An extension of Replica-Exchange Monte Carlo methods applying to matrix geometry

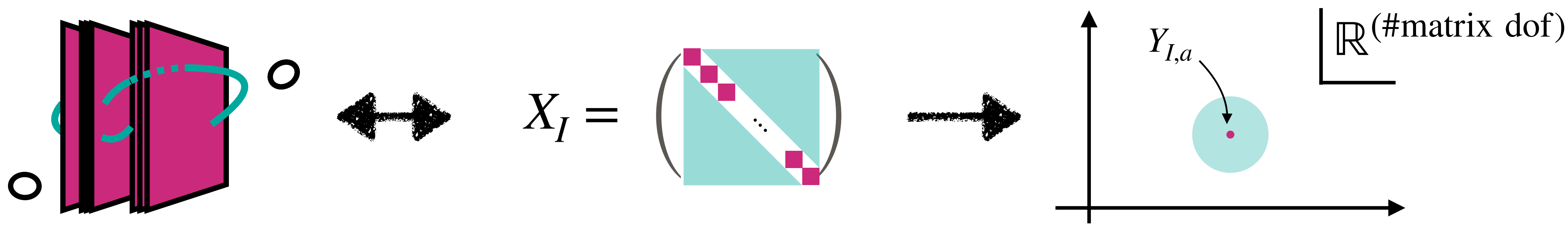
Hiromasa Watanabe（渡辺 展正)
Yukawa Institute for Theoretical Physics, Kyoto U.

Based on collaboration with
M. Hanada (Queens Mary, London), S. Kanno (Tsukuba), S. Matsuura (Keio)
in progress

2023/10/05
ExU-YITP Conference "Quantum Information and Theoretical Physics”
@Quantum Information, Quantum Matter and Quantum Gravity，YITP Kyoto U.

Short summary

Via gauge/gravity duality, how can we obtain geometric data from QFT side?
→ the slow mode plays an essential role



To determine the slow mode in a high-dim space, we compute a quantity

$$R_\infty(X;Y) := \min_U \left(\max_a \left| \left(U^\dagger X U - Y \right)_a \right| \right)$$

$X, Y : N \times N$ hermitian mat.

U : unitary mat.

which can be translated into an optimization problem.

We employ the Replica-Exchange Monte Carlo methods (REMC) and consider their extensions to solve this problem numerically.

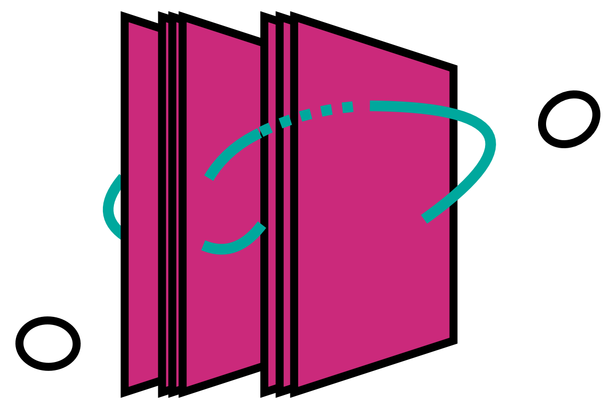
Contents

Gauge/gravity duality

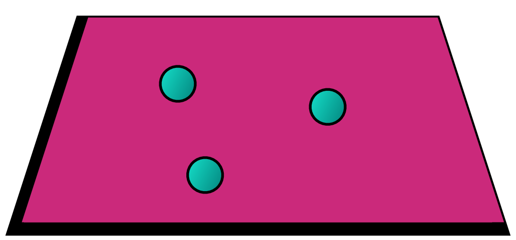
A conjecture from 2 descriptions of D-branes in string theory;

[Maldacena ('97) / Gubser, Klebanov, Polyakov, ('98) / Witten, ('98)]
[Itzhaki, Maldacena, Sonnenschein, Yankielowicz, ('98)]

D-branes & string



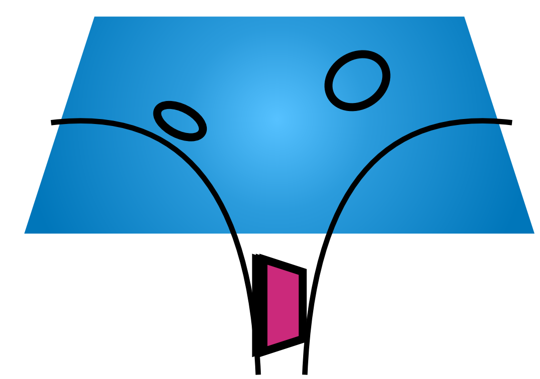
Effective theory of
open strings on branes



QFT
(Super Yang-Mills theory)

⟺

Theory of closed strings
in D-brane geometry



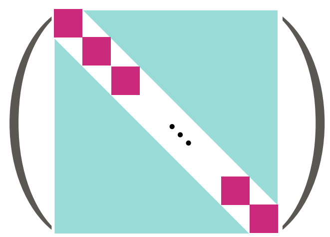
GR in curved spacetime

$$\int \mathrm{d}^{p+1}x \,\mathrm{tr}\left(\frac{1}{4}F_{\mu\nu}^{\ 2} + \frac{1}{2}(D_\mu X_I)^2 + \frac{g^2}{4}[X_I, X_J]^2 + (\text{fermion terms})\right)$$: (p+1)-d SU(N) SYM

$X_I(x)$: $N \times N$ hermitian matrices, $N \gg 1$ to satisfy the duality

Position of D-branes & open strings

For some special cases (e.g., 4d $\mathcal{N}=4$ SYM) → X : simultaneously diagonal

$X_I =$ 

diagonal : position of D-branes

off-diagonal : open string fluctuations among D-branes

[Witten, ('95)]

∵) Suppose $X_I = Y_I + \tilde{X}_I, \quad Y = \mathrm{diag}(y_1, \cdots, y_N),$

$$\mathrm{tr}\left[Y_I, X_J\right]^2 = \sum_{i,j} (Y_I X_J - X_I Y_J)^{ij} (Y_I X_J - X_I Y_J)^{ji} \supset -(y_I^i - y_I^j)^2 |\tilde{X}_I^{ij}|^2 \sim O(N^{-1})$$

And remember (open string mass) = (string tension) × (string length).

In a more generic case,

Key : Separation of the classical mode and fluctuation around it

[Polchinski, ('98, '99) / Susskind, ('99) / Hanada ('21)]

$$X_I = Y_I + \tilde{X}_I = (\text{slow mode}) + (\text{fast mode})$$

(One realization of slow mode : the center of wave packet in the matrix space)

How to identify the slow mode for generic theory?

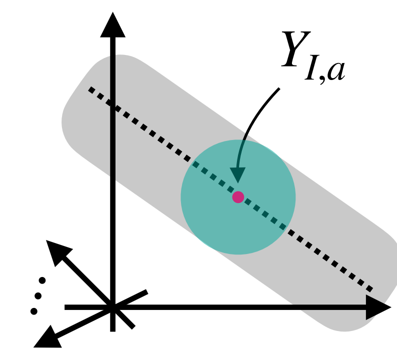
Determination of slow mode

How to identify the slow mode for generic theory?

Our proposal (in path-integral formalism) [Hanada, Kanno, Matsuura, HW, in progress]

: works for theories undefined in Hamiltonian formalism (e.g., matrix model)

Determine a specific configuration (= a point in matrix space)



c.f. [Hanada, ('21)] for a proposal in Hamiltonian formalism

- Prepare $\{X_I\}$, and find a unitary matrix U minimizing R_∞ with given $Y_I^{(\mathrm{trial})}$.

$$R_\infty(X; Y^{(\mathrm{trial})}) := \min_U \left(\max_{I,a} \left| \left(X_I^{(U)} - Y_I^{(\mathrm{trial})} \right)_a \right| \right)$$

: L_∞-distance or Chebyshev distance

($\min_U$) corresponds to the searching of Y along the gauge orbit

- Vary $Y_I^{(\mathrm{trial})}$ in order to search $\min_Y R_\infty(X, Y)$.
- Repeat above for different X_I, and take the average since $\langle R_\infty(X, Y_{\mathrm{min}}) \rangle$ is gauge invariant

: A variational approach

Contents

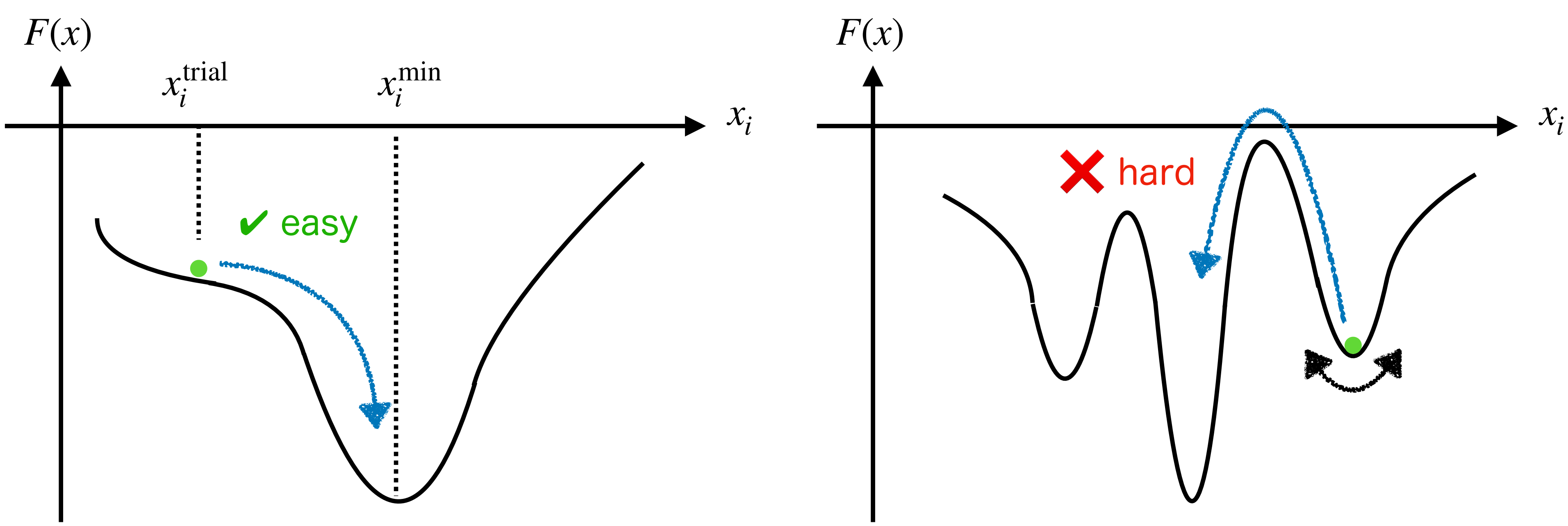
- D-brane geometry from matrix
- Monte Carlo methods to the minimization problem
- Numerical results
- Summary

MCMC & minimization

- Markov-chain Monte Carlo method enable us to generate the probability density $P(x)$ with a "potential" $F(x)$.

$$P(x) \propto \mathrm{e}^{-F(x)} \qquad x^{(0)} \to x^{(1)} \to x^{(2)} \to \cdots \to x^{(N_s)}$$

- This algorithm is powerful not only for performing integrals (e.g., lattice QCD) but also to search the minima of $F(x)$.
  - The difficulty depends on the structure of minima for $F(x)$. (: Ergodicity)
    (It always gives correct answers if we have an INFINITE computing resource!)



→concept of annealing which introduce a "temperature"

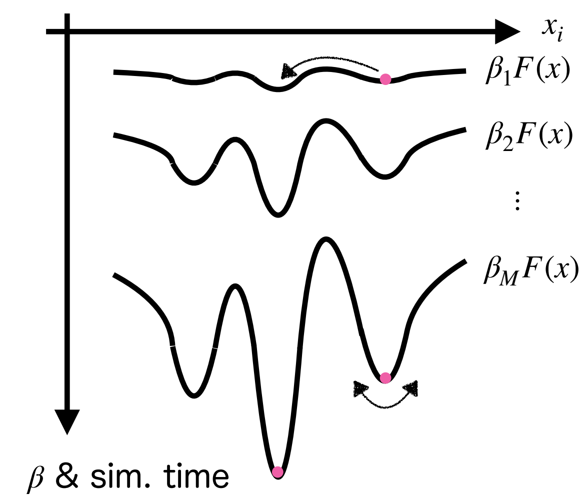
Simulated Annealing methods

- Naive SA is a method approximately searching the global minimum;

[Kirkpatrick, Gelatt, Vecchi, ('83)]

→ simulation time

MCMC with $\beta_1 F(x)$ → MCMC with $\beta_2 F(x)$ → ⋯ → MCMC with $\beta_M F(x)$



- $F(x)$: function to be minimized $\rightarrow R_\infty(U;X;Y) = \max_{I,a} |X_I^{(U)} - Y_I|_a,\ x \leftrightarrow U$
- $\beta_1 < \beta_2 < \cdots < \beta_M$: "inverse temperature" (← scaling of depth of potential)

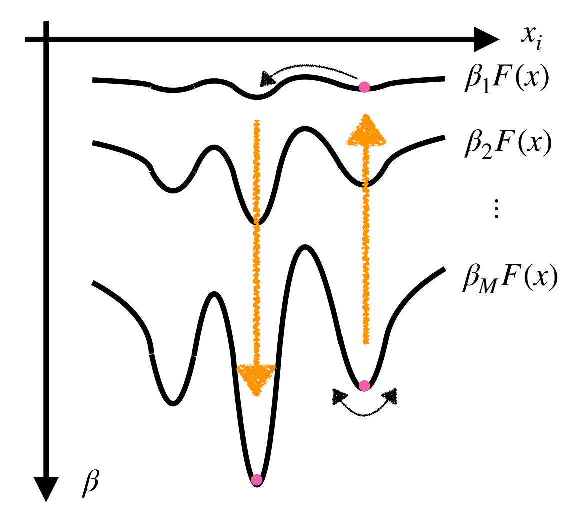
- Replica-Exchange MC method (REMC) is an upgraded method of SA;

[Swendsen, Wang, ('86)]
[Geyer, ('91)]

→ simulation time

MC with $\beta_1 F(x)$		MC with $\beta_1 F(x)$
MC with $\beta_2 F(x)$	→	MC with $\beta_2 F(x)$
⋮		⋮
MC with $\beta_M F(x)$		MC with $\beta_M F(x)$

→ ⋯



Simulations of different β simultaneously (more accurate but expensive)

Extension of replica actions

A "regularization" of the function aiming to escape from wrong convergence

original problem

$$R_\infty(U;X;Y) = \max_{I,a} |X_I^{(U)} - Y_I|_a$$

: L_∞-distance

new problem

$$R_p(U;X;Y) = \left(\sum_{I,a} |X_I^{(U)} - Y_I|_a{}^p \right)^{1/p}$$

: L_p-distance

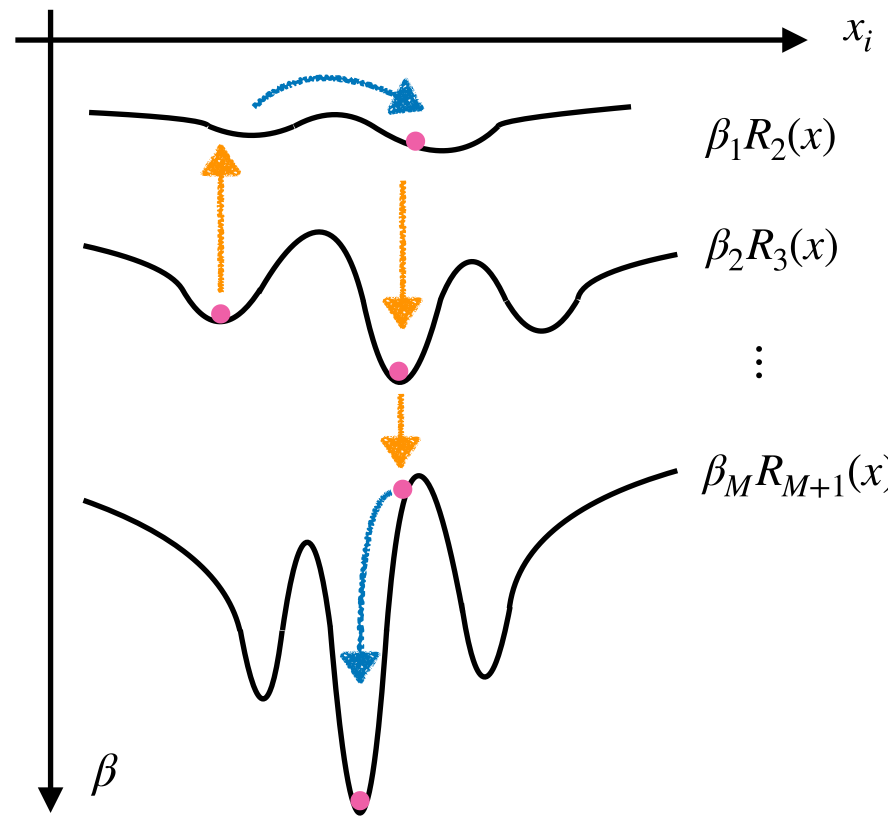
[Hanada, Kanno, Matsuura, HW, in progress]

- Different pot. structure among replicas
  → many minimizing path
- Less local minima for smaller p

∵) $R_2(X^{(U)}, Y_I) = \min_U \sqrt{\mathrm{tr}(X_I^{(U)} - Y_I)^2} = \sqrt{\mathrm{tr}(X_I^{(U)} - Y_I)^2}$

: gauge inv.

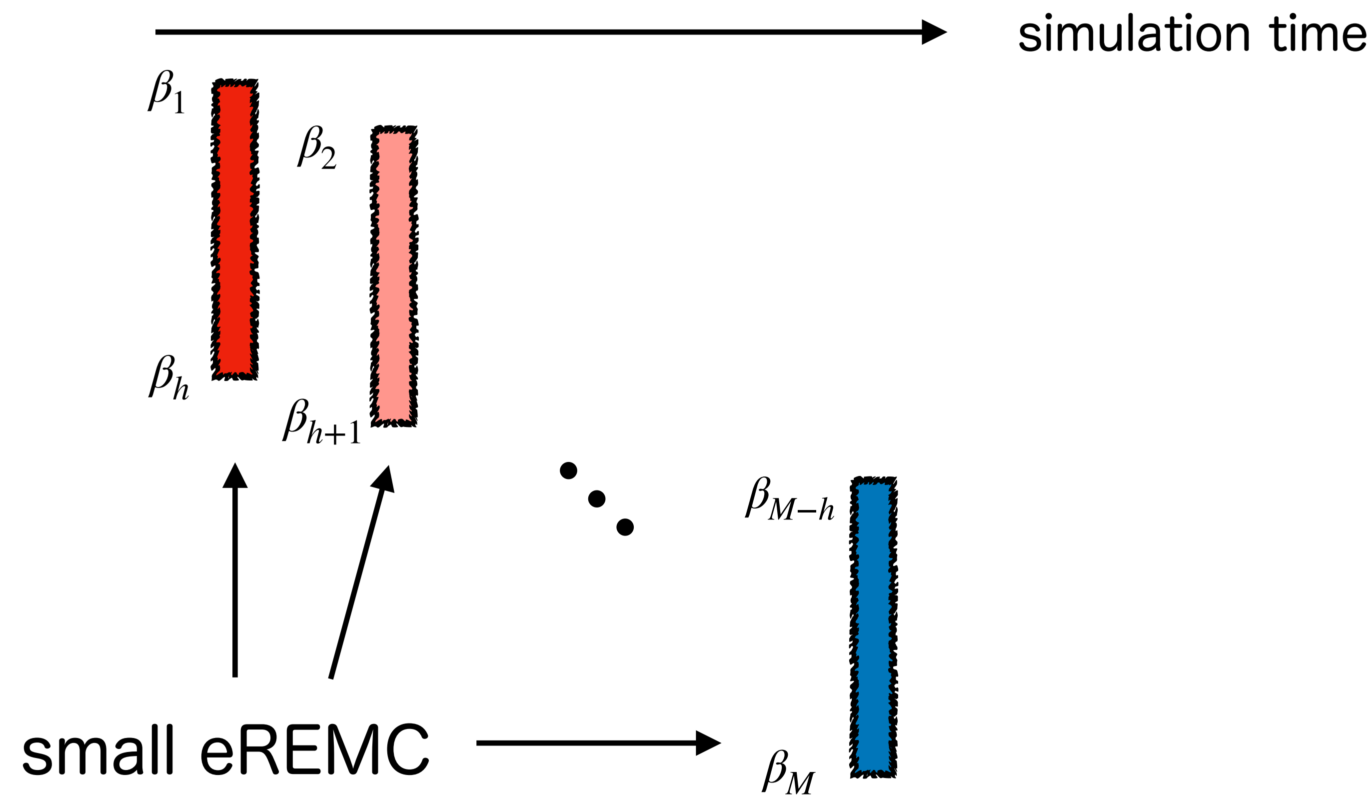
→ Expecting a gain of efficiency



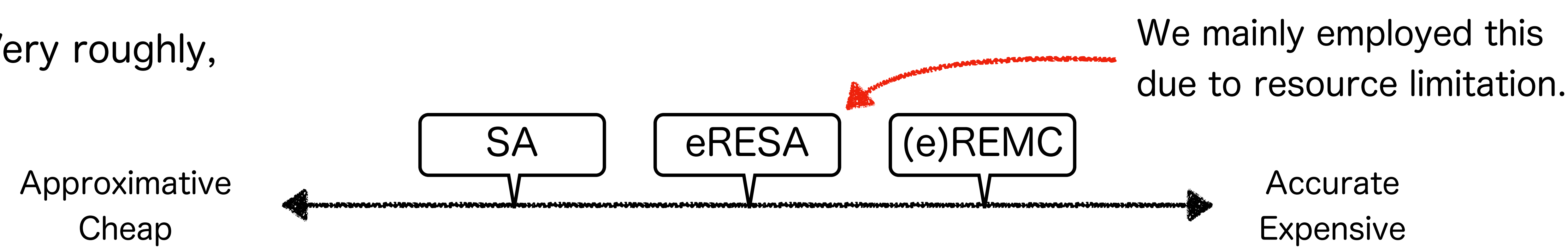
Extended Replica-Exchange SA

[Hanada, Kanno, Matsuura, HW, in progress]

eRESA = Annealing of REMC with small replicas w/ extended replica action



Very roughly,

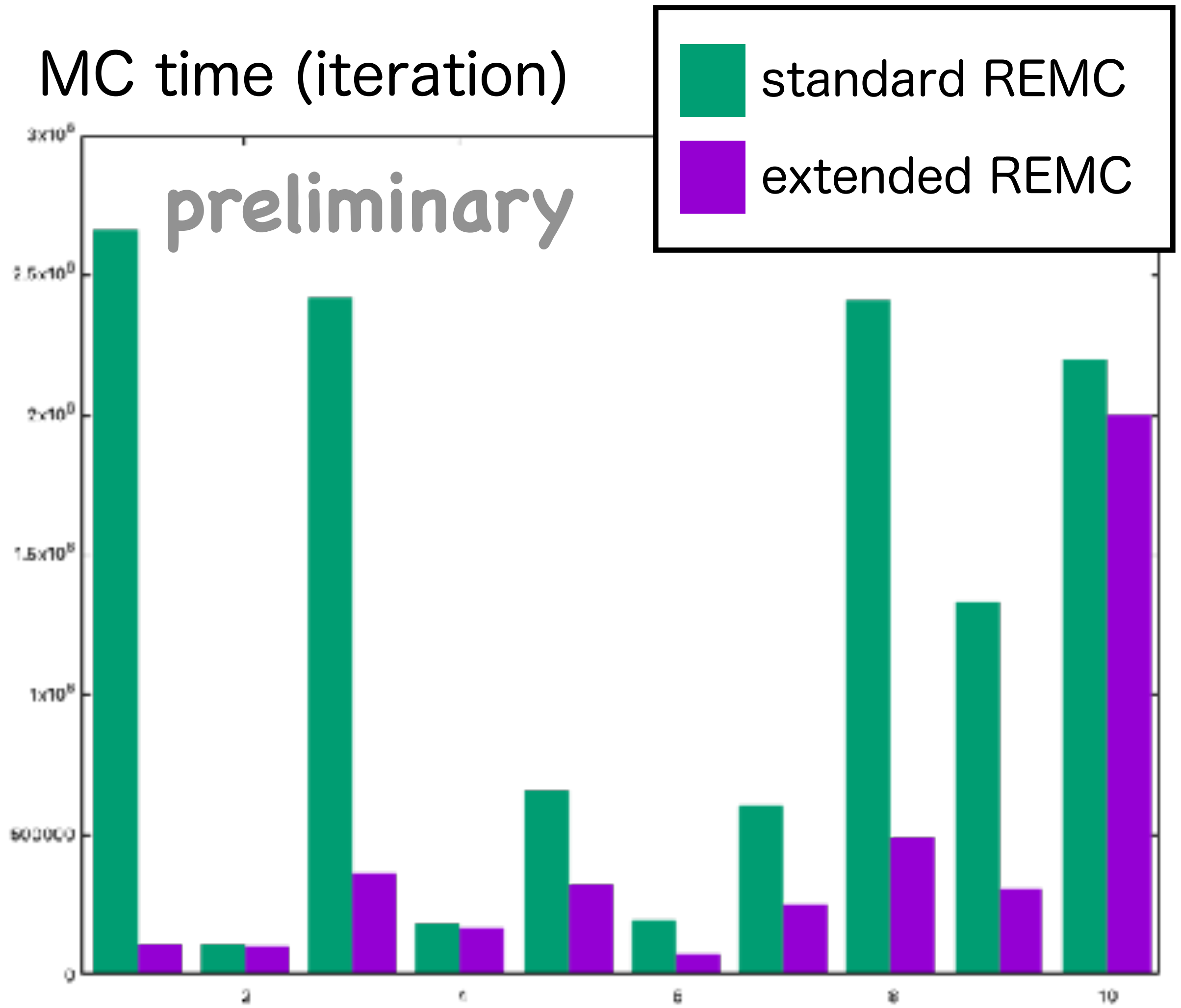


Contents

Prep.: Mock-data analysis

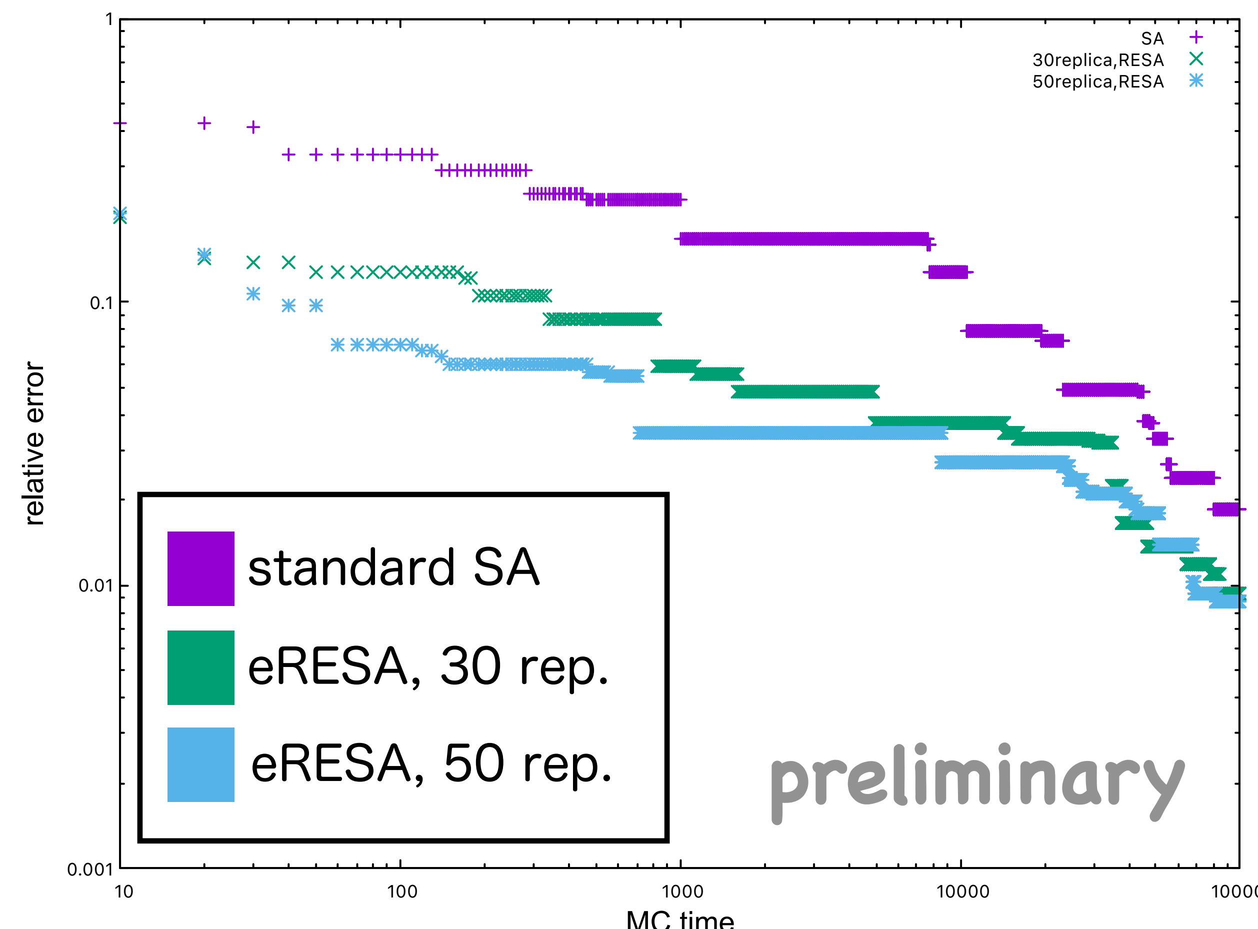
Demonstration: one 4×4 matrix in which we know the answer

MC time (iteration)



(#replica = 500, (threshold) = 1.025)

Relative error



MC time

→ Minimization by eREMC, eRESA tends faster than standard ones.

Example: Fuzzy sphere matrix model

[Iso, Kimura, Tanaka, Wakatsuki, ('01)]

A supersymmetric toy model; $(I, J, K = 1,2,3)$

$$S(X_I, \psi) = N \mathrm{tr} \left(-\frac{1}{4}[X_I, X_J]^2 + \frac{2i\mu}{3}\epsilon_{IJK} X_I X_J X_K + \frac{1}{2}\bar{\psi}\sigma^I [X_I, \psi] + \mu \bar{\psi}\psi \right)$$

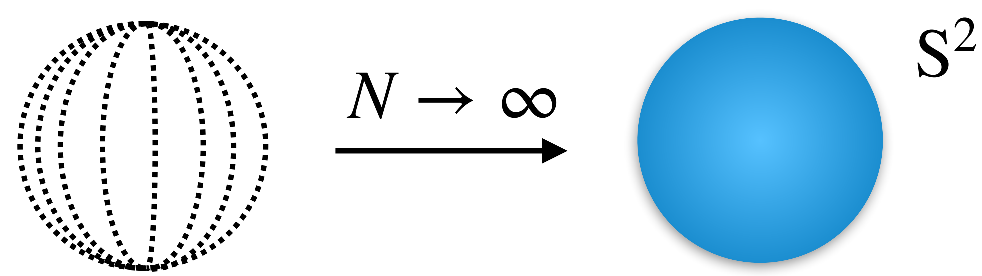
(σ_I : Pauli matrices)

: X_I's are not simultaneously diagonalizable!

Classical minima : Fuzzy sphere solution

$$X_I^{\mathrm{FS}} = \mu J_I, \qquad \left[J_I, J_J\right] = i\epsilon_{IJK} J_K$$

J_I : N-dim. irrep. of SU(2) generator



$$R_{\mathrm{FS}}^2 = \frac{1}{N}\mathrm{tr}\, X_I^2 = \frac{\mu^2}{4}(N^2 - 1)$$

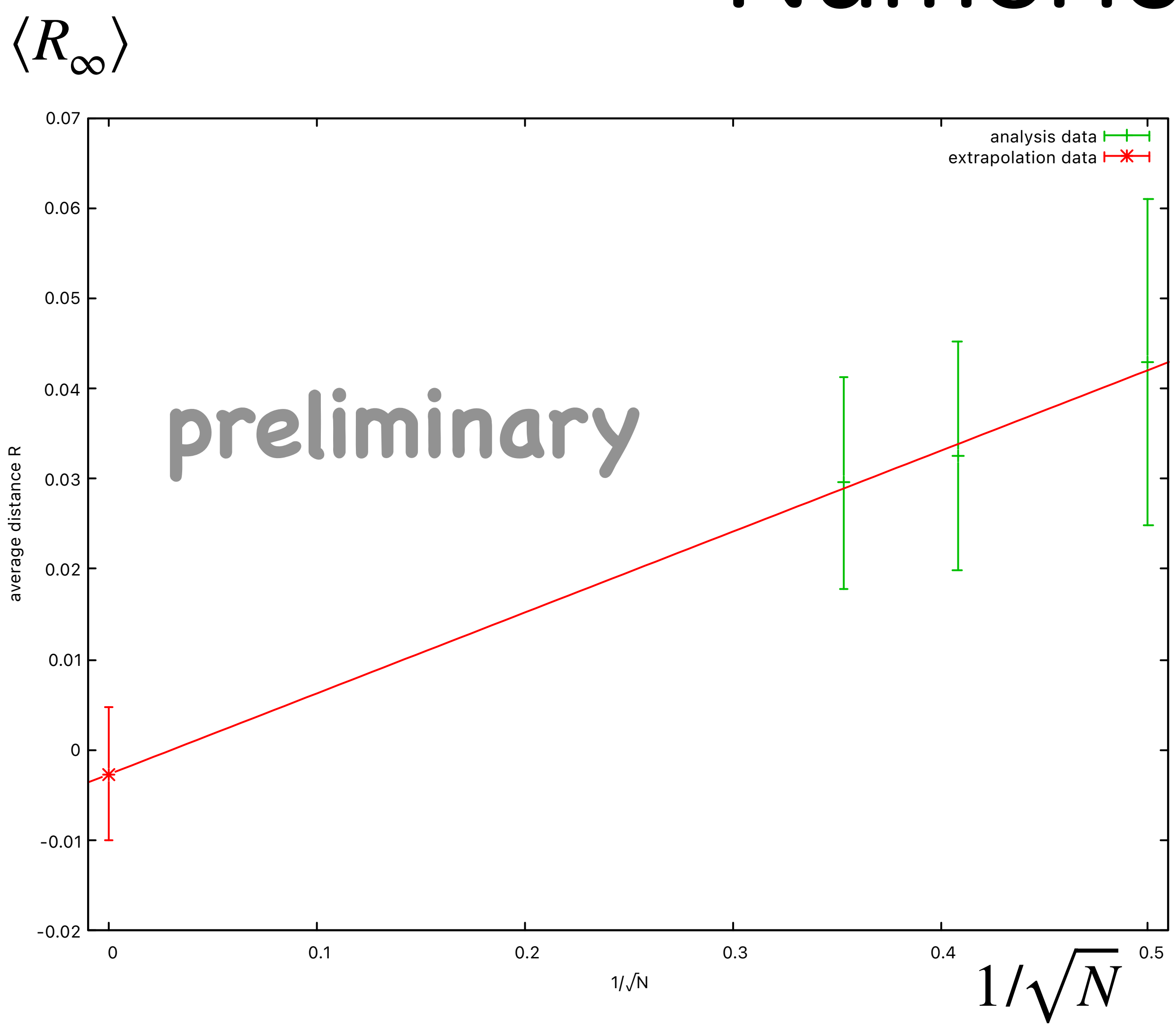
Minimization of the distance w.r.t. U by eRESA

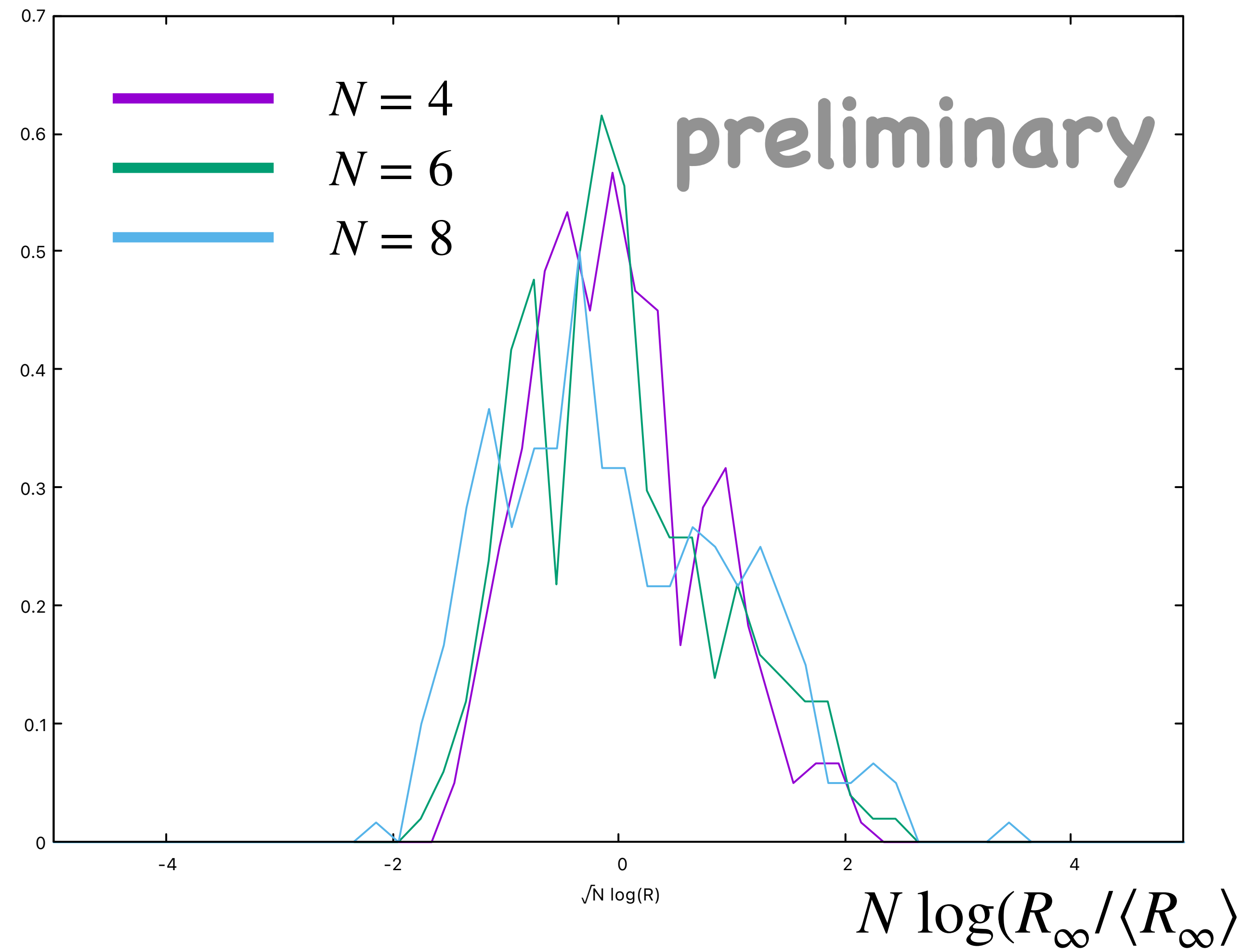
$$R_\infty(U; X; X_I^{\mathrm{FS}}) = \max_{I,a} | X_I^{(U)} - X_I^{\mathrm{FS}} |_a \longrightarrow R_p(U; X; X_I^{\mathrm{FS}}) = \left(\sum_{I,a} | X_I^{(U)} - X_I^{\mathrm{FS}} |_a^{\,p} \right)^{1/p}$$

(Under investigation for varying the ansatz of slow mode Y_I)

Numerical results

(at $\mu = 10$)





[⬅️ Left] Large N extrapolation shows an $1/\sqrt{N}$ scaling and convergence to zero.

[Right ➡️] Histogram of $R_\infty(X_I, X_I^{\mathrm{FS}})$ shows that width scales by N^{-1}.

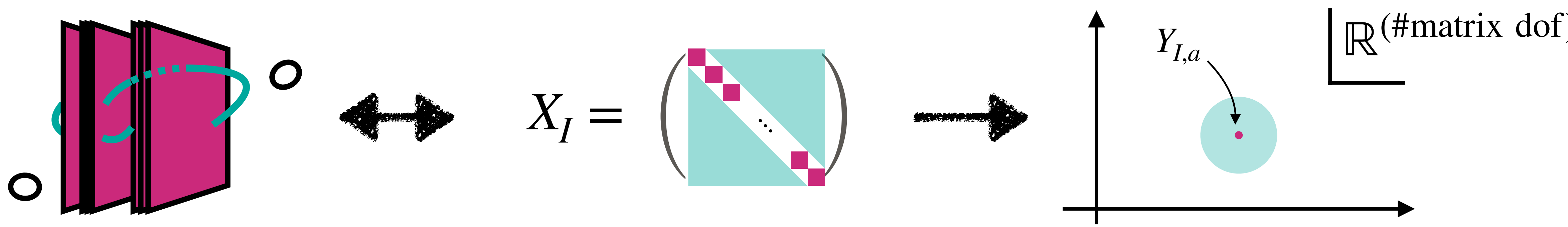
→ Consistent with the theoretical prediction!

$$\mathrm{tr}(X_I - Y_I)^2 = \sum_a \left| X_I - Y_I \right|_a^2 \sim O(N), \qquad R_\infty \sim \max_a \left| X_I - Y_I \right|_a \sim O(N^{-1/2})$$

Contents

Summary

To read off geometric information in string theory via gauge/gravity duality, we have to extract the slow mode from matrices (c.f. center of wave packet).



To determine the slow mode in a high-dim space, we compute a quantity

$$R_\infty(X;Y) := \min_U \left(\max_a \left| \left(U^\dagger X U - Y \right)_a \right| \right)$$

$X, Y : N \times N$ hermitian mat.

U : unitary mat.

We employ the Replica-Exchange Monte Carlo methods (REMC) and consider their extensions (extension of the replica action, and RE Simulated Annealing)

- Mock-data analysis
- Fuzzy sphere three-matrix model

(+ One-matrix model w/ double-well pot.)

Future directions

- More detailed analysis for (bosonic) fuzzy-sphere three-matrix model
  - Interesting in small μ region where FS seems obscure

    in bosonic model exhibiting phase transition btw large/small μ

- Investigation of (0+1)d models both in path-integral & Hamiltonian formalisms.
  - Quantum computations may be powerful to determine the quantum state of wave packet corresponding to emergent geometry.
- Necessity to clarify how to find a better ansatz for Y_I

  $\leftarrow$ Essential for analyzing (0+1)d models (e.g. BFSS-type model) and so on.

[Banks, Fischler, Shenker, Susskind, ('96)]
[Berenstein, Maldacena, Nastase, ('02)]

- Further understanding, generalization, application of the extended REMCs
  - Combination of RE method with Machine Learning?