \left[\left(z - z\eta \begin{pmatrix} G_1(z^2) & 0 \\ 0 & G_2(z^2) \end{pmatrix} - \begin{pmatrix} 0 & A \\ A^* & 0 \end{pmatrix} \right)^{-1} \right] \\ &= E_{D_{2N}} \left[\begin{pmatrix} z - z\eta_1(G_2(z^2)) & -A \\ -A^* & z - z\eta_2(G_1(z^2)) \end{pmatrix}^{-1} \right] \end{aligned}$$

This yields

$$zG_1(z) = E_{D_N}\left[\left(1 - \eta_1(G_2(z)) + A_N \frac{1}{z - z\eta_2(G_1(z))} A_N^*\right)^{-1}\right]$$

and

$$zG_2(z) = E_{D_N}\left[\left(1 - \eta_2(G_1(z)) + A_N^* \frac{1}{z - z\eta_1(G_2(z))} A_N\right)^{-1}\right]$$

This yields

$$zG_1(z) = E_{D_N}\left[\left(1 - \eta_1(G_2(z)) + A_N \frac{1}{z - z\eta_2(G_1(z))} A_N^*\right)^{-1}\right]$$

and

$$zG_2(z) = E_{D_N}\left[\left(1 - \eta_2(G_1(z)) + A_N^* \frac{1}{z - z\eta_1(G_2(z))} A_N\right)^{-1}\right]$$

These are actually the fixed point equations of

[Hachem, Loubaton, Najim, Ann. Appl. Prob. 2007]

for a **deterministic equivalent** (a la Girko) of square of random matrix with non-centered, independent Gaussians with non-constant variance as entries.

Conclusion

- many approximations (like deterministic equivalents a la Girko) consist in replacing independent Gaussians by free (semi)circulars

-

-

-

Conclusion

- many approximations (like deterministic equivalents a la Girko) consist in replacing independent Gaussians by free (semi)circulars

- operator-valued free probability allows conceptual and streamlined treatment of those

-

-

Conclusion

- many approximations (like deterministic equivalents a la Girko) consist in replacing independent Gaussians by free (semi)circulars

- operator-valued free probability allows conceptual and streamlined treatment of those

- also convergence questions might be treated more uniformly by relying on asymptotic freeness results

-

Conclusion

- many approximations (like deterministic equivalents a la Girko) consist in replacing independent Gaussians by free (semi)circulars

- operator-valued free probability allows conceptual and streamlined treatment of those

- also convergence questions might be treated more uniformly by relying on asymptotic freeness results

- this approach also allows to treat classes of random matrices with correlation between entries