`dennm-torch`: (DEN)sity of (N)on-(M)arkovian phenomena with Libtorch

The computational framework for `dennm-torch` is derived using the formalism from the publicly-available book: Worlds Of Observation. Let's begin by writing the generalised non-Markovian master equation for stochastic phenomena which was given in this chapter of the book

$$P_{\mathsf{t}+1}(x\vert z) = \int_{\Omega_{\mathsf{t}}} \mathrm{d}X' P_{\mathsf{t}+1}(X\vert z) = \int_{\Omega_{\mathsf{t}}} \mathrm{d}X' P_{\mathsf{t}}(X'\vert z) P_{(\mathsf{t}+1)\mathsf{t}}(x\vert X', z).$$

Without loss of generality, we can relate the latest probabilities to those from deeper into the past by chaining conditional probabilities together in a non-Markovian equivalent of the Chapman-Kolmogorov equation

$$\begin{aligned}
P_{\mathsf{t}+1}(x\vert z) &= \int_{\Omega_{\mathsf{t}-1}} \mathrm{d}X'' P_{\mathsf{t}-1}(X''\vert z) \int_{\omega_{\mathsf{t}}} \mathrm{d}^n x' P_{\mathsf{t}(\mathsf{t}-1)}(x'\vert X'', z) P_{(\mathsf{t}+1)\mathsf{t}}(x\vert X', z) \\
&= \int_{\Omega_{\mathsf{t}-2}} \mathrm{d}X''' P_{\mathsf{t}-2}(X'''\vert z) \int_{\omega_{\mathsf{t}-1}} \mathrm{d}^n x'' P_{(\mathsf{t}-1)(\mathsf{t}-2)}(x''\vert X''', z) \int_{\omega_{\mathsf{t}}} \mathrm{d}^n x' P_{\mathsf{t}(\mathsf{t}-1)}(x'\vert X'', z) P_{(\mathsf{t}+1)\mathsf{t}}(x\vert X', z) \\
&= \dots \\
&= \int_{\Omega_{\mathsf{t}-\mathsf{s}}} \mathrm{d}X''' P_{\mathsf{t}-\mathsf{s}}(X'''\vert z) \prod_{\mathsf{s}'=0}^{\mathsf{s}-1} \left\{ \int_{\omega_{\mathsf{t}-\mathsf{s}'}} \mathrm{d}^n x' P_{(\mathsf{t}-\mathsf{s}')(\mathsf{t}-\mathsf{s}'-1)}(x'\vert X'', z) \right\} P_{(\mathsf{t}+1)\mathsf{t}}(x\vert X', z).
\end{aligned}$$

Depending on the temporal correlation structure of the process, the conditional probabilities can be factorised. For example, processes with second or third-order temporal correlations would be described by the following expressions

$$\begin{aligned}
P_{(\mathsf{t}+1)\mathsf{t}}(x\vert X', z) &= \frac{1}{\mathsf{t}} \sum_{\mathsf{t}'=0}^{\mathsf{t}} \int_{\omega_{\mathsf{t}'}} \mathrm{d}^n x' P_{\mathsf{t}'}(x'\vert z) P_{(\mathsf{t}+1)\mathsf{t}'}(x\vert x', z) \\
P_{(\mathsf{t}+1)\mathsf{t}}(x\vert X', z) &= \frac{1}{\mathsf{t}} \sum_{\mathsf{t}'=0}^{\mathsf{t}} \frac{1}{\mathsf{t}'} \sum_{\mathsf{t}'=0}^{\mathsf{t}'} \int_{\omega_{\mathsf{t}'}} \mathrm{d}^n x' \int_{\omega_{\mathsf{t}''}} \mathrm{d}^n x'' P_{\mathsf{t}''}(x''\vert z) P_{\mathsf{t}'\mathsf{t}''}(x'\vert x'', z) P_{(\mathsf{t}+1)\mathsf{t}'\mathsf{t}''}(x\vert x', x'', z).
\end{aligned}$$

Let's imagine that x is just a scalar (as opposed to a row vector) for simplicity in the expressions. We can then discretise the 1D space over x into separate i-labelled regions such that $[P]^i_{\mathsf{t}+1} - [P]^i_{\mathsf{t}} = J^i_{\mathsf{t}+1}$, where the probability current $J^i_{\mathsf{t}+1}$ for the factorised processes above would be defined as

$$\begin{aligned}
J^i_{\mathsf{t}+1} &= -[P]^i_{\mathsf{t}} + \frac{1}{\mathsf{t}} \sum_{\mathsf{t}'=0}^{\mathsf{t}} \sum_{i'=0}^{N} \Delta x [P]^{i'}_{\mathsf{t}'} [P]^{ii'}_{(\mathsf{t}+1)\mathsf{t}'} \\
J^i_{\mathsf{t}+1} &= -[P]^i_{\mathsf{t}} + \frac{1}{\mathsf{t}} \sum_{\mathsf{t}'=1}^{\mathsf{t}} \frac{1}{\mathsf{t}'-1} \sum_{\mathsf{t}''=0}^{\mathsf{t}'-1} \sum_{i'=0}^{N} \sum_{i''=0}^{N} \Delta x^2 [P]^{i''}_{\mathsf{t}''} [P]^{i'i''}_{\mathsf{t}'\mathsf{t}''} [P]^{ii'i''}_{(\mathsf{t}+1)\mathsf{t}'\mathsf{t}''}.
\end{aligned}$$

The $[P]^{ii'i''}_{(\mathsf{t}+1)\mathsf{t}'\mathsf{t}''}$ tensor, in particular, will have $N^3\mathsf{t}(\mathsf{t}^2-1)$ elements. Note that the third-order temporal correlations can be evolved by identifying the pairwise conditional probabilities as time-dependent state variables and evolving them according to the following relation

$$[P]^{ii''}_{(\mathsf{t}+1)\mathsf{t}''} = \frac{1}{\mathsf{t}} \sum_{\mathsf{t}'=1}^{\mathsf{t}} \sum_{i'=0}^{N} \Delta x [P]^{i'i''}_{\mathsf{t}'\mathsf{t}''} [P]^{ii'i''}_{(\mathsf{t}+1)\mathsf{t}'\mathsf{t}''}.$$