Matrix concentration inequalities with dependent summands and sharp leading-order terms

Alexander Van Werde, presented at SNAPP Seminar (2023)

Joint work with Jaron Sanders

arXiv:2307.11632

TU/e EINDHOVEN UNIVERSITY OF TECHNOLOGY

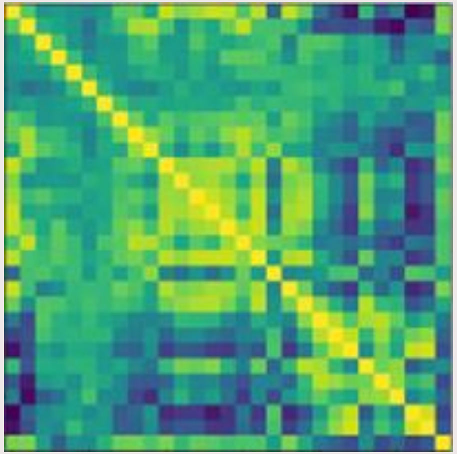
Applications of random matrices in dependent processes

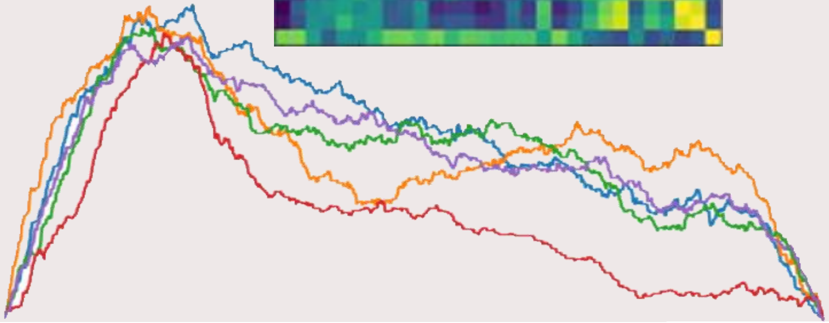
(Auto)covariance matrices of time series

- M. Banna, F. Merlevède, and P. Youssef (2016)
- F. Han, Y. Li. (2020)

Spectral clustering in block Markov chains & reinforcement learning

- J. Sanders, A. Proutière, S.-Y. Yun (2020)
- A. Senen-Cerda, J. Sanders (2023)
- Y. Jedra, J. Lee, A. Proutiere, S.-Y. Yun (2023)





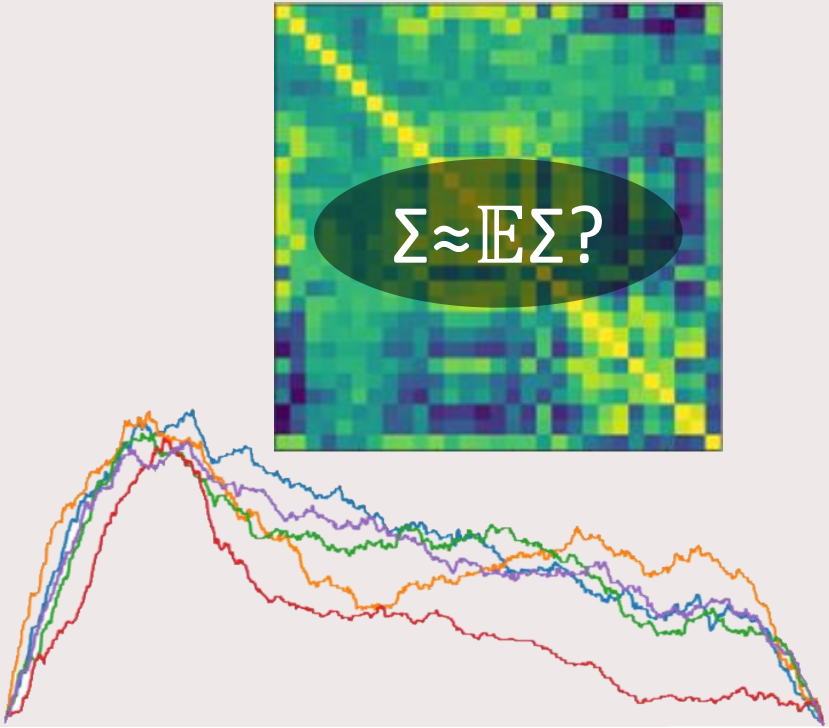
Applications of random matrices in dependent processes

(Auto)covariance matrices of time series

- M. Banna, F. Merlevède, and P. Youssef (2016)
- F. Han, Y. Li. (2020)

Spectral clustering in block Markov chains & reinforcement learning

- J. Sanders, A. Proutière, S.-Y. Yun (2020)
- A. Senen-Cerda, J. Sanders (2023)
- Y. Jedra, J. Lee, A. Proutiere, S.-Y. Yun (2023)



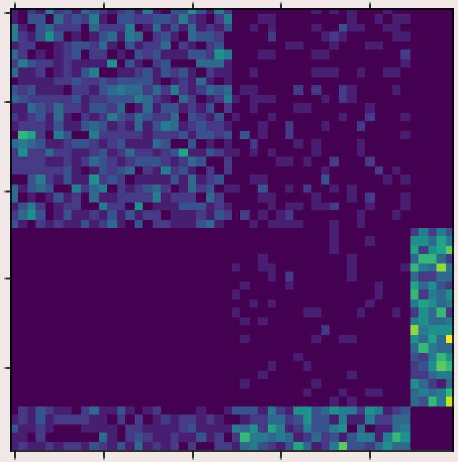
Applications of random matrices in dependent processes

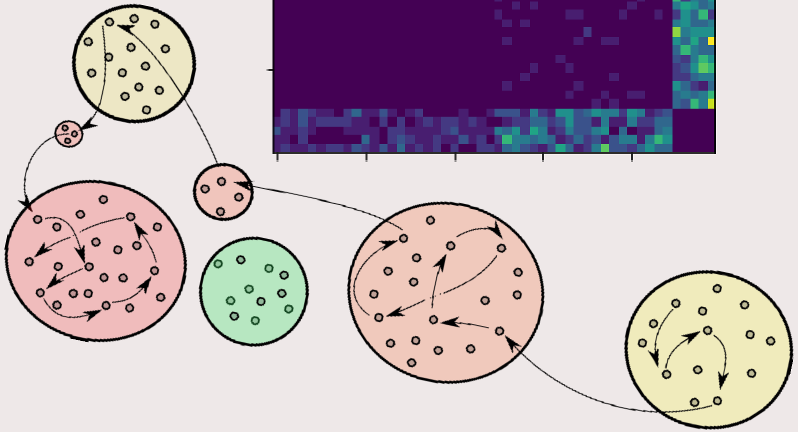
(Auto)covariance matrices of time series

- M. Banna, F. Merlevède, and P. Youssef (2016)
- F. Han, Y. Li. (2020)

Spectral clustering in block Markov chains & reinforcement learning

- J. Sanders, A. Proutière, S.-Y. Yun (2020)
- A. Senen-Cerda, J. Sanders (2023)
- Y. Jedra, J. Lee, A. Proutiere, S.-Y. Yun (2023)





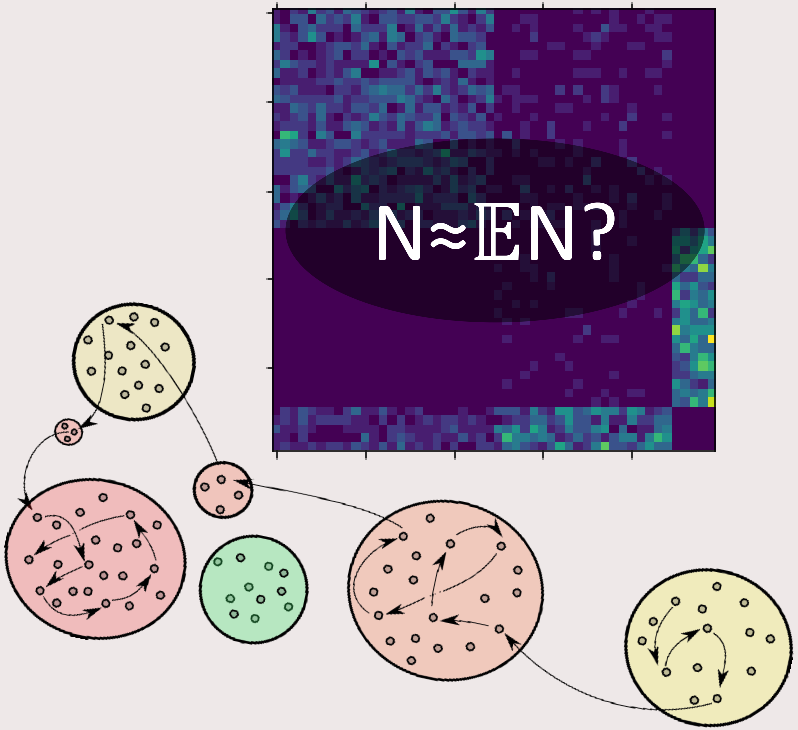
Applications of random matrices in dependent processes

(Auto)covariance matrices of time series

- M. Banna, F. Merlevède, and P. Youssef (2016)
- F. Han, Y. Li. (2020)

Spectral clustering in block Markov chains & reinforcement learning

- J. Sanders, A. Proutière, S.-Y. Yun (2020)
- A. Senen-Cerda, J. Sanders (2023)
- Y. Jedra, J. Lee, A. Proutiere, S.-Y. Yun (2023)



Applications of random matrices in dependent processes

(Auto)covariance matrices of time series
- M. Banna, F. Merlevède, and P. Youssef (2016)
- F. Han, Y. Li. (2020)

Spectral clustering in block Markov chains & reinforcement learning
- J. Sanders, A. Proutière, S.-Y. Yun (2020)
- A. Senen-Cerda, J. Sanders (2023)
- Y. Jedra, J. Lee, A. Proutiere, S.-Y. Yun (2023)

Convergence rates for Hessians in stochastic gradient descent
- J. Neeman, B. Shi, R. Ward (2023)

Randomness-efficient sampling methods
- A. Garg, Y.T. Lee, Z. Song, N. Srivastava (2018)

Random walk-based graph embedding methods
- J. Qiu, C. Wang, B. Liao, R. Peng, J. Tang (2020)

TU/e

Summary of our contributions

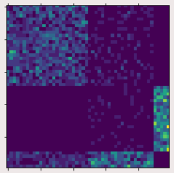
1. Non-asymptotic concentration inequality

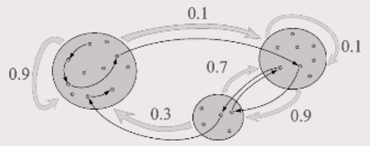
$$\|\boldsymbol{M} - \mathbb{E}\boldsymbol{M}\| \leq \|\boldsymbol{M}_{\mathrm{free}}\| + \varepsilon$$

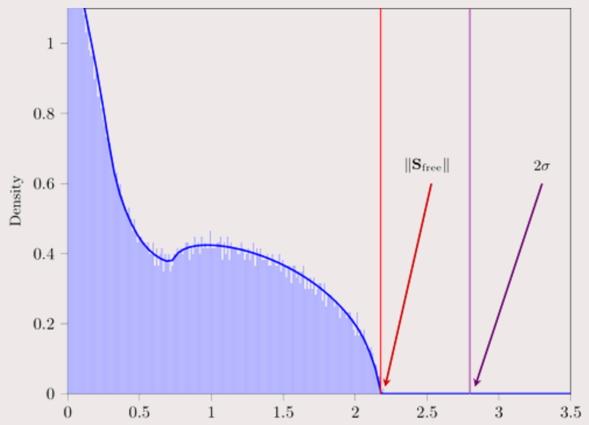
   for matrices $\boldsymbol{M}$ generated by a dependent process.

2. Asymptotically sharp and easy to use in applications!
   - Block Markov chains
   - Random matrices with heavy-tailed entries
   - Random graphs with dependent edges

3. Proof techniques to deal with dependence.
   (Boolean cumulants, change-of-measure, combinatorial estimates, …)







TU/e

Summary of our contributions

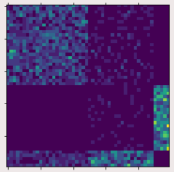
1. Non-asymptotic concentration inequality

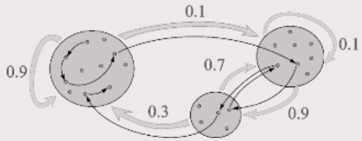
$$\|\boldsymbol{M} - \mathbb{E}\boldsymbol{M}\| \leq \|\boldsymbol{M}_{\text{free}}\| + \varepsilon$$

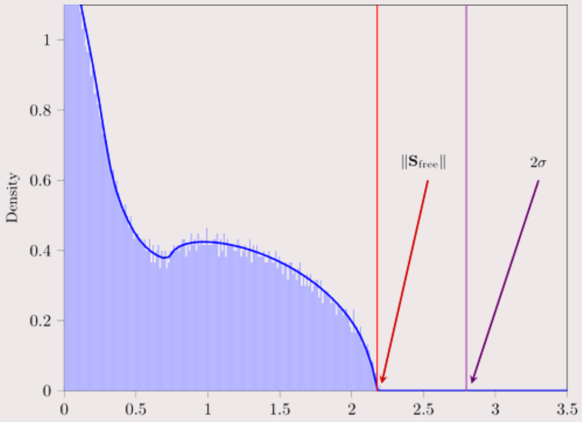
for matrices $\boldsymbol{M}$ generated by a dependent process.

2. Asymptotically sharp and easy to use in applications!
    - Block Markov chains
    - Random matrices with heavy-tailed entries
    - Random graphs with dependent edges

3. Proof techniques to deal with dependence. (Boolean cumulants, change-of-measure, combinatorial estimates, …)







Summary of our contributions

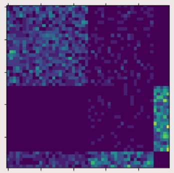
1. Non-asymptotic concentration inequality

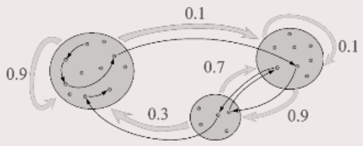
$$\|\boldsymbol{M} - \mathbb{E}\boldsymbol{M}\| \leq \|\boldsymbol{M}_{\text{free}}\| + \varepsilon$$

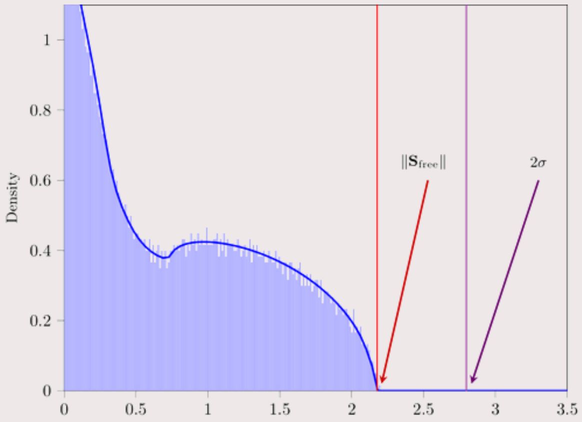
for matrices $\boldsymbol{M}$ generated by a dependent process.

2. Asymptotically sharp and easy to use in applications!
    - Block Markov chains
    - Random matrices with heavy-tailed entries
    - Random graphs with dependent edges

3. Proof techniques to deal with dependence.
(Boolean cumulants, change-of-measure, combinatorial estimates, …)







Context: suboptimality of classical matrix concentration

Matrix Khintchine (Lust–Piquard)

There exist absolute constants $c, C > 0$ so that any self-adjoint Gaussian $d \times d$ matrix $\boldsymbol{M}$ satisfies

$$c\sigma \leq \mathbb{E}\|\boldsymbol{M} - \mathbb{E}\boldsymbol{M}\| \leq C\sqrt{\ln(2d)}\,\sigma$$

where $\sigma^2 := \|\,\mathbb{E}(\boldsymbol{M} - \mathbb{E}\boldsymbol{M})^2\|$.

Context: suboptimality of classical matrix concentration

Matrix Khintchine (Lust–Piquard)

There exist absolute constants $c, C > 0$ so that any self-adjoint Gaussian $d \times d$ matrix $\boldsymbol{M}$ satisfies

$$c\sigma \leq \mathbb{E}\|\boldsymbol{M} - \mathbb{E}\boldsymbol{M}\| \leq C\sqrt{\ln(2d)}\,\sigma$$

where $\sigma^2 := \|\mathbb{E}(\boldsymbol{M} - \mathbb{E}\boldsymbol{M})^2\|$.

Context: suboptimality of classical matrix concentration

Matrix Khintchine (Lust–Piquard)

There exist absolute constants $c, C > 0$ so that any self-adjoint Gaussian $d \times d$ matrix $\boldsymbol{M}$ satisfies

$$c\sigma \leq \mathbb{E}\|\boldsymbol{M} - \mathbb{E}\boldsymbol{M}\| \leq C\sqrt{\ln(2d)}\,\sigma$$

where $\sigma^2 := \|\,\mathbb{E}(\boldsymbol{M} - \mathbb{E}\boldsymbol{M})^2\|$.

Context: suboptimality of classical matrix concentration

Matrix Khintchine (Lust–Piquard)

There exist absolute constants $c, C > 0$ so that any self-adjoint Gaussian $d \times d$ matrix $\boldsymbol{M}$ satisfies

$$c\sigma \leq \mathbb{E}\|\boldsymbol{M} - \mathbb{E}\boldsymbol{M}\| \leq C\sqrt{\ln(2d)}\,\sigma$$

where $\sigma^2 := \|\mathbb{E}(\boldsymbol{M} - \mathbb{E}\boldsymbol{M})^2\|$.

Limitations:

- Restricted to Gaussian matrices
- Upper bound has suboptimal asymptotic order in typical applications

Context: suboptimality of classical matrix concentration

Matrix Khintchine (Lust–Piquard)

There exist absolute constants $c, C > 0$ so that any self-adjoint Gaussian $d \times d$ matrix $\boldsymbol{M}$ satisfies

$$c\sigma \leq \mathbb{E}\|\boldsymbol{M} - \mathbb{E}\boldsymbol{M}\| \leq C\,\sqrt{\ln(2d)}\,\sigma$$

where $\sigma^2 := \|\,\mathbb{E}(\boldsymbol{M} - \mathbb{E}\boldsymbol{M})^2\|$.

Limitations:

- Restricted to Gaussian matrices
- Upper bound has suboptimal asymptotic order in typical applications

Context: suboptimality of classical matrix concentration

Matrix Khintchine (Lust–Piquard)

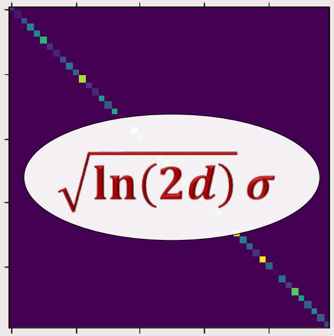
There exist absolute constants $c, C > 0$ so that any self-adjoint Gaussian $d \times d$ matrix $\boldsymbol{M}$ satisfies

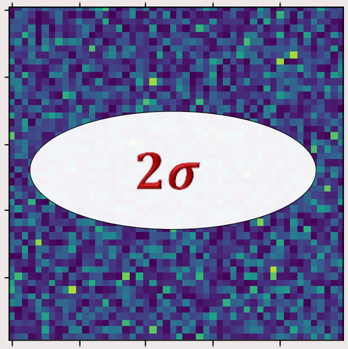
$$c\sigma \leq \mathbb{E}\|\boldsymbol{M} - \mathbb{E}\boldsymbol{M}\| \leq C\sqrt{\ln(2d)}\,\sigma$$

where $\sigma^2 := \|\mathbb{E}(\boldsymbol{M} - \mathbb{E}\boldsymbol{M})^2\|$.

Limitations:

- Restricted to Gaussian matrices
- Upper bound has suboptimal asymptotic order in typical applications





Some related literature

Matrix concentration with non-Gaussian dependent summands

- **R.I. Oliveira (2009):**
  Matrix Bernstein for matrix martingales
- **D. Paulin, L. Mackey, J.A. Tropp (2016)**
  Efron-Stein for exchangeable pairs
- **M. Banna, F. Merlevède, and P. Youssef (2016):**
  β-mixing analogue of matrix Bernstein.
- **E. Bacry, S. Gaïffas, J.F. Muzy (2018)**
  Continuous-time martingales
- **F. Han, Y. Li (2020)**
  Autocovariance matrices
- **J. Neeman, B. Shi, and R. Ward (2023):**
  Markovian matrix Bernstein.

When can $\sqrt{\ln(2d)}$ be removed?

- **J.A. Tropp (2018):**
  Introduced a difficult-to-compute but insightful parameter.
- **A.S. Bandeira, M.T. Boedihardjo, and R. van Handel (2021):**
  Breakthrough for Gaussian matrices: $v^2 := \|\mathrm{Cov}(\boldsymbol{M})\|$ and relation to free probability.
- **T. Brailovskaya and R. van Handel (2022):**
  Independent non-Gaussian summands.

Some related literature

Matrix concentration with non-Gaussian dependent summands

- **R.I. Oliveira (2009):**
  Matrix Bernstein for matrix martingales
- **D. Paulin, L. Mackey, J.A. Tropp (2016)**
  Efron-Stein for exchangeable pairs
- **M. Banna, F. Merlevède, and P. Youssef (2016):**
  β-mixing analogue of matrix Bernstein.
- **E. Bacry, S. Gaïffas, J.F. Muzy (2018)**
  Continuous-time martingales
- **F. Han, Y. Li (2020)**
  Autocovariance matrices
- **J. Neeman, B. Shi, and R. Ward (2023):**
  Markovian matrix Bernstein.

When can $\sqrt{\ln(2d)}$ be removed?

- **J.A. Tropp (2018):**
  Introduced a difficult-to-compute but insightful parameter.
- **A.S. Bandeira, M.T. Boedihardjo, and R. van Handel (2021):**
  Breakthrough for Gaussian matrices: $v^2 := \|\mathrm{Cov}(\boldsymbol{M})\|$ and relation to free probability.
- **T. Brailovskaya and R. van Handel (2022):**
  Independent non-Gaussian summands.

Some related literature

Matrix concentration with non-Gaussian dependent summands

- **R.I. Oliveira (2009):**
  Matrix Bernstein for matrix martingales
- **D. Paulin, L. Mackey, J.A. Tropp (2016)**
  Efron-Stein for exchangeable pairs
- **M. Banna, F. Merlevède, and P. Youssef (2016):**
  β-mixing analogue of matrix Bernstein.
- **E. Bacry, S. Gaïffas, J.F. Muzy (2018)**
  Continuous-time martingales
- **F. Han, Y. Li (2020)**
  Autocovariance matrices
- **J. Neeman, B. Shi, and R. Ward (2023):**
  Markovian matrix Bernstein.

When can $\sqrt{\ln(2d)}$ be removed?

- **J.A. Tropp (2018):**
  Introduced a difficult-to-compute but insightful parameter.
- **A.S. Bandeira, M.T. Boedihardjo, and R. van Handel (2021):**
  Breakthrough for Gaussian matrices: $v^2 := \|\mathrm{Cov}(\boldsymbol{M})\|$ and relation to free probability.
- **T. Brailovskaya and R. van Handel (2022):**
  Independent non-Gaussian summands.

Some related literature

Matrix concentration with non-Gaussian dependent summands

- **R.I. Oliveira (2009):**
  Matrix Bernstein for matrix martingales
- **D. Paulin, L. Mackey, J.A. Tropp (2016)**
  Efron-Stein for exchangeable pairs
- **M. Banna, F. Merlevède, and P. Youssef (2016):**
  β-mixing analogue of matrix Bernstein.
- **E. Bacry, S. Gaïffas, J.F. Muzy (2018)**
  Continuous-time martingales
- **F. Han, Y. Li (2020)**
  Autocovariance matrices
- **J. Neeman, B. Shi, and R. Ward (2023):**
  Markovian matrix Bernstein.

When can $\sqrt{\ln(2d)}$ be removed?

- **J.A. Tropp (2018):**
  Introduced a difficult-to-compute but insightful parameter.
- **A.S. Bandeira, M.T. Boedihardjo, and R. van Handel (2021):**
  Breakthrough for Gaussian matrices: $v^2 := \|\mathrm{Cov}(\boldsymbol{M})\|$ and relation to free probability.
- **T. Brailovskaya and R. van Handel (2022):**
  Independent non-Gaussian summands.

Some related literature

Matrix concentration with non-Gaussian dependent summands

- **R.I. Oliveira (2009):**
  Matrix Bernstein for matrix martingales
- **D. Paulin, L. Mackey, J.A. Tropp (2016)**
  Efron-Stein for exchangeable pairs
- **M. Banna, F. Merlevède, and P. Youssef (2016):**
  β-mixing analogue of matrix Bernstein.
- **E. Bacry, S. Gaïffas, J.F. Muzy (2018)**
  Continuous-time martingales
- **F. Han, Y. Li (2020)**
  Autocovariance matrices
- **J. Neeman, B. Shi, and R. Ward (2023):**
  Markovian matrix Bernstein.

When can $\sqrt{\ln(2d)}$ be removed?

- **J.A. Tropp (2018):**
  Introduced a difficult-to-compute but insightful parameter.
- **A.S. Bandeira, M.T. Boedihardjo, and R. van Handel (2021):**
  Breakthrough for Gaussian matrices: $v^2 := \|\mathrm{Cov}(\boldsymbol{M})\|$ and relation to free probability.
- **T. Brailovskaya and R. van Handel (2022):**
  Independent non-Gaussian summands.

Some related literature

Matrix concentration with non-Gaussian dependent summands

- **R.I. Oliveira (2009):**
  Matrix Bernstein for matrix martingales
- **D. Paulin, L. Mackey, J.A. Tropp (2016)**
  Efron-Stein for exchangeable pairs
- **M. Banna, F. Merlevède, and P. Youssef (2016):**
  β-mixing analogue of matrix Bernstein.
- **E. Bacry, S. Gaïffas, J.F. Muzy (2018)**
  Continuous-time martingales
- **F. Han, Y. Li (2020)**
  Autocovariance matrices
- **J. Neeman, B. Shi, and R. Ward (2023):**
  Markovian matrix Bernstein.

When can $\sqrt{\ln(2d)}$ be removed?

- **J.A. Tropp (2018):**
  Introduced a difficult-to-compute but insightful parameter.
- **A.S. Bandeira, M.T. Boedihardjo, and R. van Handel (2021):**
  Breakthrough for Gaussian matrices: $v^2 \coloneqq \|\mathrm{Cov}(\boldsymbol{M})\|$ and relation to free probability.
- **T. Brailovskaya and R. van Handel (2022):**
  Independent non-Gaussian summands.

Sharp matrix concentration
for Markovian model & matrix series model

Markovian model and parameters

Markovian model.

Consider a Markov chain $Z_1, \ldots, Z_n$.

Generate self-adjoint $d \times d$ matrices

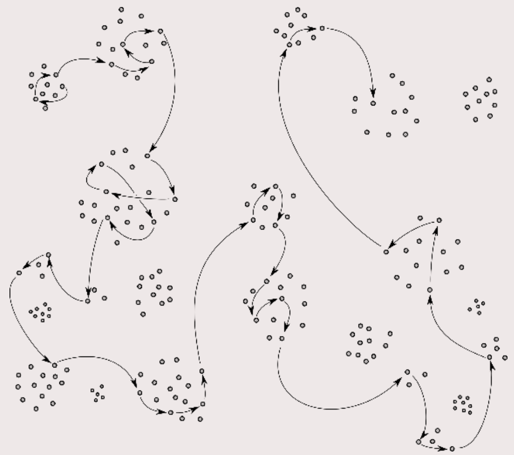
$$\boldsymbol{X}_i := f_i(Z_i) \quad \text{and} \quad \boldsymbol{M} := \sum_{i=1}^{n} \boldsymbol{X}_i$$

Further, for simplicity, also assume

- Finite state space
- Transition matrix P
  & equilibrium distribution π
- $\mathbb{E}\boldsymbol{X}_i = 0$

Dependence parameter

$$\Psi := \min\{\, t \geq 1 : \frac{|P_{x,y}^t - \pi_y|}{\pi_y} \leq \frac{1}{4} \; \forall x, y\}$$



Markovian model and parameters

Markovian model.

Consider a Markov chain $Z_1, \dots, Z_n$.

Generate self-adjoint $d \times d$ matrices

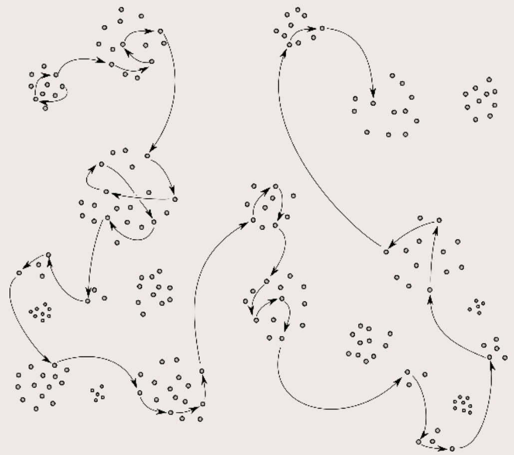
$$\boldsymbol{X}_i \coloneqq f_i(Z_i) \quad \text{and} \quad \boldsymbol{M} \coloneqq \sum_{i=1}^{n} \boldsymbol{X}_i$$

Further, for simplicity, also assume

- Finite state space
- Transition matrix P & equilibrium distribution π
- $\mathbb{E}\boldsymbol{X}_i = 0$

Dependence parameter

$$\Psi \coloneqq \min\{\, t \geq 1 : \frac{|P_{x,y}^t - \pi_y|}{\pi_y} \leq \frac{1}{4} \; \forall x, y\}$$



Markovian model and parameters

Markovian model.

Consider a Markov chain $Z_1, \dots, Z_n$.

Generate self-adjoint $d \times d$ matrices

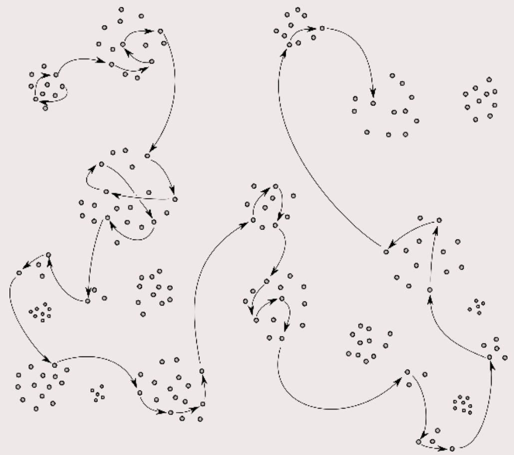
$$\boldsymbol{X}_i := f_i(Z_i) \quad \text{and} \quad \boldsymbol{M} := \sum_{i=1}^{n} \boldsymbol{X}_i$$

Further, for simplicity, also assume

- Finite state space
- Transition matrix P
  & equilibrium distribution π
- $\mathbb{E}\boldsymbol{X}_i = 0$

Dependence parameter

$$\Psi := \min\{\, t \geq 1 : \frac{|P_{x,y}^t - \pi_y|}{\pi_y} \leq \frac{1}{4} \ \forall x, y\}$$



Markovian model and parameters

Markovian model.

Consider a Markov chain $Z_1, \ldots, Z_n$.

Generate self-adjoint $d \times d$ matrices

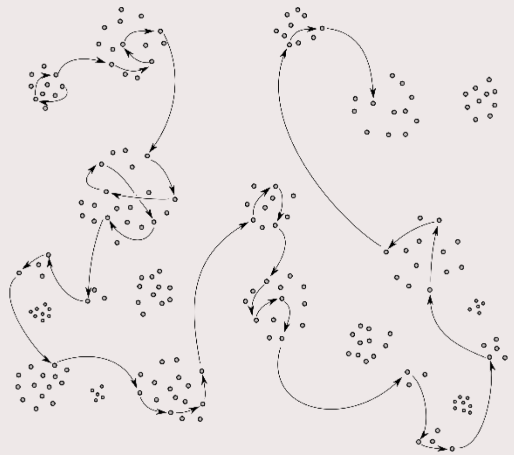
$\boldsymbol{X}_i := f_i(Z_i)$ and $\boldsymbol{M} := \sum_{i=1}^{n} \boldsymbol{X}_i$

Further, for simplicity, also assume

- Finite state space
- Transition matrix P
  & equilibrium distribution π
- $\mathbb{E}\boldsymbol{X}_i = 0$

Dependence parameter

$$\Psi := \min\{ t \geq 1 : \frac{|P_{x,y}^{t} - \pi_y|}{\pi_y} \leq \frac{1}{4} \ \forall x, y\}$$



Markovian model and parameters

Markovian model.

Consider a Markov chain $Z_1, \dots, Z_n$.

Generate self-adjoint $d \times d$ matrices

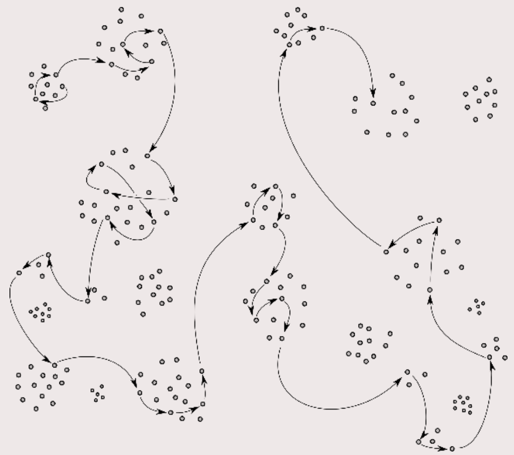
$$\boldsymbol{X}_i := f_i(Z_i) \quad \text{and} \quad \boldsymbol{M} := \textstyle\sum_{i=1}^{n} \boldsymbol{X}_i$$

Further, for simplicity, also assume

- Finite state space
- Transition matrix P
  & equilibrium distribution π
- $\mathbb{E}\boldsymbol{X}_i = 0$

Dependence parameter

$$\Psi := \min\{\, t \geq 1 : \frac{|P_{x,y}^{t} - \pi_y|}{\pi_y} \leq \frac{1}{4} \;\forall x, y\}$$



Markovian model and parameters

Markovian model.

Consider a Markov chain $Z_1, \dots, Z_n$.

Generate self-adjoint $d \times d$ matrices

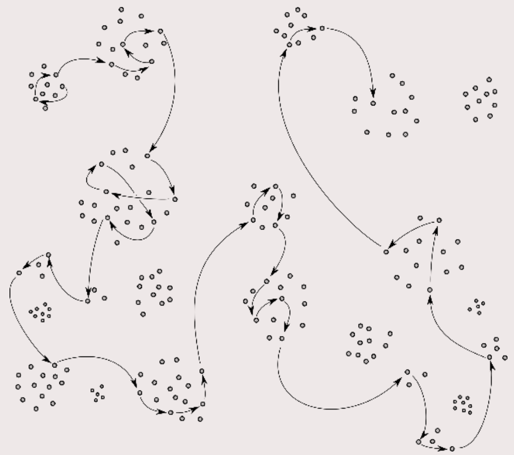
$$\boldsymbol{X}_i \coloneqq f_i(Z_i) \quad \text{and} \quad \boldsymbol{M} \coloneqq \sum_{i=1}^{n} \boldsymbol{X}_i$$

Further, for simplicity, also assume

- Finite state space
- Transition matrix P
  & equilibrium distribution π
- $\mathbb{E}\boldsymbol{X}_i = 0$

Dependence parameter

$$\Psi \coloneqq \min\{\, t \geq 1 : \frac{|P_{x,y}^{t} - \pi_y|}{\pi_y} \leq \frac{1}{4} \ \forall x, y\}$$



Markovian model and parameters

Markovian model.

Consider a Markov chain $Z_1, \dots, Z_n$.

Generate self-adjoint $d \times d$ matrices

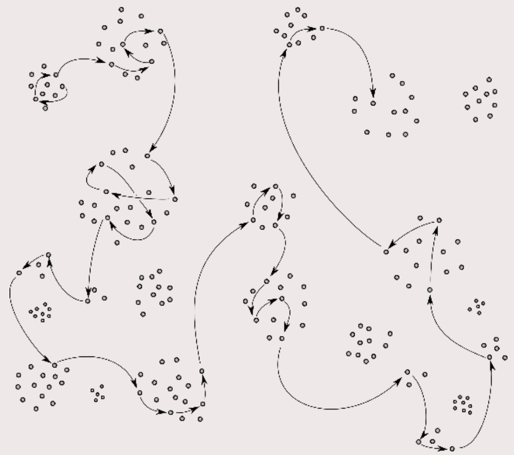
$$\boldsymbol{X}_i := f_i(Z_i) \quad \text{and} \quad \boldsymbol{M} := \sum_{i=1}^{n} \boldsymbol{X}_i$$

Further, for simplicity, also assume

- Finite state space
- Transition matrix P
  & equilibrium distribution π
- $\mathbb{E}\boldsymbol{X}_i = 0$

Dependence parameter

$$\Psi := \min\{\, t \geq 1 : \frac{|P_{x,y}^t - \pi_y|}{\pi_y} \leq \frac{1}{4} \; \forall x, y\}$$



Sharp matrix concentration in Markovian model

Main result.

There exists $c > 0$ such that for $0 < \delta \leq 1$

$$\mathbb{E}\|\boldsymbol{M}\| \leq (1+\delta)\|\boldsymbol{M}_{\text{free}}\| + c\mathcal{E}_{d,\delta}$$

where

$$\mathcal{E}_{d,\delta} := v^{1/2}\sigma^{1/2}\log_{1+\delta}(2d)^{3/4} + R^{1/3}\Psi^{2/3}\varsigma^{2/3}\log_{1+\delta}(2d)^{2/3} + R\,\Psi\log_{1+\delta}(2d)$$

Dependence parameter

$$\Psi := \min\{\, t \geq 1 : \frac{|P^t_{x,y} - \pi_y|}{\pi_y} \leq \frac{1}{4}\ \forall x, y\}$$

Free-probabilistic quantity

$$\|\boldsymbol{M}_{\text{free}}\| = \inf_{\boldsymbol{W} \succ \boldsymbol{0}} \lambda_{\max}\left(\boldsymbol{W}^{-1} + \mathbb{E}[\boldsymbol{M}\boldsymbol{W}\boldsymbol{M}]\right).$$

Matrix parameters

$$\varsigma^2 := \left\|\mathbb{E}\left[\textstyle\sum_{i=1}^n \boldsymbol{X}_i^2\right]\right\|, \quad \sigma^2 := \|\mathbb{E}[\boldsymbol{M}^2]\|$$

$$v^2 := \|\mathrm{Cov}(\boldsymbol{M})\|, \qquad R \geq \|\boldsymbol{X}_i\|.$$

Sharp matrix concentration in Markovian model

Main result.

There exists $c > 0$ such that for $0 < \delta \leq 1$

$$\mathbb{E}\|\boldsymbol{M}\| \leq (1+\delta)\|\boldsymbol{M}_{\text{free}}\| + c\mathcal{E}_{d,\delta}$$

where

$$\mathcal{E}_{d,\delta} := v^{1/2}\sigma^{1/2}\log_{1+\delta}(2d)^{3/4}$$
$$+R^{1/3}\Psi^{2/3}\varsigma^{2/3}\log_{1+\delta}(2d)^{2/3}$$
$$+\ R\,\Psi\log_{1+\delta}(2d)$$

Dependence parameter

$$\Psi := \min\{\, t \geq 1 : \frac{|P^t_{x,y} - \pi_y|}{\pi_y} \leq \frac{1}{4}\ \forall x, y\}$$

Free-probabilistic quantity (Lehner)

$$\|\boldsymbol{M}_{\text{free}}\| = \inf_{\boldsymbol{W} \succ \boldsymbol{0}} \lambda_{\max}\left(\boldsymbol{W}^{-1} + \mathbb{E}[\boldsymbol{M}\boldsymbol{W}\boldsymbol{M}]\right).$$

Matrix parameters

$$\varsigma^2 := \left\|\mathbb{E}\left[\textstyle\sum_{i=1}^n \boldsymbol{X}_i^2\right]\right\|, \quad \sigma^2 := \|\mathbb{E}[\boldsymbol{M}^2]\|$$

$$v^2 := \|\mathrm{Cov}(\boldsymbol{M})\|, \qquad R \geq \|\boldsymbol{X}_i\|.$$

TU/e

Sharp matrix concentration in Markovian model

Main result.

There exists $c > 0$ such that for $0 < \delta \leq 1$

$$\mathbb{E}\|\boldsymbol{M}\| \leq (1+\delta)\|\boldsymbol{M}_{\text{free}}\| + c\mathcal{E}_{d,\delta}$$

where

$$\mathcal{E}_{d,\delta} := v^{1/2}\sigma^{1/2}\log_{1+\delta}(2d)^{3/4}$$
$$+R^{1/3}\Psi^{2/3}\varsigma^{2/3}\log_{1+\delta}(2d)^{2/3}$$
$$+\ R\,\Psi\log_{1+\delta}(2d)$$

Dependence parameter

$$\Psi := \min\{\, t \geq 1 : \frac{|P^t_{x,y} - \pi_y|}{\pi_y} \leq \frac{1}{4}\ \forall x, y\}$$

Free-probabilistic quantity (Lehner)

$$\|\boldsymbol{M}_{\text{free}}\| = \inf_{\boldsymbol{W} \succ \boldsymbol{0}} \lambda_{\max}\left(\boldsymbol{W}^{-1} + \mathbb{E}[\boldsymbol{M}\boldsymbol{W}\boldsymbol{M}]\right).$$

Matrix parameters

$$\varsigma^2 := \left\|\mathbb{E}\left[\sum_{i=1}^n \boldsymbol{X}_i^2\right]\right\|, \quad \sigma^2 := \|\mathbb{E}[\boldsymbol{M}^2]\|$$

$$v^2 := \|\mathrm{Cov}(\boldsymbol{M})\|, \qquad R \geq \|\boldsymbol{X}_i\|.$$

Informal summary

Main result. (Informally)

We provide explicit $\varepsilon > 0$ such that

$$\mathbb{E}\|\boldsymbol{M}\| \leq \|\boldsymbol{M}_{\text{free}}\| + \varepsilon \ln(2d)$$

Informal summary

Main result. (Informally)

We provide explicit $\varepsilon > 0$ such that

$$\mathbb{E}\|\boldsymbol{M}\| \leq \|\boldsymbol{M}_{\text{free}}\| + \varepsilon \ln(2d)$$

Removal of $\sqrt{\ln(2d)}$

Due to Pisier,

$$\sigma \leq \|\boldsymbol{M}_{\text{free}}\| \leq 2\sigma$$

Asymptotically sharp

In many relevant models

$$\lim_{d\to\infty} \|\boldsymbol{M}\| = \|\boldsymbol{M}_{\text{free}}\|$$

TU/e

Informal summary

Main result. (Informally)

We provide explicit $\varepsilon > 0$ such that

$$\mathbb{E}\|\boldsymbol{M}\| \leq \|\boldsymbol{M}_{\text{free}}\| + \varepsilon \ln(2d)$$

Removal of $\sqrt{\ln(2d)}$

Due to Pisier,

$$\sigma \leq \|\boldsymbol{M}_{\text{free}}\| \leq 2\sigma$$

Asymptotically sharp

In many relevant models

$$\lim_{d\to\infty} \|\boldsymbol{M}\| = \|\boldsymbol{M}_{\text{free}}\|$$

Informal summary

Main result. (Informally)

We provide explicit $\varepsilon > 0$ such that

$$\mathbb{E}\|\boldsymbol{M}\| \leq \|\boldsymbol{M}_{\text{free}}\| + \varepsilon \ln(2d)$$

Removal of $\sqrt{\ln(2d)}$

Due to Pisier,

$$\sigma \leq \|\boldsymbol{M}_{\text{free}}\| \leq 2\sigma$$

Asymptotically sharp

In many relevant models

$$\lim_{d\to\infty} \|\boldsymbol{M}\| = \|\boldsymbol{M}_{\text{free}}\|$$

Matrix series model.

We show a similar result for a *matrix series model*.

There, summands are of the form $\boldsymbol{X}_i := Y_i \boldsymbol{A}_i$ for deterministic $\boldsymbol{A}_i$.

Proof sketch

Proof sketch.

Goal. $\mathbb{E}[\operatorname{tr} \boldsymbol{M}^{2p}] \approx \mathbb{E}[\operatorname{tr} \mathbf{G}^{2p}]$ where $\boldsymbol{G}$ is a *Gaussian model* for $\boldsymbol{M}$

Interpolate: For $t \in [0,1]$ set $\boldsymbol{M}(t) := \sqrt{t}\boldsymbol{M} + \sqrt{1-t}\boldsymbol{G}$ and show that

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}[\operatorname{tr} \boldsymbol{M}(t)^{2p}] \leq \epsilon(t)$$

Proof sketch.

Goal. $\mathbb{E}[\operatorname{tr} \boldsymbol{M}^{2p}] \approx \mathbb{E}[\operatorname{tr} \mathbf{G}^{2p}]$ where $\boldsymbol{G}$ is a *Gaussian model* for $\boldsymbol{M}$

Interpolate: For $t \in [0,1]$ set $\boldsymbol{M}(t) \coloneqq \sqrt{t}\boldsymbol{M} + \sqrt{1-t}\boldsymbol{G}$ and show that

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}[\operatorname{tr} \boldsymbol{M}(t)^{2p}] \leq \epsilon(t)$$

Brailovskaya & van Handel (independent setting)

- Expansion with *classical cumulants.*
- Independence-implies-vanishing property reduces combinatorics

We lack independence!

Proof sketch.

Goal. $\mathbb{E}[\mathrm{tr}\, \boldsymbol{M}^{2p}] \approx \mathbb{E}[\mathrm{tr}\, \mathbf{G}^{2p}]$ where $\boldsymbol{G}$ is a *Gaussian model* for $\boldsymbol{M}$

Interpolate: For $t \in [0,1]$ set $\boldsymbol{M}(t) := \sqrt{t}\boldsymbol{M} + \sqrt{1-t}\boldsymbol{G}$ and show that

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}[\,\mathrm{tr}\, \boldsymbol{M}(t)^{2p}] \leq \epsilon(t)$$

Brailovskaya & van Handel (independent setting)

- Expansion with *classical cumulants.*
- Independence-implies-vanishing property reduces combinatorics

We lack independence!

Direct modification using classical cumulants gives suboptimal results…

Summary

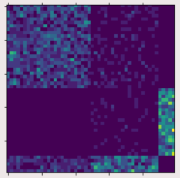
arXiv:2307.11632

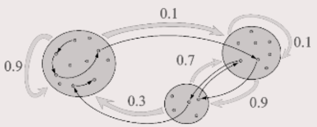
1. We achieve sharp concentration inequalities in a dependent setting

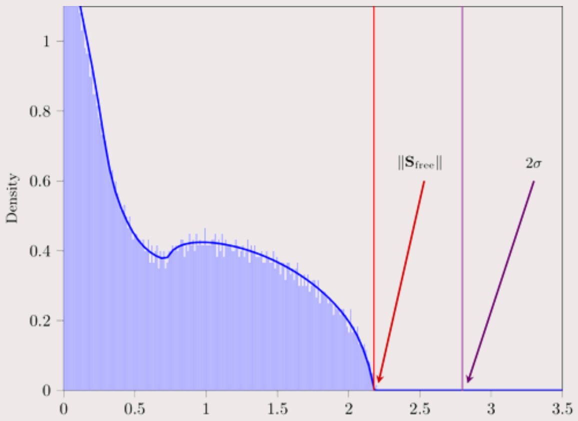
$$\|\boldsymbol{M} - \mathbb{E}\boldsymbol{M}\| \leq \|\boldsymbol{M}_{\text{free}}\| + \varepsilon$$

2. Surprisingly, classical cumulants are insufficient. We use the (more obscure) Boolean cumulants instead.

3. Applications such as clustering in Markov chains and random graphs may be found in the paper.







Thank you!

arXiv:2307.11632

1. We achieve sharp concentration inequalities in a dependent setting

$$\|\boldsymbol{M} - \mathbb{E}\boldsymbol{M}\| \leq \|\boldsymbol{M}_{\text{free}}\| + \varepsilon$$

2. Surprisingly, classical cumulants are insufficient. We use the (more obscure) Boolean cumulants instead.

3. Applications such as clustering in Markov chains and random graphs may be found in the paper.