

# Quadratic Variation of High Dimensional Itô Processes

Claudio Heinrich,  
joint work with Mark Podolskij

September 15, 2014

# Table of Contents

- 1 Setting
- 2 Statement of the Main Result
- 3 Sketch of the Proof

# Section 1

## Setting

### $p$ -dimensional Itô process without drift

Stochastic process of the form

$$X_t = X_0 + \int_0^t f_s dW_s$$

where  $W$  is a  $p$ -dimensional Brownian motion.

### $p$ -dimensional Itô process without drift

Stochastic process of the form

$$X_t = X_0 + \int_0^t f_s dW_s$$

where  $W$  is a  $p$ -dimensional Brownian motion.

**Realized Covariance Matrix:**

$$[X]_p^n = \sum_{i=1}^n \Delta_i^n X \Delta_i^n X^*,$$

where  $\Delta_i^n X = X_{\frac{i}{n}} - X_{\frac{i-1}{n}}$ .

### $p$ -dimensional Itô process without drift

Stochastic process of the form

$$X_t = X_0 + \int_0^t f_s dW_s$$

where  $W$  is a  $p$ -dimensional Brownian motion.

#### Realized Covariance Matrix:

$$[X]_p^n = \sum_{i=1}^n \Delta_i^n X \Delta_i^n X^*,$$

where  $\Delta_i^n X = X_{\frac{i}{n}} - X_{\frac{i-1}{n}}$ .

- The RCV is an estimator for the quadratic variation (in 1)

$$[X] = \int_0^1 f_s f_s^* ds.$$

### $p$ -dimensional Itô process without drift

Stochastic process of the form

$$X_t = X_0 + \int_0^t f_s dW_s$$

where  $W$  is a  $p$ -dimensional Brownian motion.

#### Realized Covariance Matrix:

$$[X]_p^n = \sum_{i=1}^n \Delta_i^n X \Delta_i^n X^*,$$

where  $\Delta_i^n X = X_{\frac{i}{n}} - X_{\frac{i-1}{n}}$ .

- The RCV is an estimator for the quadratic variation (in 1)

$$[X] = \int_0^1 f_s f_s^* ds.$$

- $[X]_p^n \xrightarrow{\mathbb{P}} [X]$  if  $n \rightarrow \infty$ .

$$[X]_p^n = \sum_{i=1}^n \Delta_i^n X \Delta_i^n X^*$$

What happens if the dimension of the process  $p$  and the number of observations  $n$  both are large but of the same order of magnitude? We investigate the behavior of  $[X]_p^n$  when  $p \rightarrow \infty$ ,  $n \rightarrow \infty$ , and  $n/p \rightarrow c \in (0, \infty)$ .



$$[X]_p^n = \sum_{i=1}^n \Delta_i^n X \Delta_i^n X^*$$

What happens if the dimension of the process  $p$  and the number of observations  $n$  both are large but of the same order of magnitude? We investigate the behavior of  $[X]_p^n$  when  $p \rightarrow \infty$ ,  $n \rightarrow \infty$ , and  $n/p \rightarrow c \in (0, \infty)$ .

- High dimensional random matrix theory provides the framework for this investigation.

$$[X]_p^n = \sum_{i=1}^n \Delta_i^n X \Delta_i^n X^*$$

What happens if the dimension of the process  $p$  and the number of observations  $n$  both are large but of the same order of magnitude? We investigate the behavior of  $[X]_p^n$  when  $p \rightarrow \infty$ ,  $n \rightarrow \infty$ , and  $n/p \rightarrow c \in (0, \infty)$ .

- High dimensional random matrix theory provides the framework for this investigation.
- We restrict ourselves to processes with volatility process of the form

$$f_t = \sum_{l=1}^m T_l 1_{[t_{l-1}, t_l)}(t) \quad (1)$$

where  $0 = t_0 < \dots < t_m = 1$ , and  $T_1, \dots, T_m$  are  $p \times p$  nonrandom matrices.

It is straightforward to show that

$$[X]_p^n \stackrel{d}{=} \frac{1}{n} \sum_{l=1}^m T_l Y_l Y_l^* T_l^*$$

where  $Y_l$  are  $p \times [n(t_l - t_{l-1})]$  matrices containing i.i.d.  $\mathcal{N}(0, 1)$  variables.

It is straightforward to show that

$$[X]_p^n \stackrel{d}{=} \frac{1}{n} \sum_{l=1}^m T_l Y_l Y_l^* T_l^*$$

where  $Y_l$  are  $p \times [n(t_l - t_{l-1})]$  matrices containing i.i.d.  $\mathcal{N}(0, 1)$  variables.  
How to analyze  $[X]_p^n$  for  $p, n \rightarrow \infty$  with  $n/p \rightarrow c$ ?

It is straightforward to show that

$$[X]_p^n \stackrel{d}{=} \frac{1}{n} \sum_{l=1}^m T_l Y_l Y_l^* T_l^*$$

where  $Y_l$  are  $p \times [n(t_l - t_{l-1})]$  matrices containing i.i.d.  $\mathcal{N}(0, 1)$  variables.  
How to analyze  $[X]_p^n$  for  $p, n \rightarrow \infty$  with  $n/p \rightarrow c$ ?

**High Dimensional Random Matrix Theory:** Analyze the limiting spectral behavior of  $[X]_p^n$ :

It is straightforward to show that

$$[X]_p^n \stackrel{d}{=} \frac{1}{n} \sum_{l=1}^m T_l Y_l Y_l^* T_l^*$$

where  $Y_l$  are  $p \times [n(t_l - t_{l-1})]$  matrices containing i.i.d.  $\mathcal{N}(0, 1)$  variables. How to analyze  $[X]_p^n$  for  $p, n \rightarrow \infty$  with  $n/p \rightarrow c$ ?

**High Dimensional Random Matrix Theory:** Analyze the limiting spectral behavior of  $[X]_p^n$ :

For a diagonalisable  $p \times p$  matrix  $A$  with real eigenvalues  $\lambda_1, \dots, \lambda_p$  the *spectral distribution*  $F^A$  is defined by the p.d.f.

$$F^A(x) = \frac{1}{p} \#\{i : \lambda_i \leq x\}.$$

In other words,  $F^A(x)$  is the proportion of eigenvalues of  $A$  which are smaller or equal to  $x$ .

## Section 2

# Statement of the Main Result

## theorem 2.1

Let  $[X]_p^n = 1/n \sum_l T_l Y_l Y_l^* T_l^*$ . Assume

- (a) that there is  $\tau_0 > 0$  such that the largest eigenvalues of  $T_l T_l^*$  are bounded by  $\tau_0$  for all  $l$ , uniformly in  $p$ .
- (b) for all  $k > 0$  and for all  $\mathbf{l} = (l_1, \dots, l_k) \in \{1, \dots, m\}^k$  the existence of the mixed limiting spectral moments

$$M_{\mathbf{l}}^k = \lim_{p \rightarrow \infty} \frac{1}{p} \operatorname{tr} \left( \prod_{i=1}^k T_{l_i} T_{l_i}^* \right).$$

Then, the spectral distributions  $F^{[X]_p^n}$  converges weakly to a nonrandom p.d.f.  $F$ , a.s., which will be specified by its moment sequence. The moments  $\beta_k$  of  $F$  are of the form

$$\beta_k = \sum_{r=1}^k c^{r-1} \sum_{\nu_1 + \dots + \nu_r = k} \sum_{\mathbf{l} \in \{1, \dots, m\}^k} c_{r, \nu, \mathbf{l}} \prod_{a=1}^r M_{\mathbf{l}^{(a)}}^{\nu_a} \prod_{l=1}^m (t_l - t_{l-1})^{s_{\mathbf{l}, \nu, \mathbf{l}'}}.$$



## Section 3

### Sketch of the Proof

The theorem extends the results of a seminal paper (*Y.Q. Yin and P.R. Krishnaiah, 1983*), where the case  $1/n TYY^* T^*$ , i.e.  $T_1 = \dots = T_m$  was considered.

The theorem extends the results of a seminal paper (*Y.Q. Yin and P.R. Krishnaiah, 1983*), where the case  $1/n TYY^* T^*$ , i.e.  $T_1 = \dots = T_m$  was considered.

- The authors worked with the assumption that  $TT^*$  has a limiting spectral distribution.

The theorem extends the results of a seminal paper (*Y.Q. Yin and P.R. Krishnaiah, 1983*), where the case  $1/n TYY^* T^*$ , i.e.  $T_1 = \dots = T_m$  was considered.

- The authors worked with the assumption that  $TT^*$  has a limiting spectral distribution.
- In our framework, assuming the existence of LSDs for  $T_1 T_1^*, \dots, T_m T_m^*$  is not sufficient.

The theorem extends the results of a seminal paper (*Y.Q. Yin and P.R. Krishnaiah, 1983*), where the case  $1/n TYY^*T^*$ , i.e.  $T_1 = \dots = T_m$  was considered.

- The authors worked with the assumption that  $TT^*$  has a limiting spectral distribution.
- In our framework, assuming the existence of LSDs for  $T_1T_1^*, \dots, T_mT_m^*$  is not sufficient.
- The existence of all mixed limiting spectral moments implies the existence of LSDs for  $T_1T_1^*, \dots, T_mT_m^*$ .

The theorem extends the results of a seminal paper (*Y.Q. Yin and P.R. Krishnaiah, 1983*), where the case  $1/n TYY^* T^*$ , i.e.  $T_1 = \dots = T_m$  was considered.

- The authors worked with the assumption that  $TT^*$  has a limiting spectral distribution.
- In our framework, assuming the existence of LSDs for  $T_1 T_1^*, \dots, T_m T_m^*$  is not sufficient.
- The existence of all mixed limiting spectral moments implies the existence of LSDs for  $T_1 T_1^*, \dots, T_m T_m^*$ .
- Additionally, it ensures a nice 'joint spectral behavior' of  $T_1 T_1^*, \dots, T_m T_m^*$ .

**One way to determine limiting spectral distributions:**

## One way to determine limiting spectral distributions:

Moment convergence theorem,

### theorem 3.1

Let  $(F_n)$  be a sequence of p.d.f.s with finite moments of all orders  $\beta_{k,n} = \int x^k dF_n(x)$ . Assume  $\beta_{k,n} \rightarrow \beta_k$  for  $n \rightarrow \infty$  for  $k = 0, 1, \dots$  where

(a)  $\beta_k < \infty$  for all  $k$  and

(b)  $\sum_{k=0}^{\infty} [\beta_{2k}(F)]^{-\frac{1}{2k}} = \infty$ .

Then,  $F_n$  converges weakly to the unique probability distribution function  $F$  with moment sequence  $(\beta_k)_{k=0, \dots}$ .



## One way to determine limiting spectral distributions:

Moment convergence theorem,

### theorem 3.1

Let  $(F_n)$  be a sequence of p.d.f.s with finite moments of all orders  $\beta_{k,n} = \int x^k dF_n(x)$ . Assume  $\beta_{k,n} \rightarrow \beta_k$  for  $n \rightarrow \infty$  for  $k = 0, 1, \dots$  where

(a)  $\beta_k < \infty$  for all  $k$  and

(b)  $\sum_{k=0}^{\infty} [\beta_{2k}(F)]^{-\frac{1}{2k}} = \infty$ .

Then,  $F_n$  converges weakly to the unique probability distribution function  $F$  with moment sequence  $(\beta_k)_{k=0, \dots}$ .

Let  $\lambda_1 \leq \dots \leq \lambda_p$  be the eigenvalues of  $A$ , then

$$\beta_k(F^A) = \frac{1}{p} \sum_{i=1}^p \lambda_i^k = \frac{1}{p} \text{tr}(A^k).$$

In order to determine  $F = \lim F^{[X]_p^n}$  it is sufficient to find

$$\lim_{p, n \rightarrow \infty} \beta_k(F^{[X]_p^n}) = \lim_{p, n \rightarrow \infty} \frac{1}{p} \text{tr}([X]_p^n)^k,$$

for  $k = 1, 2, \dots$  .

In order to determine  $F = \lim F^{[X]_\rho^n}$  it is sufficient to find

$$\lim_{\rho, n \rightarrow \infty} \beta_k(F^{[X]_\rho^n}) = \lim_{\rho, n \rightarrow \infty} \frac{1}{\rho} \text{tr}([X]_\rho^n)^k,$$

for  $k = 1, 2, \dots$ .

It holds that

$$\mathbb{E}[1/\rho \text{tr}([X]_\rho^n)^k - \mathbb{E}[1/\rho \text{tr}([X]_\rho^n)^k]]^4$$

is summable.

In order to determine  $F = \lim F^{[X]_\rho^n}$  it is sufficient to find

$$\lim_{\rho, n \rightarrow \infty} \beta_k(F^{[X]_\rho^n}) = \lim_{\rho, n \rightarrow \infty} \frac{1}{\rho} \text{tr}([X]_\rho^n)^k,$$

for  $k = 1, 2, \dots$ .

It holds that

$$\mathbb{E}[1/\rho \text{tr}([X]_\rho^n)^k - \mathbb{E}[1/\rho \text{tr}([X]_\rho^n)^k]]^4$$

is summable. Therefore, by virtue of Borel-Cantelli Lemma, we have

$$\lim_{\rho, n \rightarrow \infty} \frac{1}{\rho} \text{tr}([X]_\rho^n)^k = \lim_{\rho, n \rightarrow \infty} \mathbb{E} \left[ \frac{1}{\rho} \text{tr}([X]_\rho^n)^k \right],$$

almost surely.

In order to determine  $F = \lim F^{[X]_\rho^n}$  it is sufficient to find

$$\lim_{\rho, n \rightarrow \infty} \beta_k(F^{[X]_\rho^n}) = \lim_{\rho, n \rightarrow \infty} \frac{1}{\rho} \operatorname{tr}([X]_\rho^n)^k,$$

for  $k = 1, 2, \dots$ .

It holds that

$$\mathbb{E}[1/\rho \operatorname{tr}([X]_\rho^n)^k] - \mathbb{E}[1/\rho \operatorname{tr}([X]_\rho^n)^k]^4$$

is summable. Therefore, by virtue of Borel-Cantelli Lemma, we have

$$\lim_{\rho, n \rightarrow \infty} \frac{1}{\rho} \operatorname{tr}([X]_\rho^n)^k = \lim_{\rho, n \rightarrow \infty} \mathbb{E} \left[ \frac{1}{\rho} \operatorname{tr}([X]_\rho^n)^k \right],$$

almost surely. Hence

$$\lim_{\rho, n \rightarrow \infty} \frac{1}{\rho} \operatorname{tr}([X]_\rho^n)^k = \lim_{\rho, n \rightarrow \infty} \frac{1}{\rho} \sum_{\mathbf{l} \in \{1, \dots, m\}^k} \mathbb{E} \left[ \operatorname{tr} \left( \prod_{i=1}^k T_{l_i} Y_{l_i} Y_{l_i}^* T_{l_i}^* \right) \right],$$

where  $\mathbf{l} = (l_1, \dots, l_k)$ .

Expansion leads to

$$\mathbb{E}[\beta_k(F^{[X]_p^n})]$$

$$= p^{-1} n^{-k} \sum_{\mathbf{i}, \mathbf{j}, \mathbf{l}} \mathbb{E} \left[ T_{l_1, i_1 i_2} \underbrace{Y_{l_1, i_2 j_1}}_{\sim \mathcal{N}(0,1)} Y_{l_1, j_1 i_3}^* T_{l_1, i_3 i_4}^* \cdots T_{l_k, i_{3k-2} i_{3k-1}} Y_{l_k, i_{3k-1} j_k} Y_{l_k, j_k i_{3k}}^* T_{l_k, i_{3k} i_1}^* \right]$$

where the summation runs for all  $\mathbf{i} \in \{1, \dots, p\}^{3k}$ ,  
 $\mathbf{l} = (l_1, \dots, l_k) \in \{1, \dots, m\}^k$ , and  $j_a \in [n(t_{l_a} - t_{l_{a-1}})]$

Expansion leads to

$$\mathbb{E}[\beta_k(F^{[X]_p^n})] = p^{-1} n^{-k} \sum_{\mathbf{i}, \mathbf{j}, \mathbf{l}} \mathbb{E} \left[ T_{l_1, i_1 i_2} \underbrace{Y_{l_1, i_2 j_1}}_{\sim \mathcal{N}(0,1)} Y_{l_1, j_1 i_3}^* T_{l_1, i_3 i_4}^* \cdots T_{l_k, i_{3k-2} i_{3k-1}} Y_{l_k, i_{3k-1} j_k} Y_{l_k, j_k i_{3k}}^* T_{l_k, i_{3k} i_1}^* \right]$$

where the summation runs for all  $\mathbf{i} \in \{1, \dots, p\}^{3k}$ ,

$\mathbf{l} = (l_1, \dots, l_k) \in \{1, \dots, m\}^k$ , and  $j_a \in [n(t_{l_a} - t_{l_{a-1}})]$

$\Rightarrow$  Combinatorial problem, solution uses graph theory.

We assign a colored graph to every summand

$$\mathbb{E} \left[ T_{l_1, i_1 i_2} Y_{l_1, i_2 j_1} Y_{l_1, j_1 i_3}^* T_{l_1, i_3 i_4}^* \cdots T_{l_k, i_{3k-2} i_{3k-1}} Y_{l_k, i_{3k-1} j_k} Y_{l_k, j_k i_{3k}}^* T_{l_k, i_{3k} i_1}^* \right].$$

We choose  $m$  colors which correspond to  $l_a \in \{1, \dots, m\}$ .



We assign a colored graph to every summand

$$\mathbb{E} \left[ T_{l_1, i_1 i_2} Y_{l_1, i_2 j_1} Y_{l_1, j_1 i_3}^* T_{l_1, i_3 i_4}^* \cdots T_{l_k, i_{3k-2} i_{3k-1}} Y_{l_k, i_{3k-1} j_k} Y_{l_k, j_k i_{3k}}^* T_{l_k, i_{3k} i_1}^* \right].$$

We choose  $m$  colors which correspond to  $l_a \in \{1, \dots, m\}$ . Here an example for  $m = 2$  and  $t_1 = 1/2$ :

We assign a colored graph to every summand

$$\mathbb{E} \left[ T_{l_1, i_1 i_2} Y_{l_1, i_2 j_1} Y_{l_1, j_1 i_3}^* T_{l_1, i_3 i_4}^* \cdots T_{l_k, i_{3k-2} i_{3k-1}} Y_{l_k, i_{3k-1} j_k} Y_{l_k, j_k i_{3k}}^* T_{l_k, i_{3k} i_1}^* \right].$$

We choose  $m$  colors which correspond to  $l_a \in \{1, \dots, m\}$ . Here an example for  $m = 2$  and  $t_1 = 1/2$ :

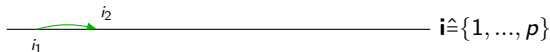
\_\_\_\_\_  $\hat{i} \in \{1, \dots, p\}$

\_\_\_\_\_  $\hat{j} \in \{1, \dots, \lfloor n/2 \rfloor\}$

We assign a colored graph to every summand

$$\mathbb{E} \left[ T_{l_1, i_1 i_2} Y_{l_1, i_2 j_1} Y_{l_1, j_1 i_3}^* T_{l_1, i_3 i_4}^* \cdots T_{l_k, i_{3k-2} i_{3k-1}} Y_{l_k, i_{3k-1} j_k} Y_{l_k, j_k i_{3k}}^* T_{l_k, i_{3k} i_1}^* \right].$$

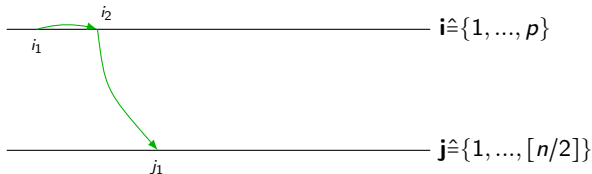
We choose  $m$  colors which correspond to  $l_a \in \{1, \dots, m\}$ . Here an example for  $m = 2$  and  $t_1 = 1/2$ :



We assign a colored graph to every summand

$$\mathbb{E} \left[ T_{l_1, i_1 i_2} Y_{l_1, i_2 j_1} Y_{l_1, j_1 i_3}^* T_{l_1, i_3 i_4}^* \cdots T_{l_k, i_{3k-2} i_{3k-1}} Y_{l_k, i_{3k-1} j_k} Y_{l_k, j_k i_{3k}}^* T_{l_k, i_{3k} i_1}^* \right].$$

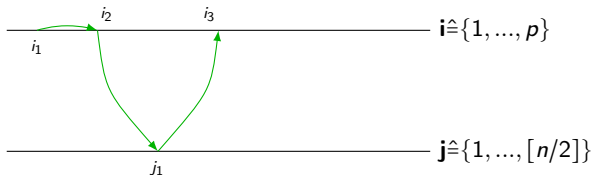
We choose  $m$  colors which correspond to  $l_a \in \{1, \dots, m\}$ . Here an example for  $m = 2$  and  $t_1 = 1/2$ :



We assign a colored graph to every summand

$$\mathbb{E} \left[ T_{l_1, i_1 i_2} Y_{l_1, i_2 j_1} Y_{l_1, j_1 i_3}^* T_{l_1, i_3 i_4}^* \cdots T_{l_k, i_{3k-2} i_{3k-1}} Y_{l_k, i_{3k-1} j_k} Y_{l_k, j_k i_{3k}}^* T_{l_k, i_{3k} i_1}^* \right].$$

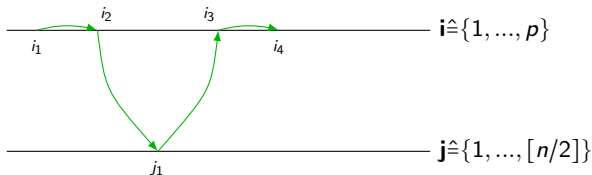
We choose  $m$  colors which correspond to  $l_a \in \{1, \dots, m\}$ . Here an example for  $m = 2$  and  $t_1 = 1/2$ :



We assign a colored graph to every summand

$$\mathbb{E} \left[ T_{l_1, i_1 i_2} Y_{l_1, i_2 j_1} Y_{l_1, j_1 i_3}^* T_{l_1, i_3 i_4}^* \cdots T_{l_k, i_{3k-2} i_{3k-1}} Y_{l_k, i_{3k-1} j_k} Y_{l_k, j_k i_{3k}}^* T_{l_k, i_{3k} i_1}^* \right].$$

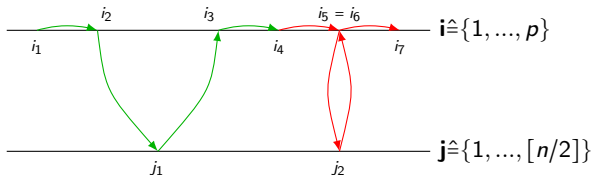
We choose  $m$  colors which correspond to  $l_a \in \{1, \dots, m\}$ . Here an example for  $m = 2$  and  $t_1 = 1/2$ :



We assign a colored graph to every summand

$$\mathbb{E} \left[ T_{l_1, i_1 i_2} Y_{l_1, i_2 j_1} Y_{l_1, j_1 i_3}^* T_{l_1, i_3 i_4}^* \cdots T_{l_k, i_{3k-2} i_{3k-1}} Y_{l_k, i_{3k-1} j_k} Y_{l_k, j_k i_{3k}}^* T_{l_k, i_{3k} i_1}^* \right].$$

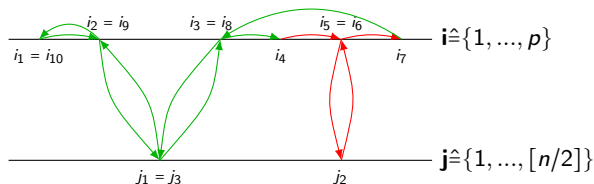
We choose  $m$  colors which correspond to  $l_a \in \{1, \dots, m\}$ . Here an example for  $m = 2$  and  $t_1 = 1/2$ :



We assign a colored graph to every summand

$$\mathbb{E} \left[ T_{l_1, i_1 i_2} Y_{l_1, i_2 j_1} Y_{l_1, j_1 i_3}^* T_{l_1, i_3 i_4}^* \cdots T_{l_k, i_{3k-2} i_{3k-1}} Y_{l_k, i_{3k-1} j_k} Y_{l_k, j_k i_{3k}}^* T_{l_k, i_{3k} i_1}^* \right].$$

We choose  $m$  colors which correspond to  $l_a \in \{1, \dots, m\}$ . Example for  $m = 2$  and  $t_1 = 1/2$ :

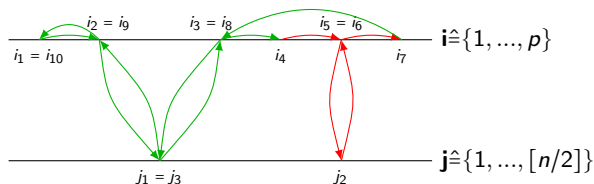




We assign a colored graph to every summand

$$\mathbb{E} \left[ T_{l_1, i_1 i_2} Y_{l_1, i_2 j_1} Y_{l_1, j_1 i_3}^* T_{l_1, i_3 i_4}^* \cdots T_{l_k, i_{3k-2} i_{3k-1}} Y_{l_k, i_{3k-1} j_k} Y_{l_k, j_k i_{3k}}^* T_{l_k, i_{3k} i_1}^* \right].$$

We choose  $m$  colors which correspond to  $l_a \in \{1, \dots, m\}$ . Example for  $m = 2$  and  $t_1 = 1/2$ :

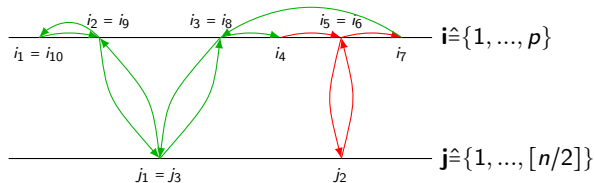


**Advantage:** One can divide the graphs into several categories according to their shape.

Example: Summing all summands

$$\mathbb{E} \left[ T_{l_1, i_1 i_2} Y_{l_1, i_2 j_1} Y_{l_1, j_1 i_3}^* T_{l_1, i_3 i_4}^* \cdots T_{l_k, i_{3k-2} i_{3k-1}} Y_{l_k, i_{3k-1} j_k} Y_{l_k, j_k i_{3k}}^* T_{l_k, i_{3k} i_1}^* \right]$$

corresponding to a graph with the shape

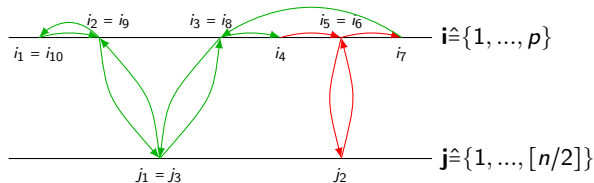


gives  $\approx \text{tr}(T_1 T_1^*) \text{tr}(T_1 T_1^* T_2 T_2^*)$  if green  $\hat{=}$  1 and red  $\hat{=}$  2.

Example: Summing all summands

$$\mathbb{E} \left[ T_{l_1, i_1 i_2} Y_{l_1, i_2 j_1} Y_{l_1, j_1 i_3}^* T_{l_1, i_3 i_4}^* \cdots T_{l_k, i_{3k-2} i_{3k-1}} Y_{l_k, i_{3k-1} j_k} Y_{l_k, j_k i_{3k}}^* T_{l_k, i_{3k} i_1}^* \right]$$

corresponding to a graph with the shape



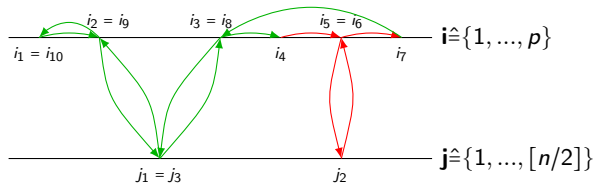
gives  $\approx \text{tr}(T_1 T_1^*) \text{tr}(T_1 T_1^* T_2 T_2^*)$  if green  $\hat{=}$  1 and red  $\hat{=}$  2.

- Expectation factor  $\mathbb{E}[Y_{l_1, i_2 j_1} \cdots Y_{l_k, j_k i_{3k}}^*]$  is 1.

Example: Summing all summands

$$\mathbb{E} \left[ T_{l_1, i_1 i_2} Y_{l_1, i_2 j_1} Y_{l_1, j_1 i_3}^* T_{l_1, i_3 i_4}^* \cdots T_{l_k, i_{3k-2} i_{3k-1}} Y_{l_k, i_{3k-1} j_k} Y_{l_k, j_k i_{3k}}^* T_{l_k, i_{3k} i_1}^* \right]$$

corresponding to a graph with the shape



gives  $\approx \text{tr}(T_1 T_1^*) \text{tr}(T_1 T_1^* T_2 T_2^*)$  if green  $\hat{=}$  1 and red  $\hat{=}$  2.

- Expectation factor  $\mathbb{E}[Y_{l_1, i_2 j_1} \cdots Y_{l_k, j_k i_{3k}}^*]$  is 1.
- $1/p^2 \text{tr}(T_1 T_1^*) \text{tr}(T_1 T_1^* T_2 T_2^*)$  converges to  $M_{(1)}^1 M_{(1,2)}^2$ .

**Possible directions for future research:**

## Possible directions for future research:

- Does the theorem hold without the restriction  $\|T_l\| \leq \tau_0$  for all  $l, p$ ?

### Possible directions for future research:

- Does the theorem hold without the restriction  $\|T_l\| \leq \tau_0$  for all  $l, p$ ?
- Is there a relation between the limiting spectral distribution of  $[X]_p^n$  and the limiting spectral distribution of the true covariance matrix

$$(t_1 - t_0) T_1 T_1^* + \cdots + (t_m - t_{m-1}) T_m T_m^*?$$

## Possible directions for future research:

- Does the theorem hold without the restriction  $\|T_l\| \leq \tau_0$  for all  $l, p$ ?
- Is there a relation between the limiting spectral distribution of  $[X]_p^n$  and the limiting spectral distribution of the true covariance matrix

$$(t_1 - t_0) T_1 T_1^* + \cdots + (t_m - t_{m-1}) T_m T_m^*?$$

- For  $T_1 = \dots = T_m$  such a relation exists (Marčenko-Pastur equation), allowing the construction of consistent spectrum estimators (N. El Karoui, 2007)



# Bibliography

Thank you for your attention!



N. El Karoui (2008): Spectrum estimation for large dimensional covariance matrices using random matrix theory. *Annals of Statistics*, 36, 2757–2790.



C. Heinrich and M. Podolskij (2014, Work in progress): On spectral distribution of high dimensional covariation matrices.



Y.Q. Yin and P.R. Krishnaiah (1983): A limit theorem for the eigenvalues of product of two random matrices. *Journal of Multivariate Analysis* 13, 489—507.