Geodesics

In the paper (`https://arxiv.org/pdf/2011.13456`), they show that forward diffusion (VP specifically) is modeled by the stochastic differential equation

$$dx = -\frac{1}{2}\beta(t)x\ dt + \sqrt{\beta(t)}\ dW_t$$

There is a fact that this is the same as (if $\epsilon \sim \mathcal{N}(0, I)$)

$$X_t = X_0 e^{-\frac{1}{2}\int_0^t \beta(s)\ ds} + \sqrt{1 - e^{-\int_0^t \beta(s)\ ds}}\epsilon$$

And any given datapoint will be pushed through the forward process by

$$X_t[i] = X_0[i] e^{-\frac{1}{2}\int_0^t \beta(s)\ ds} + \sqrt{1 - e^{-\int_0^t \beta(s)\ ds}}\epsilon$$

Clearly if x, y are in our data manifold, and z is the geodesic midpoint on the data manifold, we can find exactly what that midpoint should be at any time t by just plugging the original point in (and what x and y should be at t by plugging them in).

There is an algorithm for approximating geodesics on a sampled manifold M called the Fast Marching Method (`https://www.pnas.org/doi/epdf/10.1073/pnas.95.15.8431`). How this works is by setting up an equation for some specified point $x_0 \in M$

$$|\nabla T_{x_0}(x)| = 1,\ T(x_0) = 0$$
$$T_{x_0}(x) = d(x_0, x) \text{ ¡- Actual manifold distance}$$

Then the Fast Marching Method solves this equation (I won't explain how, it is in the paper I linked).

Once it is solved, we merely compute $T_x(z)$ and $T_y(z)$ for our datapoints x and y, then we solve the following minimization problem for some small tolerance $\epsilon > 0$

$$z = \arg \min_{\substack{p \in M \\ |T_x(p) - T_y(p)| < \epsilon}} T_x(p) + T_y(p)$$

Thus we can take two points $x, y \in M$ from our dataset, use the Fast Marching Method to find a z that is approximately the midpoint of x and y on M, then pass x, y, z through our forward SDE to recover what a good guess for the true midpoint of x and y on M should be at any given timestep.

To illustrate my point that we really need to know the noise ϵ to get a good guess at high timesteps (even if we know the schedule $\beta(s)$), we show that $\lim_{t\to\infty} \mathbb{E}(X_t[z]) = 0$ and $\lim_{t\to\infty} \text{Std}(X_t[z]) = I$.

We can take the expected value here to find

$$\mathbb{E}(X_t[z]) = X_0[z]e^{-\frac{1}{2}\int\limits_0^s \beta(s)\ ds}$$

This is going to tend to zero, which makes sense because our whole thing is diffusing to Gaussian noise (with expected value zero). The variance is (first term vanishes because it is constant)

$$\text{Var}(X_t[z]) = \text{Var}\left(\sqrt{1 - e^{-\int\limits_0^t \beta(s)\ ds}}\epsilon\right) = \left(\sqrt{1 - e^{-\int\limits_0^t \beta(s)\ ds}}\right)^2 \text{Var}(\epsilon) = \left(1 - e^{-\int\limits_0^t \beta(s)\ ds}\right) I$$

And the standard deviation is

$$\text{Std}(X_t[z]) = I\sqrt{1 - e^{-\int\limits_0^t \beta(s)\ ds}}$$

These are going to converge to the variance and standard deviation of Gaussian noise, which also makes sense.