

2^{nd} year PhD report

Mauro D’Arcangelo
Supervisor: Prof. J. W. Barrett

Contents

1	Introduction	2
1.1	Non-commutative geometry	2
1.2	Fuzzy spaces	2
1.3	Random geometries	3
2	Results	4
2.1	Improvements in the numerical algorithm	4
2.1.1	Hybrid Monte Carlo	4
2.1.2	Adapting HMC to fuzzy spaces	5
2.1.3	The case $p = 2$	5
2.1.4	The case $p = 4$	6
2.1.5	Testing HMC against exact results	9
2.1.6	Comparing the speed of HMC and Metropolis	10
2.2	Yang-Mills matrix models from Dirac operators	13
2.2.1	Yang-Mills matrix models	13
2.2.2	Type $(0, 3)$ geometry as a Yang-Mills matrix model	13
2.3	Reduced model for the $(2, 0)$ fuzzy Dirac operator	16
2.3.1	Notation	16
2.3.2	Stationary manifold	16
2.3.3	Small oscillations and quadratic action	17
3	Future directions	19
3.1	Including matter fields	19

1 Introduction

This project is concerned with the characterization of *random fuzzy spaces* by means of Markov chain Monte Carlo simulations. The chapter presents a brief introduction to the basic concepts (non-commutative, fuzzy and random).

1.1 Non-commutative geometry

The fundamental object of non-commutative geometry is a *spectral triple* (A, H, D) where A is an algebra with a representation in a Hilbert space H and D is an operator on H , called Dirac operator. A Riemannian spin manifold can be fully characterized by the commutative algebra A of functions on the manifold and by the Dirac operator, which encodes the metric [11] [10]. One could then consider a generalization in which the algebra is allowed to be non-commutative. Such geometries arise naturally in physics, and are tightly related to gauge theories [23]. Indeed it has been shown that the Standard Model has the structure of a non-commutative geometry [8] [3] [9]. This suggests a possible path to quantum gravity by replacing the ordinary commutative spacetime by a non-commutative one that presents a commutative behaviour as a limiting case.

1.2 Fuzzy spaces

There is a class of non-commutative geometries called *fuzzy spaces*, where the algebra is taken to be $M_n(\mathbb{C})$, the algebra of $n \times n$ complex matrices, and the Hilbert space is finite dimensional. The Dirac operator of a fuzzy space takes the form [4]:

$$D = \sum_j \alpha_j \otimes [L_j, \cdot] + \sum_k \tau_k \otimes \{H_k, \cdot\} \quad (1)$$

where:

1. τ_k and α_j are respectively Hermitian and anti-Hermitian basis elements of the algebra generated by a (p, q) Clifford module;
2. H_k and L_j are $n \times n$ Hermitian and anti-Hermitian matrices respectively;
3. $[\cdot, \cdot]$ indicates a commutator and $\{\cdot, \cdot\}$ an anti-commutator.

Fuzzy spaces are classified by the pair of integers (p, q) of the Clifford module. It is worth noting that any choice of H_k and L_j gives an admissible Dirac operator, as long as the Hermitian or anti-Hermitian character is preserved. See Appendix A for some remarks on the notation used.

1.3 Random geometries

A *random geometry* is a spectral triple (A, H, D) in which the Dirac operator fluctuates according to a certain probability measure. Here the probability measure is taken to be proportional to:

$$e^{-S[D]} dD \quad (2)$$

for a certain choice of $S[D]$. The expectation value of an observable $f(D)$ on a random geometry is given by:

$$\langle f(D) \rangle = \int f(D) e^{-S[D]} dD. \quad (3)$$

Since D encodes the metric, this is in clear analogy with the Euclidean path integral of Quantum Field Theory.

So far no assumption has been made on the choice of Dirac operator. The purpose of this project is to study the path integral when D is taken to be the Dirac operator of a fuzzy space.

Fuzzy spaces provide an alternative type of regularization that is non-lattice [6]. Therefore the study of such random geometries is especially interesting in connection with models of (Euclidean) quantum gravity.

This line of research first appeared in [6], where the following action was considered:

$$S[D] = g_2 \operatorname{Tr} D^2 + \operatorname{Tr} D^4, \quad g_2 \in \mathbb{R}. \quad (4)$$

In the remainder, the action is taken to be of the form of Eq.(4).

The expectation value (3) is computed numerically using Monte Carlo methods. Eq.(3) is therefore replaced by:

$$\langle f(D) \rangle \approx \frac{1}{N} \sum_{i=1}^N f(D_i) \quad (5)$$

where $\{D_i\}$ is a set of Dirac operators sampled from the distribution (2).

2 Results

2.1 Improvements in the numerical algorithm

The evaluation of integrals such as (3) requires a way to sample the most relevant configurations (the *typical set*) out of the entirety of parameter space. The simplest algorithm based on Markov chains for doing so is Metropolis-Hastings [16]. Two crucial shortcomings of Metropolis are a rather slow exploration of the typical set based on a random walk in parameter space, and the large correlation between adjacent samples. A more sophisticated approach is Hybrid Monte Carlo, or HMC [14]: originally developed for lattice QCD computations, it allows a faster and more uniform exploration of the typical set by transposing the problem of sampling from a distribution to Hamiltonian evolution in a fictitious phase space.

After a brief explanation of the idea behind HMC, the adaptation to the fuzzy space path integral is discussed and the results from numerical simulations are presented with a comparison between HMC and Metropolis.

2.1.1 Hybrid Monte Carlo

In Markov chain theory one is interested in a system specified by a finite set of parameters (q_1, \dots, q_N) , $q_i \in \mathbb{R}$, collectively referred to as \mathbf{q} . The probability that the system be in a particular configuration is given by some probability measure $\pi(\mathbf{q})d\mathbf{q}$. Given an initial configuration \mathbf{q} , a Markov chain establishes a transition $\mathbf{q} \rightarrow \mathbf{q}'$ from the old configuration to a new one in such a way that \mathbf{q}' is chosen with the desired probability.

This situation is analogous to the Monte Carlo estimation of the integral (3), where the parameters are the independent degrees of freedom of the Dirac operator and the probability measure is $e^{-S[D]}dD$.

The first step of Hybrid Monte Carlo is to enlarge parameter space by introducing a “conjugate momentum” p_i to each parameter q_i , thus effectively working in a fictitious phase space. The probability measure is extended to include the new variables $\pi(\mathbf{q}) \rightarrow \pi(\mathbf{q}, \mathbf{p})$. By defining the “Hamiltonian” $H(\mathbf{q}, \mathbf{p}) \equiv -\log \pi(\mathbf{q}, \mathbf{p})$, a configuration is then evolved along a Hamiltonian trajectory by integrating Hamilton’s equations:

$$\frac{dq_i}{dt} = \frac{\partial H}{\partial p_i} \tag{6}$$

$$\frac{dp_i}{dt} = -\frac{\partial H}{\partial q_i} \tag{7}$$

where t denotes a fictitious time that will be referred to as Monte Carlo time. This specifies a transition $(\mathbf{q}(0), \mathbf{p}(0)) \rightarrow (\mathbf{q}(t), \mathbf{p}(t))$ from an initial

configuration to a new one. The new configuration is then accepted with probability $\min[1, \exp(H(t) - H(0))]$.

Note that Hamiltonian dynamics preserves the value of the energy of the system, therefore $H(t) = H(0)$ and $\min[1, \exp(H(t) - H(0))] = 1$. However, numerical integration of Hamilton's equations is a non-trivial matter that introduces errors, therefore $H(t) \neq H(0)$ in general. The standard choice of numerical integrator in HMC is the so-called *leapfrog* integrator [19] [7].

2.1.2 Adapting HMC to fuzzy spaces

In the matrix integral considered here, the dynamical variables are the $n \times n$ Hermitian matrices M_i , to which a corresponding set of Hermitian matrices P_i is added. The Hamiltonian is chosen to be simply:

$$H(M_i, P_i) = S[M_i] + \sum_i \frac{1}{2} P_i^2. \quad (8)$$

Schematically, the algorithm goes as follows:

1. extract the momenta P_i according to $\exp(-P_i^2/2)$;
2. integrate Hamilton's equations for a certain time t ;
3. accept the new configuration with probability $\min[1, \exp(H(t) - H(0))]$.

The non-trivial step in the leapfrog integrator is the evaluation of the force term in Eq.(YYYY), which requires to take derivatives such as:

$$\frac{\partial S[M_i]}{\partial M_k} \quad (9)$$

which amounts to finding formulas for terms like:

$$\frac{\partial \text{Tr } D(M_i)^p}{\partial M_k}. \quad (10)$$

For the definition of matrix derivative see Appendix B. In the following, formulas for $p = 2$ and $p = 4$ are developed.

2.1.3 The case $p = 2$

When $p = 2$ the M_i matrices are decoupled:

$$\text{Tr } D^2 = \sum_i \text{Tr } \omega_i^2 (2n \text{Tr } M_i^2 + 2\epsilon_i (\text{Tr } M_i)^2). \quad (11)$$

Taking a derivative with respect to M_k yields:

$$\begin{aligned} \frac{\partial}{\partial M_k} \left(\sum_i \text{Tr} \omega_i^2 (2n \text{Tr} M_i^2 + 2\epsilon_i (\text{Tr} M_i)^2) \right) = \\ \sum_i \delta_{ik} \text{Tr} \omega_i^2 (4n M_i^T + 4\epsilon_i (\text{Tr} M_i) I) = \\ 4C (n M_k^T + \epsilon_k (\text{Tr} M_k) I) \end{aligned} \quad (12)$$

where $C \equiv \text{Tr} \omega_i^2$ is the dimension of the Clifford module.

2.1.4 The case $p = 4$

First expand $\text{Tr} D^4$:

$$\begin{aligned} \text{Tr} D^4 = \sum_{i_1, i_2, i_3, i_4} \text{Tr}(\omega_{i_1} \omega_{i_2} \omega_{i_3} \omega_{i_4}) \cdot \\ \left(n[1 + \epsilon *] \text{Tr}(M_{i_1} M_{i_2} M_{i_3} M_{i_4}) + \right. \\ \epsilon_{i_1} \text{Tr} M_{i_1} [1 + \epsilon *] \text{Tr}(M_{i_2} M_{i_3} M_{i_4}) + \\ \epsilon_{i_2} \text{Tr} M_{i_2} [1 + \epsilon *] \text{Tr}(M_{i_1} M_{i_3} M_{i_4}) + \\ \epsilon_{i_3} \text{Tr} M_{i_3} [1 + \epsilon *] \text{Tr}(M_{i_1} M_{i_2} M_{i_4}) + \\ \epsilon_{i_4} \text{Tr} M_{i_4} [1 + \epsilon *] \text{Tr}(M_{i_1} M_{i_2} M_{i_3}) + \\ \epsilon_{i_1} \epsilon_{i_2} [1 + \epsilon] \text{Tr}(M_{i_1} M_{i_2}) \text{Tr}(M_{i_3} M_{i_4}) + \\ \epsilon_{i_1} \epsilon_{i_3} [1 + \epsilon] \text{Tr}(M_{i_1} M_{i_3}) \text{Tr}(M_{i_2} M_{i_4}) + \\ \left. \epsilon_{i_1} \epsilon_{i_4} [1 + \epsilon] \text{Tr}(M_{i_1} M_{i_4}) \text{Tr}(M_{i_2} M_{i_3}) \right) \end{aligned} \quad (13)$$

where $*$ denotes complex conjugation of everything that appears on the right, ϵ is defined as the product $\epsilon \equiv \epsilon_{i_1} \epsilon_{i_2} \epsilon_{i_3} \epsilon_{i_4}$, and the relation $M^T = M^*$ has been used. Since D is Hermitian, the expression must be real. It is not immediate to see that this is the case because of the $\epsilon = \pm 1$ factor inside the square brackets. Reality nonetheless holds, and becomes manifest by observing that a simultaneous index exchange $i_1 \leftrightarrow i_4$ and $i_2 \leftrightarrow i_3$ is equivalent to taking the complex conjugate (in fact, this is not the only index exchange that amounts to complex conjugation).

Taking a matrix derivative with respect to M_k results in non-vanishing con-

tributions when $k = i_1, k = i_2, k = i_3$ or $k = i_4$:

$$\begin{aligned} \frac{\partial}{\partial M_k} \text{Tr } D^4 &= \sum_{i_1, i_2, i_3, i_4} \text{Tr}(\omega_{i_1} \omega_{i_2} \omega_{i_3} \omega_{i_4}) \cdot \\ &\quad \left(\delta_{ki_1} A(i_1, i_2, i_3, i_4)^T + \delta_{ki_2} A(i_2, i_3, i_4, i_1)^T + \right. \\ &\quad \left. \delta_{ki_3} A(i_3, i_4, i_1, i_2)^T + \delta_{ki_4} A(i_4, i_1, i_2, i_3)^T \right) \end{aligned} \quad (14)$$

where $A(a, b, c, d)$ is the following $n \times n$ matrix:

$$\begin{aligned} A(a, b, c, d) &\equiv n[1 + \epsilon^\dagger] M_b M_c M_d + \\ &\quad \epsilon_a I[1 + \epsilon^*] \text{Tr } M_b M_c M_d + \\ &\quad \epsilon_b \text{Tr } M_b [1 + \epsilon^\dagger] M_c M_d + \\ &\quad \epsilon_c \text{Tr } M_c [1 + \epsilon^\dagger] M_b M_d + \\ &\quad \epsilon_d \text{Tr } M_d [1 + \epsilon^\dagger] M_b M_c + \\ &\quad \epsilon_a \epsilon_b M_b [1 + \epsilon] \text{Tr } M_c M_d + \\ &\quad \epsilon_a \epsilon_c M_c [1 + \epsilon] \text{Tr } M_b M_d + \\ &\quad \epsilon_a \epsilon_d M_d [1 + \epsilon] \text{Tr } M_b M_c \end{aligned} \quad (15)$$

and \dagger denotes Hermitian conjugation of everything that appears on the right. Upon relabeling the indices and cycling the ω matrices in the trace, the equation becomes:

$$\begin{aligned} \frac{\partial}{\partial M_k} \text{Tr } D^4 &= 4 \sum_{i_1, i_2, i_3, i_4} \delta_{ki_1} \text{Tr}(\omega_{i_1} \omega_{i_2} \omega_{i_3} \omega_{i_4}) A(i_1, i_2, i_3, i_4)^T \\ &= 4 \sum_{i_1, i_2, i_3} \text{Tr}(\omega_k \omega_{i_1} \omega_{i_2} \omega_{i_3}) A(k, i_1, i_2, i_3)^T \equiv 4 \sum_{i_1, i_2, i_3} \mathcal{B}_k(i_1, i_2, i_3) \end{aligned} \quad (16)$$

with $\mathcal{B}_k(a, b, c)$ denoting the generic term in the sum.

To see that Eq.(16) defines a Hermitian matrix, notice that an exchange of indices $i_1 \leftrightarrow i_3$ is equivalent to taking the Hermitian conjugate:

$$\mathcal{B}_k(i_1, i_2, i_3)^\dagger = \mathcal{B}_k(i_3, i_2, i_1) \quad (17)$$

therefore the sum in Eq.(16) reduces to:

$$\sum_{\substack{i_1 > i_3 \\ i_2}} [1 + \dagger] \mathcal{B}_k(i_1, i_2, i_3) + \sum_{i_1, i_2} \mathcal{B}_k(i_1, i_2, i_1). \quad (18)$$

In fact, by looking at the form of \mathcal{B} , it is clear that terms in the sum are qualitatively different based on the number of indices that coincide. Therefore it would be computationally convenient to write Eq.(18) in a way that emphasises this difference.

The only terms that contribute when all indices are different are the following:

$$\sum_{i_1 > i_2 > i_3} [1 + \dagger] \left(\mathcal{B}_k(i_1, i_2, i_3) + \mathcal{B}_k(i_1, i_3, i_2) + \mathcal{B}_k(i_2, i_1, i_3) \right). \quad (19)$$

The three inequivalent permutations of indices that appear in this formula are based on a group-theoretical argument that will generalize easily to powers of D higher than 4. First consider the symmetric group of order three S_3 acting on the set of indices $\{i_1, i_2, i_3\}$, and the subgroup of permutations that induce a simple change in \mathcal{B} , which in this case is $H = \{(), (13)\} \cong S_2$ (the first element being the identical permutation, and the second the exchange $i_1 \leftrightarrow i_3$ which induces $\mathcal{B} \rightarrow \mathcal{B}^\dagger$). The idea is then to restrict the sum to $i_1 > i_2 > i_3$ and quotient out the action of H by introducing a suitable pre-factor that accounts for it (in this case $[1 + \dagger]$). Practically, the inequivalent permutations of indices that appear in Eq.(19) are found by computing the (left or right) cosets of $H \subset S_3$ and acting on $\{i_1, i_2, i_3\}$ with a representative from each coset. In this case the representatives were chosen to be $()$, (23) , (12) .

What is left are terms in which at least two indices are equal. These are:

$$\begin{aligned} \sum_{i_1 > i_2} [1 + \dagger] \left(\mathcal{B}_k(i_1, i_1, i_2) + \mathcal{B}_k(i_1, i_2, i_2) \right) + \\ \sum_{i_1 \neq i_2} \mathcal{B}_k(i_1, i_2, i_1) + \sum_i \mathcal{B}_k(i, i, i). \end{aligned} \quad (20)$$

At this point, a useful property of the ω matrices can be exploited to simplify both Eq.(19) and Eq.(20):

$$\text{Tr}(\omega_{\sigma(i_1)} \omega_{\sigma(i_2)} \omega_{\sigma(j)} \omega_{\sigma(k)}) \propto \text{Tr}(\omega_j \omega_k) = 0 \quad \text{if } i_1 = i_2 \text{ and } j \neq k \quad (21)$$

for any permutation σ acting on $\{i_1, i_2, j, k\}$. In other words, if two indices are the same and the other two are different, the trace on the ω matrices vanishes.

Putting together Eq.(19), Eq.(20) and Eq.(21), the final formula for $\partial_k \text{Tr } D^4$

reads:

$$\begin{aligned} \frac{\partial}{\partial M_k} \text{Tr } D^4 = 4 \left[\sum_{\substack{i_1 > i_2 > i_3 \\ i_1, i_2, i_3 \neq k}} [1 + \dagger] \left(\mathcal{B}_k(i_1, i_2, i_3) + \mathcal{B}_k(i_1, i_3, i_2) + \mathcal{B}_k(i_2, i_1, i_3) \right) \right. \\ \left. + \sum_{\substack{i \\ i \neq k}} \left([1 + \dagger] \mathcal{B}_k(i, i, k) + \mathcal{B}_k(i, k, i) \right) + \mathcal{B}_k(k, k, k) \right]. \end{aligned} \quad (22)$$

The explicit form of $\mathcal{B}_k(i, i, k)$, $\mathcal{B}_k(i, k, i)$ and $\mathcal{B}_k(k, k, k)$ is given in Appendix C.

2.1.5 Testing HMC against exact results

Although not many analytical results are available for the matrix integrals considered here, there is a non-trivial observable whose expectation value is known exactly for any geometry and any action polynomial in D . Computing this observable therefore provides a good test for the numerical implementation.

Consider the following vanishing integral:

$$0 = \int \frac{\partial}{\partial D_{ij}} (D_{ij} e^{-S[D]}) dD \quad (23)$$

where D_{ij} is one of the independent degrees of freedom of the Dirac operator and $S[D]$ is in general sum of terms like $g_p \text{Tr } D^p$.

An explicit calculation of the right-hand side of Eq.(23) gives:

$$0 = \int \left(1 - \sum_p g_p D_{ij} \frac{\partial}{\partial D_{ij}} \text{Tr } D^p \right) e^{-S[D]} dD \quad (24)$$

dividing by the partition function and rearranging the terms:

$$1 = \frac{1}{\int e^{-S[D]}} \int \left(\sum_p g_p D_{ij} \frac{\partial}{\partial D_{ij}} \text{Tr } D^p \right) e^{-S[D]} dD \quad (25)$$

and finally summing over all i, j corresponding to independent degrees of freedom:

$$\# \text{d.o.f.}(D) = \frac{1}{\int e^{-S[D]}} \int \left(\sum_p g_p p \text{Tr } D^p \right) e^{-S[D]} dD = \left\langle \sum_p g_p p \text{Tr } D^p \right\rangle \quad (26)$$

where the following identity, a proof of which is given in Appendix D, was used:

$$\sum_{ij} D_{ij} \frac{\partial}{\partial D_{ij}} \text{Tr } D^p = p \text{Tr } D^p. \quad (27)$$

To estimate the number of degrees of freedom recall that in the decomposition (65) the M_i matrices are all Hermitian, therefore $\# \text{d.o.f.}(D) = n^2 m$ where n is the dimension and m the total number of M_i matrices.

Numerical tests of HMC and Metropolis agree with the prediction (26). As an example, a test on the $(0, 3)$ Dirac operator is reported in Fig.(1).

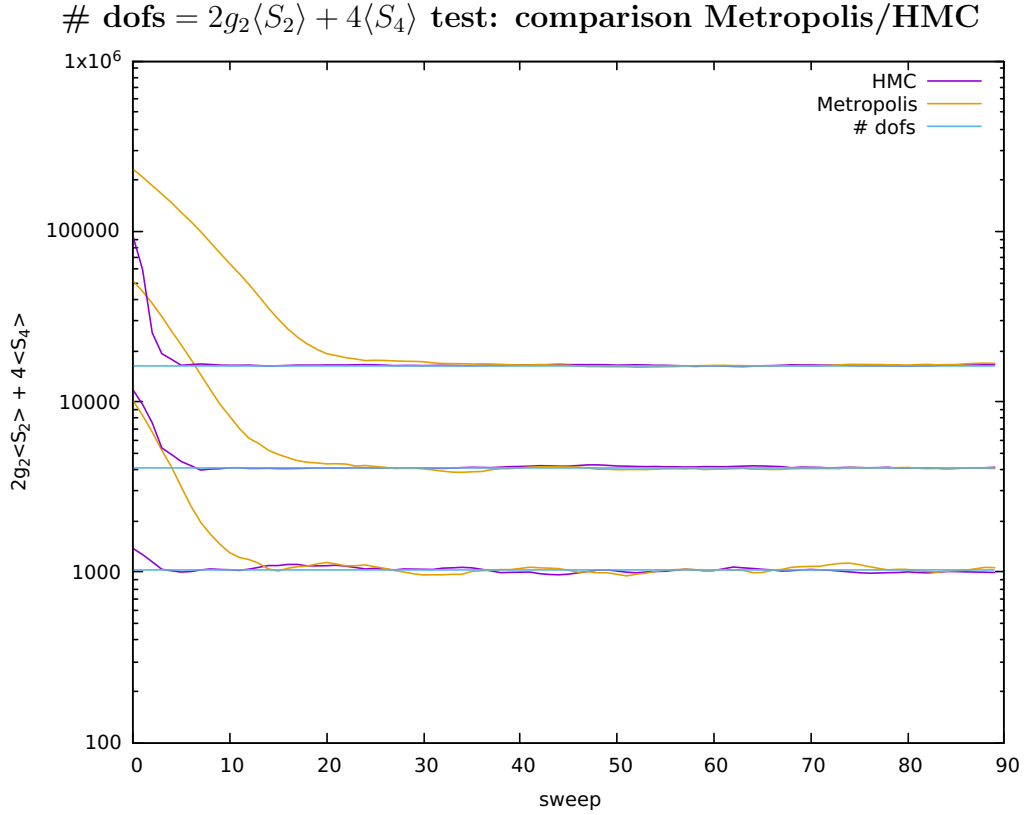


Figure 1: Convergence of the expectation value of the observable $2g_2 \text{Tr } D^2 + 4 \text{Tr } D^4$ to the number of degrees of freedom of the Dirac operator for geometry $(0, 3)$. The three sets correspond to matrix size $n = 16 \times 16$, 32×32 and 64×64 .

2.1.6 Comparing the speed of HMC and Metropolis

The samples produced by a Markov chain are correlated with each other. In general HMC generates less correlated samples if compared to Metropolis,

resulting in higher quality data for a given computational cost.

When computing a specific observable O , the efficiency of an algorithm in producing uncorrelated samples can be quantified by the *time-displaced autocorrelation*:

$$\chi(t) = \int dt' O(t') O(t' + t) - \langle O \rangle^2 \quad (28)$$

where $O(t)$ represents the value of the observable at Monte Carlo time t . The autocorrelation function is expected to fall off exponentially:

$$\chi(t) \sim e^{-\frac{t}{\tau}} \quad (29)$$

and τ is known as the autocorrelation time. Two values of the observable taken at $\sim 2\tau$ distance from each other are considered independent.

The autocorrelation time tends to be dramatically larger near second-order phase transitions. For this reason the $(2, 0)$ geometry at the phase transition was chosen to test HMC against Metropolis. The order parameter for the phase transition is:

$$F = \frac{(\text{Tr } H_1)^2 + (\text{Tr } H_2)^2}{n(\text{Tr } H_1^2 + \text{Tr } H_2^2)} \quad (30)$$

where H_i are the two free Hermitian matrices of the model. The plot of the time-displaced autocorrelation and the exponential fit for HMC and Metropolis are reported in Fig.(2). The exponential fit gives an autocorrelation time of ~ 250 sweeps for Metropolis and ~ 25 sweeps for HMC, meaning that in this example Metropolis needs to run 10 times longer than HMC for collecting the same amount of independent samples.

Of course Metropolis might be more than 10 times faster than HMC to collect a single sample, therefore the raw speed of the algorithm needs to be taken into account as well. In order to test this, the speed of HMC and Metropolis in generating 100 samples for the $(0, 3)$ geometry was recorder for matrix dimension 16×16 , 32×32 and 64×64 . The results are reported in Tab.(1). For small matrix size the two algorithms are comparable in speed, while for larger matrices HMC greatly outperforms Metropolis. Linearly fitting the data in logarithmic scale allows to estimate how the two algorithms scale with the matrix dimension, giving a computational cost of $O(n^{4.6})$ and $O(n^{2.6})$ for Metropolis and HMC respectively.

Autocorrelation of order parameter $(2,0)$ test: comparison Metropolis/HMC

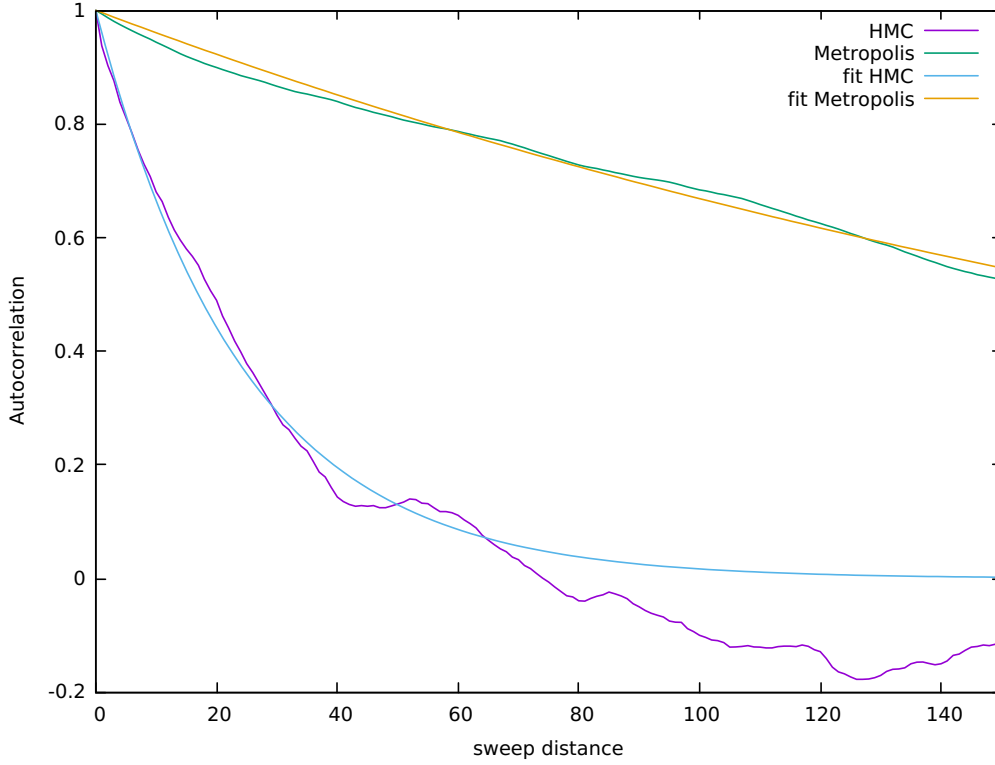


Figure 2: Autocorrelation $\chi(t)$ (normalized by $\chi(0)$) of the $(2,0)$ phase transition order parameter F . The coupling constant $g_2 = 2.725$ was chosen to be at the phase transition. The exponential fit gives an autocorrelation time of ~ 250 sweeps for Metropolis and ~ 25 sweeps for HMC.

Matrix dimension	time HMC (sec)	time Metropolis (sec)
16	23	17
32	100	428
64	1612	21218

Table 1: Time for 100 sweeps of HMC and Metropolis for geometry $(0,3)$. Fitting of the data gives a computational cost of $O(n^{4.6})$ and $O(n^{2.6})$ for Metropolis and HMC respectively.

2.2 Yang-Mills matrix models from Dirac operators

Yang-Mills matrix models have been studied extensively in the past [1] [12] [13] [20] as models of emergent geometry. Fuzzy spaces arise as classical solutions, and perturbations around them give rise to non-commutative gauge theories [22].

The purpose of this section is to show a formal similarity between Yang-Mills matrix models and a certain choice of Dirac operator, despite the very different origin of the two models.

2.2.1 Yang-Mills matrix models

A Yang-Mills matrix model is given by the following action:

$$S \propto -\frac{1}{4} \text{Tr}[M_i, M_j]^2 + \frac{2}{3} i g_3 \epsilon_{ijk} \text{Tr} M_i M_j M_k + \frac{1}{2} g_2 \text{Tr} M_i^2 \quad (31)$$

where $M_i, i = 1, 2, 3$, are $n \times n$ Hermitian matrices and a sum is intended on repeated indices. Actions without the quadratic or cubic term can also be considered. Note that in general one can consider models with more than three matrices (for example the IKKT model [18] has 10 matrices).

The variation of this action gives the following classical equation of motion for the generic matrix M_i :

$$[M_j, [M_j, M_i]] + i g_3 \epsilon_{ijk} [M_j, M_k] + g_2 M_i = 0 \quad (32)$$

where again sum on repeated indices is implied. Diagonal matrices satisfy Eq.(32), but a more interesting class of solutions is given by matrices forming an n -dimensional representation of the $su(2)$ Lie algebra:

$$[M_i, M_j] = i \epsilon_{ijk} M_k. \quad (33)$$

In this sense fuzzy spheres are classical solutions of (31).

2.2.2 Type (0, 3) geometry as a Yang-Mills matrix model

Consider the following Dirac operator:

$$D = \sum_i \sigma_i \otimes M_i \quad (34)$$

where $\sigma_i, i = 1, 2, 3$ are the Pauli matrices and M_i are Hermitian matrices. Note that no commutators or anti-commutators appear in the Dirac operator,

meaning that the bimodule structure of the algebra on itself is ignored and only the left (or right) action is taken into account. Eq.(34) is nonetheless an example of type (0,3) Dirac operator, although not the most general one. Consider then the usual action for D (with an additional cubic term and some convenient numerical factors):

$$S[D] = \frac{1}{4} \text{Tr } D^4 + \frac{1}{3} g_3 \text{Tr } D^3 + \frac{1}{4} g_2 \text{Tr } D^2. \quad (35)$$

It follows from the properties of the Pauli matrices that such action, rewritten explicitly in terms of the M_i matrices, takes a form very similar to (31):

$$\begin{aligned} S[D(M_i)] = & -\frac{1}{4} \text{Tr}[M_i, M_j]^2 - \frac{1}{2} \text{Tr } M_i^2 M_j^2 \\ & + \frac{2}{3} i g_3 \epsilon_{ijk} \text{Tr } M_i M_j M_k \\ & + \frac{1}{2} g_2 \text{Tr } M_i^2. \end{aligned} \quad (36)$$

The only difference is in fact an extra term $\text{Tr } M_i^2 M_j^2$ in the quartic part. The role of this term becomes clear by writing the equation of motion:

$$[M_j, [M_j, M_i]] - \{M_j^2, M_i\} + i g_3 \epsilon_{ijk} [M_j, M_k] + g_2 M_i = 0 \quad (37)$$

which still has the $su(2)$ algebra solution (33), provided that $\sum_j M_j^2$ is a Casimir. The extra term in the action for the Dirac operator therefore forces the representation to be either irreducible or direct sum of copies of the same irreducible.

It is interesting to see what happens if the commutator structure is reintroduced, so that the Dirac operator is now:

$$D = \sum_i \sigma_i \otimes [M_i, \cdot] \quad (38)$$

which is formally equivalent to replacing M_i with $[M_i, \cdot]$ in each formula so far. First notice that making this replacement in the equation of motion for M_i and ignoring the extra quartic term $\{[M_j, \cdot]^2, [M_i, \cdot]\}$ makes the $su(2)$ algebra still a viable classical solution:

$$[M_i, M_j] = i \epsilon_{ijk} M_k \implies [[M_i, \cdot], [M_j, \cdot]] = [[M_i, M_j], \cdot] = i \epsilon_{ijk} [M_k, \cdot] \quad (39)$$

however, the presence the extra quartic term now imposes the much stricter constraint of the representation to be the two-dimensional one (or direct sum

of copies thereof). One way to see this is to expand the commutators as in (64) and calculate the contribution of the extra term explicitly:

$$\begin{aligned} \{[M_j, \cdot]^2, [M_i, \cdot]\} = & \{\{M_j^2, \cdot\}, [M_i, \cdot]\} \\ & - 2\{M_i, M_j\} \otimes M_j^T + 2M_j \otimes \{M_i^T, M_j^T\}. \end{aligned} \quad (40)$$

The first term on the left-hand side of Eq.(40) is proportional to $[M_i, \cdot]$ when $\sum_j M_j^2$ is a Casimir, but the remaining two terms only combine to be proportional to $[M_i, \cdot]$ when $\{M_i, M_j\} \propto \delta_{ij}$, which is satisfied by the 2×2 representation of $su(2)$ but not by higher dimensional ones.

2.3 Reduced model for the $(2, 0)$ fuzzy Dirac operator

The purpose of this section is to apply the tools of random matrix theory to the study of the simplest non-trivial model, namely the type $(2, 0)$ Dirac operator. The general idea is to identify degrees of freedom that can be integrated out, and to write a simpler model in terms of the remaining ones. First a more convenient notation for the model is introduced, then the classical solutions are found and finally the simplified model is obtained by considering small perturbations of these.

This project is work in progress with Professor Sven Gnutzmann.

2.3.1 Notation

Throughout this section a more convenient notation is used in order to properly identify the relevant degrees of freedom.

Recall that a type $(2, 0)$ Dirac operator is written as:

$$D = \gamma_1 \otimes \{H_1, \cdot\} + \gamma_2 \otimes \{H_2, \cdot\} \quad (41)$$

for Hermitian matrices H_1 and H_2 . Define $\gamma_{\pm} := \frac{1}{2}(\gamma_1 \mp i\gamma_2)$ and $W := H_1 + iH_2$ and rewrite D in these terms:

$$D = \gamma_+ \otimes \{W, \cdot\} + \gamma_- \otimes \{W^\dagger, \cdot\}. \quad (42)$$

The quadratic and quartic terms in the action read:

$$\text{Tr } D^2 = 4(n \text{Tr } WW^\dagger + \text{Tr } W \text{Tr } W^\dagger) \quad (43)$$

$$\begin{aligned} \text{Tr } D^4 = 4[& n \text{Tr}(WW^\dagger)^2 + \text{Tr } W^2 \text{Tr } W^{\dagger 2} + 2(\text{Tr } WW^\dagger)^2 \\ & + 2 \text{Tr } W^2 W^\dagger \text{Tr } W^\dagger + 2 \text{Tr } W^{\dagger 2} W \text{Tr } W]. \end{aligned} \quad (44)$$

2.3.2 Stationary manifold

Stationary points are solutions to the equations of motion:

$$\delta S = S[W + \delta W] - S[W] = 0 \quad \forall \delta W \ll 1. \quad (45)$$

To first order in δW and δW^\dagger the terms $\text{Tr } D^2$ and $\text{Tr } D^4$ give:

$$\delta \text{Tr } D^2 = 4 \text{Tr} \left\{ [1 + \dagger][nW + (\text{Tr } W)I]\delta W^\dagger \right\} \quad (46)$$

$$\begin{aligned} \delta \text{Tr } D^4 = 4 \text{Tr} \left\{ [1 + \dagger][& 2nWW^\dagger W + (2 \text{Tr } W^2)W^\dagger + (4 \text{Tr } WW^\dagger)W + \right. \\ & (2 \text{Tr } W^2 W^\dagger)I + (2 \text{Tr } W^\dagger)W^2 + \\ & \left. (2 \text{Tr } W)(W^\dagger W + WW^\dagger)]\delta W^\dagger \right\}. \end{aligned} \quad (47)$$

Solutions are found by putting the coefficient of δW or δW^\dagger to zero. They are of the form:

$$W = I\rho e^{i\theta} \quad (48)$$

where

$$\rho^2 = 0 \quad \text{or} \quad \rho^2 = -\frac{g}{8} \quad (49)$$

and θ is unconstrained.

2.3.3 Small oscillations and quadratic action

Write $W = e^{i\theta}(\rho + \epsilon V)$ for small ϵ and traceless V . Order by order in ϵ , the contributions to the action are:

$$\begin{aligned} \text{Tr } D^2 \\ O(1) : \quad 8n^2\rho^2 \end{aligned} \quad (50)$$

$$O(\epsilon) : \quad 8n\rho(\text{Tr } V^\dagger + \text{Tr } V) \quad (51)$$

$$O(\epsilon^2) : \quad 4(n \text{Tr } VV^\dagger + \text{Tr } V \text{Tr } V^\dagger) \quad (52)$$

$$(53)$$

$$\begin{aligned} \text{Tr } D^4 \\ O(1) : \quad 32n^2\rho^4 \end{aligned} \quad (54)$$

$$O(\epsilon) : \quad 48n\rho^3(\text{Tr } V^\dagger + \text{Tr } V) \quad (55)$$

$$O(\epsilon^2) : \quad 16\rho^2(2n \text{Tr } VV^\dagger + n \text{Tr } VV \quad (56)$$

$$+ 2 \text{Tr } V \text{Tr } V^\dagger + \text{Tr } V \text{Tr } V + \text{c.c.})$$

Since V is traceless, the action to second order in ϵ reads:

$$S = 8n^2(g\rho^2 + 4\rho^4) + 2n\epsilon^2 \text{Tr}(\vec{V}^T M \vec{V}) \quad (57)$$

with $\vec{V}^T = (V, V^\dagger)$ and:

$$M = \begin{bmatrix} 8\rho^2 & g + 16\rho^2 \\ g + 16\rho^2 & 8\rho^2 \end{bmatrix}. \quad (58)$$

At this point the idea would be to perform the gaussian integral in V and remain with an effective theory in ρ . As it turns out, the contribution of V cannot be integrated out due to the fact that M has a zero eigenvalue on the stationary solutions. This signals the presence of modes that are not confined by a quadratic potential (*massless zero modes*). The issue can be made explicit by decomposing $V = A + iB$ for Hermitian and traceless A and B . In these variables the action (57) reads:

$$S = 8n^2(g\rho^2 + 4\rho^4) + 4n\epsilon^2 \left[(g + 24\rho^2) \text{Tr } A^2 + (g + 8\rho^2) \text{Tr } B^2 \right] \quad (59)$$

where the coefficient of the B part vanishes on the stationary solutions $\rho^2 = -g/8$. This suggests that one needs to go to higher orders in perturbation theory to find a confining potential for B . Numerical simulations confirm this picture (see Fig.(3)).

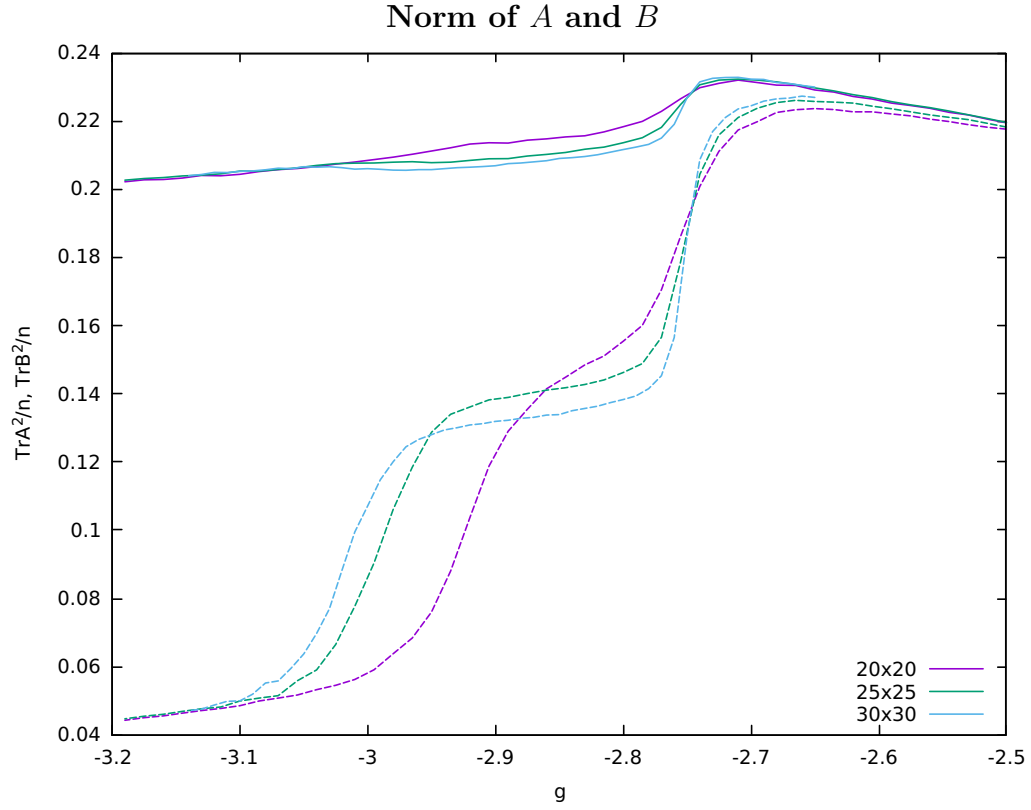


Figure 3: $\text{Tr } A^2/n$ (dotted lines) and $\text{Tr } B^2/n$ (solid lines) versus coupling constant g for $n = 20$ (purple), 25 (green), 30 (blue) at the phase transition.

3 Future directions

3.1 Including matter fields

The matrix model analyzed so far should ideally describe a non-commutative analogue of a space-time manifold in a finite theory of quantum gravity. A number of results are being collected in support of this claim, the latest being spectral estimators for dimension and volume [5] and still unpublished results on the Lie algebraic structures that emerge among the matrices of certain models.

A rather unsatisfactory aspect of the model is that no matter fields have entered the picture yet. A priori there is no guarantee that the model without matter should even be of physical relevance, and the presence of matter could produce dramatic deviations from the behaviour observed so far.

The proposal for the introduction of matter fields follows the spirit of Sakharov's work on induced gravity [21], and represents the final step in a logical progression that started in the early days of non-commutative geometry applied to the Standard Model. In [10] Connes showed that there is no need to introduce gauge fields explicitly in the spectral action, as the bosonic sector of the Standard Model emerges naturally from inner fluctuations of the Dirac operator. Schematically, the spectral action Connes considered had the following form:

$$S = \text{Tr } \chi \left(\frac{D}{\Lambda} \right) + \langle \Psi, D\Psi \rangle \quad (60)$$

where the first term would produce the Einstein-Hilbert action plus the bosonic and Higgs sector, and the second would produce the couplings between gauge fields and matter. Later on, in [2] it was observed that Sakharov's mechanism could take care of the whole bosonic sector (gauge fields plus gravity), and a more minimalistic spectral action was proposed:

$$S = \langle \Psi, D\Psi \rangle. \quad (61)$$

At this point space-time is still a commutative classical manifold M . By analogy, the natural way to introduce matter in the context of fuzzy spaces is to replace the manifold with a fuzzy space while keeping the same action (61). Grassmann integration over the fermions yields a determinant, therefore a convergency term in the action of the type $\text{Tr } D^{2p}$ is needed as well. The simplest choice would then be:

$$S = g_2 \text{Tr } D^2 - \frac{1}{2} \text{Tr } \log D^2 \quad (62)$$

where the determinant has been expressed as a trace log.

The inclusion of fermions entails both a theoretical and a numerical aspect:

1. on the theoretical side, one might carry out explicitly the calculation of the action and identify the gravity, gauge and matter sectors of the theory;
2. on the numerical side, a Monte Carlo simulation with a fermion determinant can be performed in various ways, the state of the art being an extension of HMC called Rational Hybrid Monte Carlo [17].

Appendix A: Notation

1. To avoid dealing with both Hermitian and anti-Hermitian matrices in Eq.(1), it is convenient to redefine $i\tilde{L}_j \equiv L_j$ and $\tilde{\alpha}_j \equiv i\alpha_j$. Eq.(1) becomes:

$$D = \sum_j \tilde{\alpha}_j \otimes [\tilde{L}_j, \cdot] + \sum_k \tau_k \otimes \{H_k, \cdot\} \quad (63)$$

and all the matrices are Hermitian.

2. Commutators and anti-commutators are represented in matrix form as:

$$[A, \cdot] = A \otimes I - I \otimes A^T$$

$$\{A, \cdot\} = A \otimes I + I \otimes A^T.$$

As a shorthand, the following notation will be used:

$$[A, \cdot]_\epsilon \equiv A \otimes I + \epsilon I \otimes A^T. \quad (64)$$

The final form of Eq.(1) is then:

$$D = \sum_{i \in I} \omega_i \otimes [M_i, \cdot]_{\epsilon_i} \quad (65)$$

for Hermitian matrices M_i , with $\omega_i \in \{\tilde{\alpha}_j\} \cup \{\tau_k\}$, and $\epsilon_i = \pm 1$ depending on ω_i being a τ matrix or a $\tilde{\alpha}$ matrix.

Appendix B: Matrix derivatives

Let $A \in M_n(\mathbb{C})$ and $f(A)$ be a complex valued function of A . The derivative of f with respect to A is defined in components as the $n \times n$ matrix:

$$\left(\frac{\partial f}{\partial A} \right)_{lm} \equiv \frac{\partial f}{\partial A_{lm}}. \quad (66)$$

The two special cases of interest here are:

$$\frac{\partial \text{Tr } A}{\partial A} = I \quad (67)$$

$$\frac{\partial \text{Tr } AB}{\partial A} = B^T. \quad (68)$$

Appendix C

The explicit form of $\mathcal{B}_k(i, i, k)$, $\mathcal{B}_k(i, k, i)$ and $\mathcal{B}_k(k, k, k)$ is given.

$$\begin{aligned}\mathcal{B}_k(i, i, k) &= \text{Tr}(\omega_k \omega_i \omega_i \omega_k) A(k, i, i, k)^T = C A(k, i, i, k)^T \\ \mathcal{B}_k(i, k, i) &= \text{Tr}(\omega_k \omega_i \omega_k \omega_i) A(k, i, k, i)^T \\ \mathcal{B}_k(k, k, k) &= \text{Tr}(\omega_k \omega_k \omega_k \omega_k) A(k, k, k, k)^T = C A(k, k, k, k)^T\end{aligned}$$

where C is the dimension of the Clifford module and the A matrices are:

$$\begin{aligned}A(k, i, i, k) &= n[1 + \dagger] M_i^2 M_k + \\ &\quad 2\epsilon_k I \text{Tr} M_i^2 M_k + \\ &\quad 2\epsilon_i \text{Tr} M_i [1 + \dagger] M_i M_k + \\ &\quad 4\epsilon_k \epsilon_i M_i \text{Tr} M_i M_k + \\ &\quad 2\epsilon_k M_i^2 \text{Tr} M_k + \\ &\quad 2M_k \text{Tr} M_i^2\end{aligned}\tag{69}$$

$$\begin{aligned}A(k, i, k, i) &= 2n M_i M_k M_i + \\ &\quad 2\epsilon_k I \text{Tr} M_i^2 M_k + \\ &\quad 2\epsilon_i \text{Tr} M_i [1 + \dagger] M_i M_k + \\ &\quad 4\epsilon_k \epsilon_i M_i \text{Tr} M_i M_k + \\ &\quad 2\epsilon_k M_i^2 \text{Tr} M_k + \\ &\quad 2M_k \text{Tr} M_i^2\end{aligned}\tag{70}$$

$$\begin{aligned}A(k, k, k, k) &= 2n M_k^3 + 2\epsilon_k I \text{Tr} M_k^3 + \\ &\quad 6M_k \text{Tr} M_k^2 + 6\epsilon_k M_k^2 \text{Tr} M_k.\end{aligned}\tag{71}$$

Appendix D

Suppose D is a $n \times n$ matrix in which m entries are linearly independent. An arbitrary entry can then be written as:

$$D_{ab} = \sum_{ij} c_{ab}^{ij} D_{ij}\tag{72}$$

where c_{ab}^{ij} are coefficients and the sum runs over the independent components. This is the case for fuzzy Dirac operators.

The following identity will be proven:

$$\sum_{ij} D_{ij} \frac{\partial}{\partial D_{ij}} \text{Tr} D^p = p \text{Tr} D^p.\tag{73}$$

An explicit calculation gives:

$$\begin{aligned}
\sum_{ij} D_{ij} \frac{\partial}{\partial D_{ij}} \text{Tr } D^p &= \sum_{ij} D_{ij} \frac{\partial}{\partial D_{ij}} \sum_{a_1 \cdots a_p} D_{a_1 a_2} \cdots D_{a_p a_1} = \\
&= \sum_{ij} D_{ij} \sum_{a_1 \cdots a_p} \left(c_{a_1 a_2}^{ij} D_{a_2 a_3} \cdots D_{a_p a_1} + \cdots + D_{a_1 a_2} \cdots D_{a_{p-1} a_p} c_{a_p a_1}^{ij} \right) = \\
&= \sum_{a_1 \cdots a_p} \left(\left(\sum_{ij} c_{a_1 a_2}^{ij} D_{ij} \right) D_{a_2 a_3} \cdots D_{a_p a_1} + \cdots + D_{a_1 a_2} \cdots D_{a_{p-1} a_p} \left(\sum_{ij} c_{a_p a_1}^{ij} D_{ij} \right) \right) = \\
&= p \sum_{a_1 \cdots a_p} D_{a_1 a_2} \cdots D_{a_p a_1} = p \text{Tr } D^p \tag{74}
\end{aligned}$$

References

- [1] T. Azuma, S. Bal, K. Nagao and J. Nishimura, “Nonperturbative studies of fuzzy spheres in a matrix model with the Chern-Simons term”, *JHEP* **0405**, 005 (2004).
- [2] J. W. Barrett, “State sum models, induced gravity and the spectral action”, arXiv:1101.6078 [hep-th]
- [3] J. W. Barrett, “Lorentzian version of the noncommutative geometry of the standard model of particle physics”, *J. Math. Phys.* **48**, 012303 (2007).
- [4] J. W. Barrett, “Matrix geometries and fuzzy spaces as finite spectral triples”, *J. Math. Phys.* **56**, 082301 (2015).
- [5] J. W. Barrett, P. Druce and L. Glaser, “Spectral estimators for finite non-commutative geometries”, *J. Phys. A* **52**, 275203 (2019).
- [6] J. W. Barrett and L. Glaser, “Monte Carlo simulations of random non-commutative geometries”, *J. Phys. A* **49**, 24, 245001 (2016).
- [7] M. Betancourt, “A Conceptual Introduction to Hamiltonian Monte Carlo”, (arXiv:1701.02434v1, 2017).
- [8] A. H. Chamseddine and A. Connes, “The spectral action principle”, *Commun. Math. Phys.* **186**, 731–750 (1997).
- [9] A. H. Chamseddine, A. Connes and M. Marcolli, “Gravity and the standard model with neutrino mixing”, *Adv. Theor. Math. Phys.* **11**, 991–1089 (2007).
- [10] A. Connes, “Gravity coupled with matter and the foundation of noncommutative geometry”, *Commun. Math. Phys.* **183**, 1, 155–176 (1996).
- [11] A. Connes, “On the spectral characterization of manifolds”, *J. Noncommut. Geom.* **7**, 1–82 (2013).
- [12] R. Delgadillo-Blando, D. OConnor and B. Ydri, “Geometry in transition: A model of emergent geometry”, *Phys. Rev. Lett.* **100**, 201601 (2008).
- [13] R. Delgadillo-Blando, D. OConnor and B. Ydri, “Matrix Models, Gauge Theory and Emergent Geometry”, *JHEP* **0905**, 049 (2009).

- [14] S. Duane, A. D. Kennedy, B. J. Pendleton and D. Roweth, “Hybrid Monte Carlo”, *Phys. Lett. B* **195**, 216–222 (1987).
- [15] L. Glaser, “Scaling behaviour in random non-commutative geometries”, *J. Phys. A* **50**, 27, 275201 (2017).
- [16] W. K. Hastings, “Monte Carlo sampling methods using Markov chains and their applications”, *Biometrika* **57**, 1, 97–109 (1970).
- [17] I. Horváth, A. d. Kennedy and S. Sint, “A new exact method for dynamical fermion computations with non-local actions”, *Nucl. Phys. B Proc. Suppl.* **73**, 834–836 (1999)
- [18] N. Ishibashi, H. Kawai, Y. Kitazawa and A. Tsuchiya, “A large-N reduced model as Superstring”, *Nucl. Phys. B* **498**, 467 (1997).
- [19] R. M. Neal, *Handbook of Markov Chain Monte Carlo* (Chapman & Hall / CRC Press, 2011).
- [20] D. OConnor and B. Ydri, “Monte Carlo Simulation of a NC Gauge Theory on The Fuzzy Sphere”, *JHEP* **0611**, 016 (2006).
- [21] A. D. Sakharov, “Vacuum quantum fluctuations in curved space and the theory of gravitation”, *Sov. Phys. Usp.* **34**, 394 (1991).
- [22] H. Steinacker, “Emergent geometry and gravity from matrix models: an introduction”, *Class. Quant. Grav.* **27**, 133001 (2010).
- [23] W. D. Suijlekom, *Noncommutative Geometry and Particle Physics* (Springer, 2014).