1 Hamiltonian Monte Carlo for fuzzy spaces

1.1 Statement of the problem

The fuzzy space action considered here is:

$$S[D] = g_2 \operatorname{Tr} D^2 + \operatorname{Tr} D^4 \tag{1}$$

where $g_2 \in \mathbb{R}$ and D is of the form:

$$D = \sum_i \omega_i \otimes (M_i \otimes I + \epsilon_i I \otimes M_i^T) \tag{2}$$

for Hermitian ω_i and M_i, and $\epsilon_i = \pm 1$.
The dynamical variables in the Monte Carlo are the $n \times n$ matrices M_i. Hamiltonian Monte Carlo requires to take derivatives such as:

$$\frac{\partial S[M_i]}{\partial M_k} \tag{3}$$

which amounts to finding formulas for terms like:

$$\frac{\partial \operatorname{Tr} D^p}{\partial M_k}. \tag{4}$$

In the following, formulas for $p = 2$ and $p = 4$ are developed.

1.2 Matrix calculus

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:

$$\left(\frac{\partial f}{\partial A}\right)_{lm} \equiv \frac{\partial f}{\partial A_{lm}}. \tag{5}$$

The two special cases of interest here are:

$$\frac{\partial \operatorname{Tr} A}{\partial A} = I \tag{6}$$

$$\frac{\partial \operatorname{Tr} AB}{\partial A} = B^T. \tag{7}$$

1.3 The case $p = 2$

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

$$\operatorname{Tr} D^2 = \sum_i \operatorname{Tr} \omega_i^2 (2n \operatorname{Tr} M_i^2 + 2\epsilon_i (\operatorname{Tr} M_i)^2). \tag{8}$$

Taking a derivative with respect to M_k yields:

$$\begin{aligned} \frac{\partial}{\partial M_k} \left(\sum_i \operatorname{Tr} \omega_i^2 (2n \operatorname{Tr} M_i^2 + 2\epsilon_i (\operatorname{Tr} M_i)^2) \right) &= \\ \sum_i \delta_{ik} \operatorname{Tr} \omega_i^2 \left(4n M_i^T + 4\epsilon_i (\operatorname{Tr} M_i) I \right) &= \\ 4C \left(n M_k^T + \epsilon_k (\operatorname{Tr} M_k) I \right) \end{aligned} \tag{9}$$

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

1.4 The case $p = 4$

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

$$\begin{aligned} \operatorname{Tr} D^4 = \sum_{i_1, i_2, i_3, i_4} & \operatorname{Tr}(\omega_{i_1} \omega_{i_2} \omega_{i_3} \omega_{i_4}) \cdot \\ \Big(& n[1 + \epsilon\, *] \operatorname{Tr}(M_{i_1} M_{i_2} M_{i_3} M_{i_4}) \ + \\ & \epsilon_{i_1} \operatorname{Tr} M_{i_1} [1 + \epsilon\, *] \operatorname{Tr}(M_{i_2} M_{i_3} M_{i_4}) \ + \\ & \epsilon_{i_2} \operatorname{Tr} M_{i_2} [1 + \epsilon\, *] \operatorname{Tr}(M_{i_1} M_{i_3} M_{i_4}) \ + \\ & \epsilon_{i_3} \operatorname{Tr} M_{i_3} [1 + \epsilon\, *] \operatorname{Tr}(M_{i_1} M_{i_2} M_{i_4}) \ + \\ & \epsilon_{i_4} \operatorname{Tr} M_{i_4} [1 + \epsilon\, *] \operatorname{Tr}(M_{i_1} M_{i_2} M_{i_3}) \ + \\ & \epsilon_{i_1} \epsilon_{i_2} [1 + \epsilon] \operatorname{Tr}(M_{i_1} M_{i_2}) \operatorname{Tr}(M_{i_3} M_{i_4}) \ + \\ & \epsilon_{i_1} \epsilon_{i_3} [1 + \epsilon] \operatorname{Tr}(M_{i_1} M_{i_3}) \operatorname{Tr}(M_{i_2} M_{i_4}) \ + \\ & \epsilon_{i_1} \epsilon_{i_4} [1 + \epsilon] \operatorname{Tr}(M_{i_1} M_{i_4}) \operatorname{Tr}(M_{i_2} M_{i_3}) \ \Big) \end{aligned} \tag{10}$$

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

$$\begin{aligned}\frac{\partial}{\partial M_k}\operatorname{Tr} D^4 = \sum_{i_1,i_2,i_3,i_4} & \operatorname{Tr}(\omega_{i_1}\omega_{i_2}\omega_{i_3}\omega_{i_4})\cdot \\ & \Big(\delta_{ki_1} A(i_1,i_2,i_3,i_4)^T + \delta_{ki_2} A(i_2,i_3,i_4,i_1)^T + \\ & \delta_{ki_3} A(i_3,i_4,i_1,i_2)^T + \delta_{ki_4} A(i_4,i_1,i_2,i_3)^T\Big)\end{aligned} \tag{11}$$

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

$$\begin{aligned}A(a,b,c,d) \equiv\ & n[1+\epsilon\ \dagger] M_b M_c M_d + \\ & \epsilon_a I [1+\epsilon\ *] \operatorname{Tr} M_b M_c M_d + \\ & \epsilon_b \operatorname{Tr} M_b [1+\epsilon\ \dagger] M_c M_d + \\ & \epsilon_c \operatorname{Tr} M_c [1+\epsilon\ \dagger] M_b M_d + \\ & \epsilon_d \operatorname{Tr} M_d [1+\epsilon\ \dagger] M_b M_c + \\ & \epsilon_a \epsilon_b M_b [1+\epsilon] \operatorname{Tr} M_c M_d + \\ & \epsilon_a \epsilon_c M_c [1+\epsilon] \operatorname{Tr} M_b M_d + \\ & \epsilon_a \epsilon_d M_d [1+\epsilon] \operatorname{Tr} M_b M_c\end{aligned} \tag{12}$$

and $\dagger$ denotes Hermitian conjugation of everything that appears on the right. Again, to see that Eq.(11) defines a Hermitian matrix, it is useful to rewrite it symbolically as:

$$\frac{\partial}{\partial M_k}\operatorname{Tr} D^4 = \sum_{i_1,i_2,i_3,i_4} \mathcal{B}(i_1,i_2,i_3,i_4) \tag{13}$$

and notice that a simultaneous exchange of indices $i_1 \leftrightarrow i_4$ and $i_2 \leftrightarrow i_3$ is equivalent to taking the Hermitian conjugate:

$$\mathcal{B}(i_1,i_2,i_3,i_4)^\dagger = \mathcal{B}(i_4,i_3,i_2,i_1). \tag{14}$$

In fact, if S_4 is the symmetric group of order four acting on the set of indices $\{i_1, i_2, i_3, i_4\}$, there is a larger subgroup of S_4 whose action produces a simple change in $\mathcal{B}(i_1, i_2, i_3, i_4)$. Namely, the Klein four-group, a normal subgroup of S_4, which in terms of permutations reads:

$$K_4 = \{(), (13)(24), (12)(34), (14)(23)\} \equiv \{\iota, \rho, \sigma, \tau\}. \tag{15}$$

The change in $\mathcal{B}$ induced by the action of K_4 is the following:

$$\mathcal{B}(i_{\rho(1)}, i_{\rho(2)}, i_{\rho(3)}, i_{\rho(4)}) = \mathcal{B}(i_1, i_2, i_3, i_4) \tag{16}$$

$$\mathcal{B}(i_{\sigma(1)}, i_{\sigma(2)}, i_{\sigma(3)}, i_{\sigma(4)}) = \mathcal{B}(i_1, i_2, i_3, i_4)^\dagger \tag{17}$$

$$\mathcal{B}(i_{\tau(1)}, i_{\tau(2)}, i_{\tau(3)}, i_{\tau(4)}) = \mathcal{B}(i_1, i_2, i_3, i_4)^\dagger \tag{18}$$

This property can be exploited to gain some computational efficiency.
Start by considering the terms in Eq.(13) in which all indices are different from each other:

$$\sum_{\substack{i_j \neq i_k \\ j,k=1,2,3,4}} \mathcal{B}(i_1, i_2, i_3, i_4). \tag{19}$$

Many terms in this sum are the same (possibly up to Hermitian conjugation) due to the action of K_4. Therefore the sum can be written in the following way:

$$\sum_{i_1>i_2>i_3>i_4} 2[1+\dagger]\Big\{\mathcal{B}(i_1, i_2, i_3, i_4) + \mathcal{B}(i_3, i_1, i_2, i_4) + \mathcal{B}(i_4, i_1, i_3, i_2) + \\ \mathcal{B}(i_2, i_1, i_3, i_4) + \mathcal{B}(i_3, i_2, i_1, i_4) + \mathcal{B}(i_4, i_2, i_3, i_1)\Big\} \tag{20}$$

where the sum is now restricted to $i_1 > i_2 > i_3 > i_4$ and the six terms inside curly braces are obtained by acting on $\{i_1, i_2, i_3, i_4\}$ with a representative from the cosets of $K_4 \subset S_4$. This form minimizes the number of calculations. Next, One can introduce an equivalence relation

$$(i_1, i_2, i_3, i_4) \sim (i_{\rho^m(1)}, i_{\rho^m(2)}, i_{\rho^m(3)}, i_{\rho^m(4)}), \quad m \in \mathbb{N} \tag{21}$$

and restrict the sum on the equivalence classes:

$$\frac{\partial}{\partial M_k} \operatorname{Tr} D^4 = \sum_{[i_1,i_2,i_3,i_4]} \frac{|[i_1, i_2, i_3, i_4]|}{2} [1 + \dagger] \mathcal{B}(i_1, i_2, i_3, i_4) \tag{22}$$

where $|[i_1, i_2, i_3, i_4]|$ is the cardinality of the class. Practically, ρ^{2m} is the identity permutation, and the equivalence classes have one or two elements. The factor involving the cardinality of the class prevents from overcounting terms of the type (a, b, b, a) (which are the only elements with class of cardinality one).
The sum of Eq.(22) can be written explicitly in the following way:

$$\begin{aligned}
\frac{\partial}{\partial M_k} \operatorname{Tr} D^4 = \\
&\sum_{\substack{i_4<i_1 \\ i_3<i_2}} [1 + \dagger] \big(\mathcal{B}(i_1, i_2, i_3, i_4) + \mathcal{B}(i_1, i_3, i_2, i_4)\big) + \\
&\sum_{\substack{i_4=i_1 \\ i_3<i_2}} [1 + \dagger] \mathcal{B}(i_1, i_2, i_3, i_4) + \sum_{\substack{i_4<i_1 \\ i_3=i_2}} [1 + \dagger] \mathcal{B}(i_1, i_2, i_3, i_4) + \\
&\sum_{\substack{i_4=i_1 \\ i_3=i_2}} \mathcal{B}(i_1, i_2, i_3, i_4)
\end{aligned} \tag{23}$$

and simplified noticing that the second line vanishes due to the properties of the ω matrices:

$$\operatorname{Tr}(\omega_{\sigma(i_1)} \omega_{\sigma(i_2)} \omega_{\sigma(j)} \omega_{\sigma(k)}) \propto \operatorname{Tr}(\omega_j \omega_k) = 0 \qquad \text{if} \;\; i_1 = i_2 \text{ and } j \neq k \tag{24}$$

for any permutation σ acting on $\{i_1, i_2, j, k\}$.
Therefore Eq.(23) simplifies to:

$$\frac{\partial}{\partial M_k} \operatorname{Tr} D^4 = \sum_{\substack{i_4 < i_1 \\ i_3 < i_2}} [1 + \dagger]\big(\mathcal{B}(i_1, i_2, i_3, i_4) + \mathcal{B}(i_1, i_3, i_2, i_4)\big) + \sum_{i_1, i_2} \mathcal{B}(i_1, i_2, i_2, i_1). \tag{25}$$

For the same reason, terms in the first sum for which ($i_1 = i_2$ and $i_3 \neq i_4$) or ($i_1 = i_3$ and $i_2 \neq i_4$) need not be computed as they vanish.

1.5 A general formula for every p

The first problem is to write $\operatorname{Tr} D^p$ in a useful form, along the lines of Eq.(10). $\operatorname{Tr} D^p$ expands to:

$$\operatorname{Tr} D^p = \sum_{i_1 \dots i_p} \operatorname{Tr} \omega_{i_1} \dots \omega_{i_p}.$$

$$\operatorname{Tr} \big((M_{i_1} \otimes I + \epsilon_{i_1} I \otimes M_{i_1}^T) \dots (M_{i_p} \otimes I + \epsilon_{i_p} I \otimes M_{i_p}^T)\big) \tag{26}$$

Ignoring (for now) the trace over the ω matrices, a typical term in the sum is:

$$\operatorname{Tr} \big(\epsilon_B A \otimes B^* + \epsilon_A B \otimes A^*\big) \tag{27}$$

where A and B are related to the product $M_{i_1} \dots M_{i_p}$ in the following way:

1. pick $r \geq 0$ numbers $k_1 < \dots < k_r$ from $\{1, \dots, p\}$ and call the remaining $p - r$ numbers $j_1 < \dots < j_{p-r}$;

2. define $A = M_{i_{k_1}} \dots M_{i_{k_r}}$ and $B = M_{i_{j_1}} \dots M_{i_{j_{p-r}}}$ (if $r = 0$, $A = I$);

3. define $\epsilon_A = \epsilon_{i_{k_1}} \dots \epsilon_{i_{k_r}}$ and $\epsilon_B = \epsilon_{i_{j_1}} \dots \epsilon_{i_{j_{p-r}}} = \epsilon_A \epsilon_{i_1} \dots \epsilon_{i_p}$.

In particular, a choice of A completely characterizes B.
By varying r from 0 to $\left[\frac{p}{2}\right]$ and summing over all possible choices of $k_1 \dots k_r$, every term in $\operatorname{Tr} D^p$ is generated.

One can verify that every term in Eq.(10) ($p = 4$) is of that type. For example:

$$
\begin{aligned}
&\operatorname{Tr} M_{i_1}[\epsilon_{i_1} + \epsilon_{i_2}\epsilon_{i_3}\epsilon_{i_4} *] \operatorname{Tr}(M_{i_2}M_{i_3}M_{i_4}) = \\
&\qquad \epsilon_{i_2}\epsilon_{i_3}\epsilon_{i_4} \operatorname{Tr} M_{i_1} \operatorname{Tr}(M_{i_2}M_{i_3}M_{i_4})^* + \epsilon_{i_1} \operatorname{Tr} M_{i_1} \operatorname{Tr}(M_{i_2}M_{i_3}M_{i_4}) = \\
&\qquad \epsilon_{i_2}\epsilon_{i_3}\epsilon_{i_4} \operatorname{Tr} M_{i_1} \operatorname{Tr}(M_{i_2}M_{i_3}M_{i_4})^* + \epsilon_{i_1} \operatorname{Tr}(M_{i_1})^* \operatorname{Tr}(M_{i_2}M_{i_3}M_{i_4}) = \\
&\qquad \operatorname{Tr}\big(\epsilon_{i_2}\epsilon_{i_3}\epsilon_{i_4} M_{i_1} \otimes (M_{i_2}M_{i_3}M_{i_4})^* + \epsilon_{i_1} M_{i_2}M_{i_3}M_{i_4} \otimes M_{i_1}^*\big)
\end{aligned} \tag{28}
$$

which is of the form of Eq.(27) upon identifying M_{i_1} with A and $M_{i_2}M_{i_3}M_{i_4}$ with B (in the second equality the reality of $\operatorname{Tr} M_{i_1}$ has been used).
A way of expressing $\operatorname{Tr} B$ given A is using a modified derivative operator D_i defined as:

$$
\mathrm{D}_i \equiv \operatorname{Tr} \circ \frac{\partial}{\partial M_i} \tag{29}
$$

which allows to write:

$$
A = M_{i_{k_1}} \dots M_{i_{k_r}} \implies \operatorname{Tr} B = \mathrm{D}_{i_{k_r}} \dots \mathrm{D}_{i_{k_1}} \operatorname{Tr}(M_{i_1} \dots M_{i_p}). \tag{30}
$$

Therefore Eq.(27) becomes:

$$
\begin{aligned}
&\epsilon_A[1 + \epsilon_{i_1} \dots \epsilon_{i_p} *](\operatorname{Tr} A)^* \operatorname{Tr} B = \\
&\quad \epsilon_{i_{k_1}} \dots \epsilon_{i_{k_r}}[1 + \epsilon_{i_1} \dots \epsilon_{i_p} *](\operatorname{Tr} M_{i_{k_1}} \dots M_{i_{k_r}})^* \big(\mathrm{D}_{i_{k_r}} \dots \mathrm{D}_{i_{k_1}} \operatorname{Tr}(M_{i_1} \dots M_{i_p})\big).
\end{aligned} \tag{31}
$$

There are some special cases that make the expression simpler, namely:

1. $r = 0$ gives a factor $\operatorname{Tr} I = n$;
2. $r = 1, 2$ make $\operatorname{Tr} A$ real;
3. $p - r = 1, 2$ (which can only occur for $p = 2, 4$) make $\operatorname{Tr} B$ real.

Putting everything together, $\operatorname{Tr} D^p$ can be written as:

$$
\begin{aligned}
\operatorname{Tr} D^p = \sum_{i_1 \dots i_p} \operatorname{Tr} \omega_{i_1} \dots \omega_{i_p} \Bigg[&\sum_{r=0}^{[\frac{p}{2}]} \sum_{k_1 < \dots < k_r = 1}^{p} \epsilon_{i_{k_1}} \dots \epsilon_{i_{k_r}} [1 + \epsilon_{i_1} \dots \epsilon_{i_p} *] \\
&(\operatorname{Tr} M_{i_{k_1}} \dots M_{i_{k_r}})^* \big(\mathrm{D}_{i_{k_r}} \dots \mathrm{D}_{i_{k_1}} \operatorname{Tr}(M_{i_1} \dots M_{i_p})\big) \Bigg].
\end{aligned} \tag{32}
$$

where:

$$r = 0 \quad \longrightarrow \quad n[1 + \epsilon_{i_1} \ldots \epsilon_{i_p} *] \, \mathrm{Tr}(M_{i_1} \ldots M_{i_p}) \tag{33}$$

$$r = 1 \quad \longrightarrow \quad \sum_{k_1=1}^{p} \epsilon_{i_{k_1}} \, \mathrm{Tr}(M_{i_{k_1}})[1 + \epsilon_{i_1} \ldots \epsilon_{i_p} *] \, \mathrm{D}_{i_{k_1}} \, \mathrm{Tr}(M_{i_1} \ldots M_{i_p}) \tag{34}$$

$$r = 2 \quad \longrightarrow \quad \sum_{k_1<k_2=1}^{p} \epsilon_{i_{k_1}} \epsilon_{i_{k_2}} \, \mathrm{Tr}(M_{i_{k_1}} M_{i_{k_2}})[1 + \epsilon_{i_1} \ldots \epsilon_{i_p} *] \, \mathrm{D}_{i_{k_2}} \, \mathrm{D}_{i_{k_1}} \, \mathrm{Tr}(M_{i_1} \ldots M_{i_p}). \tag{35}$$