# Approximate Bayesian inference through the lens of large deviations

One part technical discussion

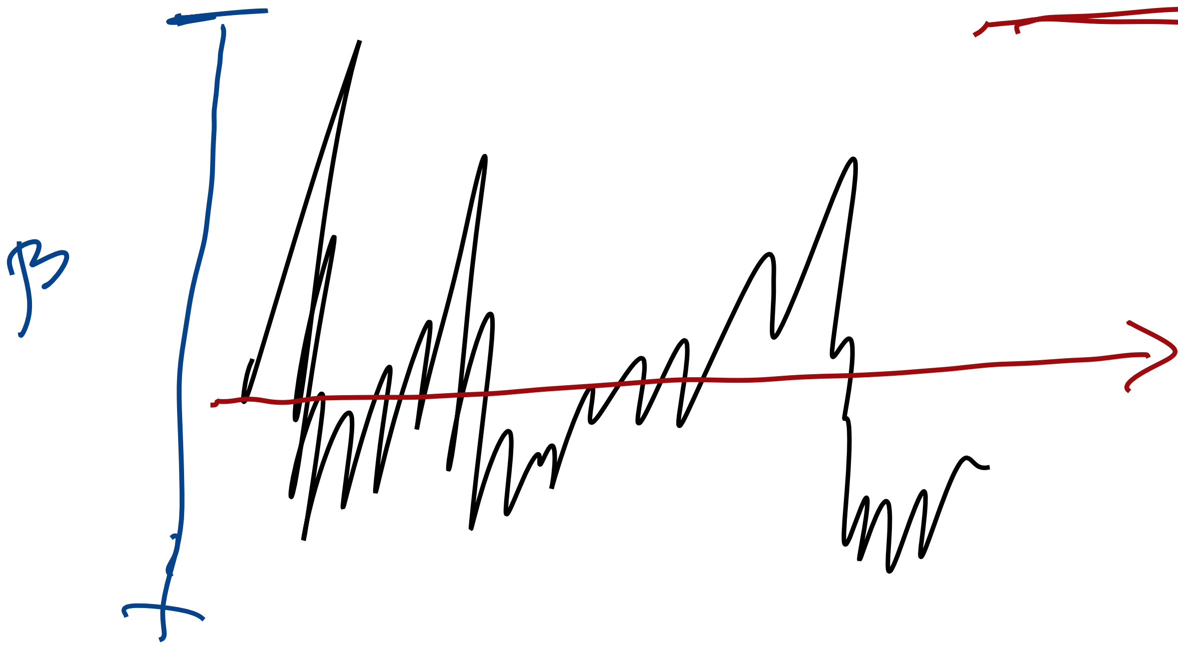
One part "Where do we go next?"

$$dX_t = \underline{f(X_t)\,dt} + \underline{\beta\, dW_t}$$

Itô SDE

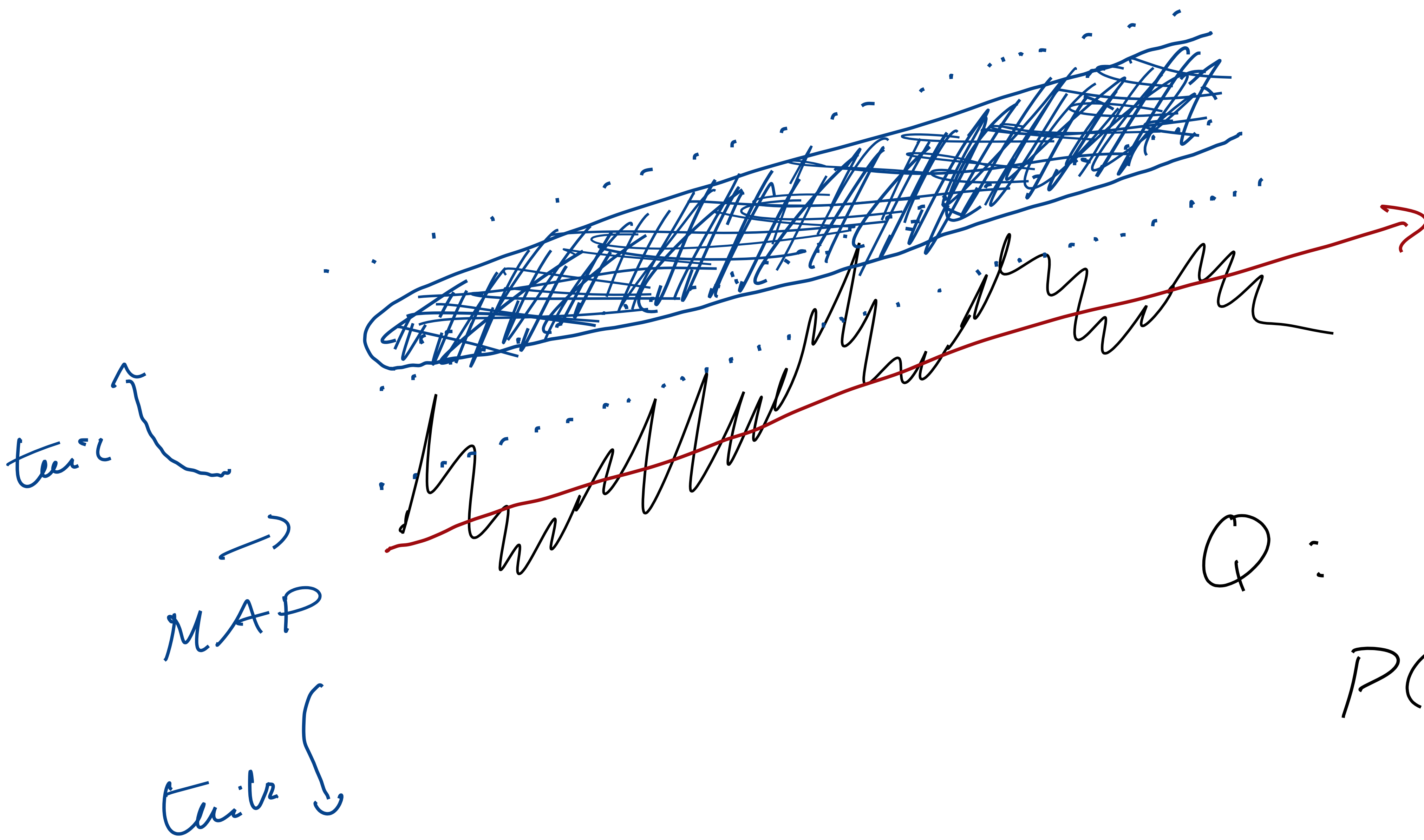
$\downarrow$

Rough paths




$$\int_0^T f(X_t)\,dt$$

$$dX_t = f(X_t)\,dt + \beta(X_t)\,dY_t$$



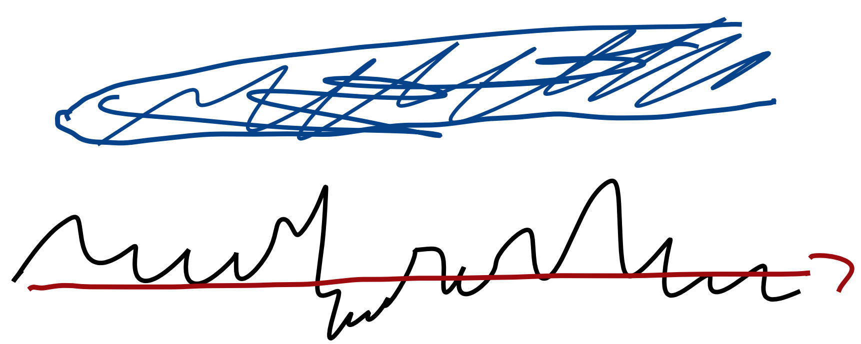

Q: What is $P(x_t \in A)$

$$\sim e^{-C\|x-m\|^2}$$

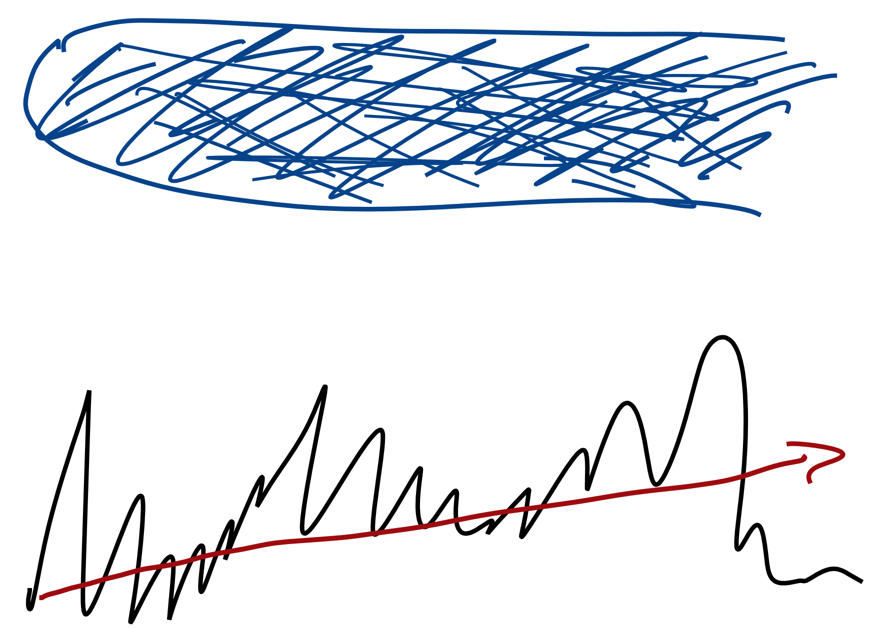
$A$ (intuitive): - how large is $\beta$

- how far away is $A$

for small $\beta$ and/or far away $A$, $P \approx \emptyset$.



$\beta \ll 1$



Formalise: $p \sim e^{-\frac{1}{\beta} \underline{I}(A)}$

$P(X_t \in A)$ depends on

. dist . $|\mu|$

$\inf A = F$

$$e^{-\frac{1}{\beta} \|F - m\|_T^2}$$

Norm on the path space (of $W_t$)

$$\left( \int_0^T \dot{x}_t^2 \, dt \right)^{1/2}$$

$\langle a, a \rangle = \|a\|$

$a \cdot a =$

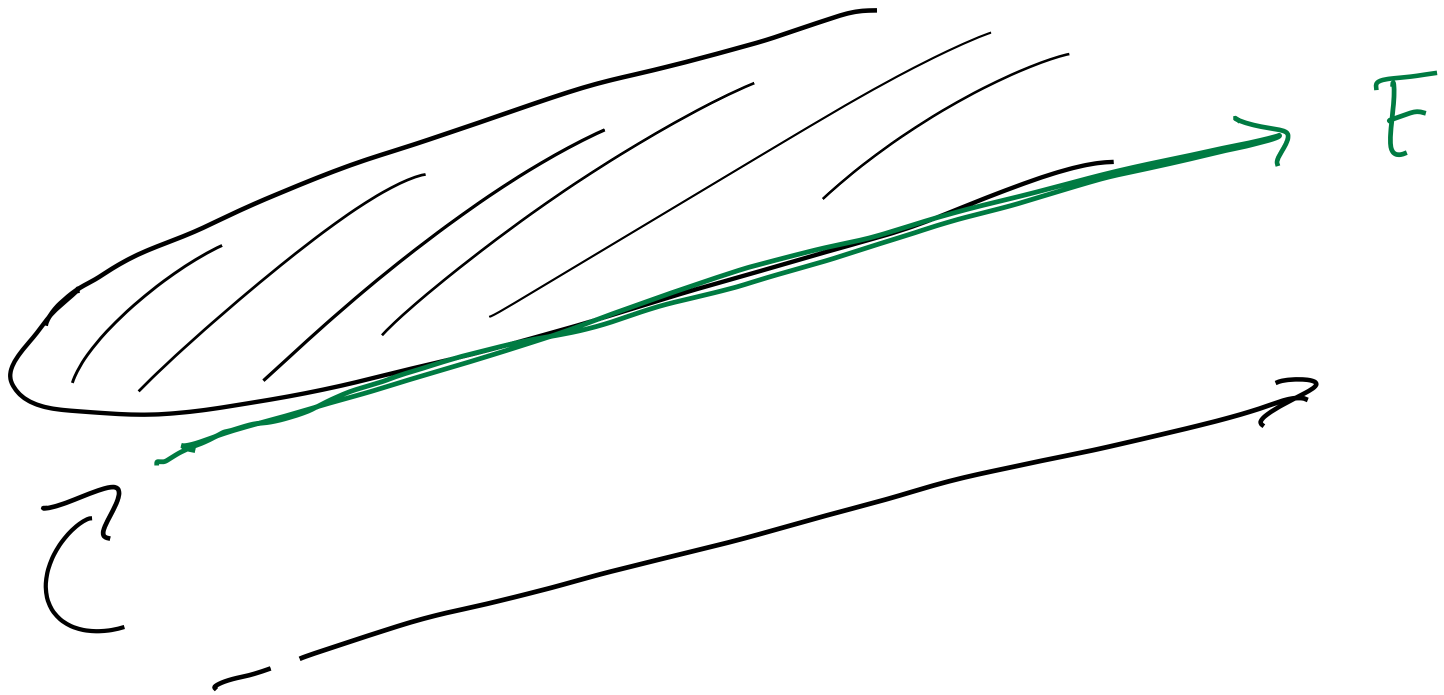
$\|a\| = a^2$

$$\| x_T \|_T^2 = \int_0^T \dot{x}_t^2 \, dt$$

Cameron – Martin then

$\Longrightarrow$

$$I(x_t) = \int_0^T (\dot{x}_t - \dot{m}_t)^2 \, dt$$




$$I(F) = \int_0^T (\dot{F}_t - \dot{m}_t)^2 \, dt$$

$$P(x_t \in A) \sim e^{-c_\beta I(\inf A)}$$

$c_\beta \to \infty$ as $\beta \to \infty$

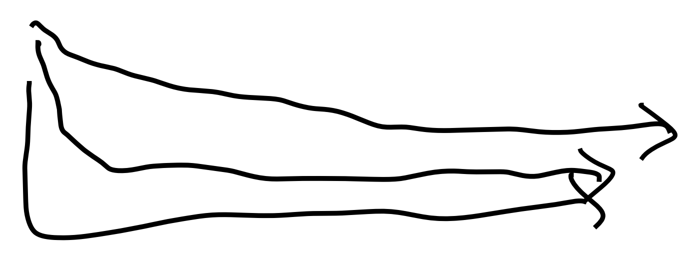
$$e^{-\infty I(\inf A)}$$

$$I = 0$$

$$P(x_t \in F) + P(x_t \in F+h) + P(x_t + 2h) + \dots$$

$$e^{-\infty \cdot 0} = e^{0} = 1$$

$$e^{-C I(F)} + e^{-C I(F+h)} + \dots$$



Slap.

Unlikely events happen in
the least unlikely unlikely way

– F Hollander

## 2.5 other examples of LDPs

Stat mech

Let

Thermo

- $x$ a microstate
- $M(x)$ = a macrostate
- $\arg\max P(M(x))$ = an eq. state
- $n$ = system size.

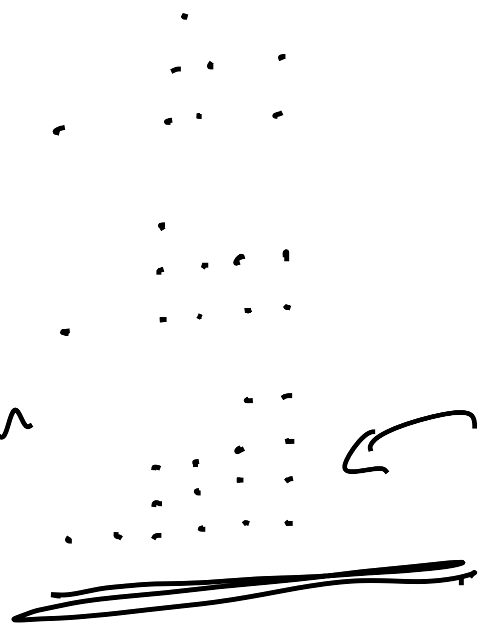
# Thermodynamics

as $\underbrace{n \to \infty}_{\text{Thermo. limit}}$ the system goes to eq state

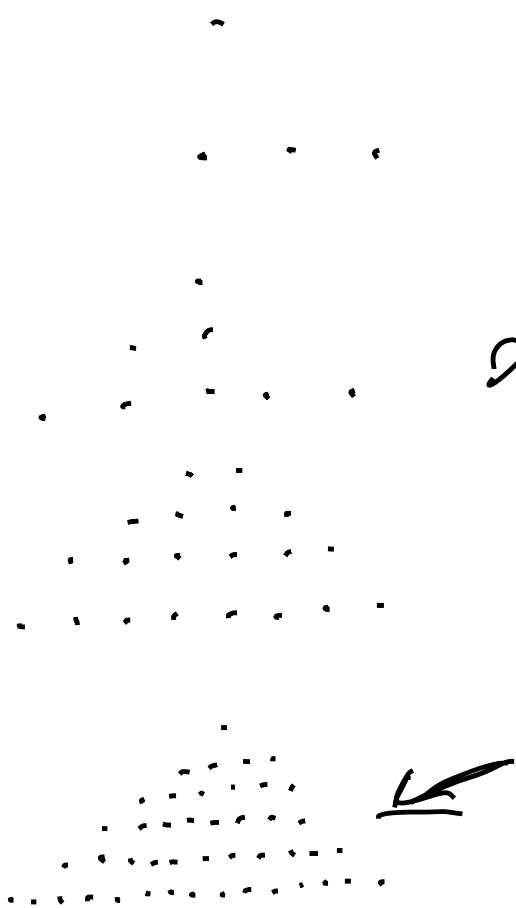
$\implies$ fluctuations away from min $I \to 0$

$\implies$ Thermo satisfies a LDP

exp unlikely when measure is Gaussian



$n \to \infty$



Another example:

as $n \to \infty$

emp. distn converges to the stationary meas.

$$D_{KL}(P_{emp} \,\|\, p) \to 0$$

$$x \to \gamma$$

$$\boxed{\text{Sanov's thm}}$$

Lemma: Sanov's thm holds on the path space.

ensemble of sample paths

Remark

$C([0,T], \mathbb{R})$ is a Polish space

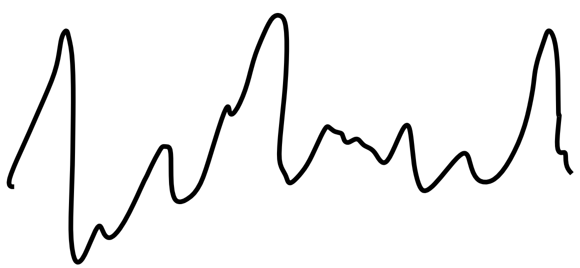
as $N \to \infty$ $D_{KL}\left(\rho(\gamma)_{emp} \,||\, \rho(\gamma)\right)$

$\Longrightarrow$ max cal principle

$\alpha < 1/3$

$\alpha < 1/2$

$\alpha > 2/3$



Further remark:

The space of geometric rough paths for any Hölder exp. is Polish

$\Longrightarrow$ Can prove Sanov's thm

$$C^{\alpha}_{g}(\;-\;)$$

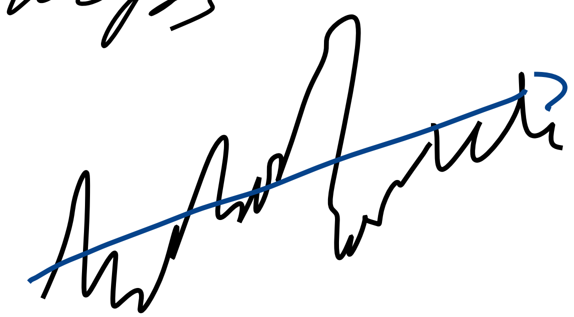
Main theorem

Start w 2 maps

Remark: whole conv. will be valid for Strat. given an app. correction term.

$t \mapsto x_t$

$t \mapsto Y_t$



– families of RV indexed by time

– driven by an Itô SDE w state-dep diffusion coeff.

Max Cal $\Rightarrow$ the stats. of sample paths tend towards the ensemble $p$

If the sample paths themselves are constrained by some cost

$\Rightarrow$ the cost is reflected in $P$(sample)

Intertwinement of 2 limits

So as $n \to \infty$ for fixed $\beta$,
Sanov's thm $\Rightarrow$ $-\log p = \frac{1}{\beta} V$
to first order in $\beta$

and $n \to \infty$ with $\beta \to \phi$ $\Rightarrow$
$p \to \delta_{p.l.a(x_t)}$

$\Longrightarrow$ Sanov's thm produces state-wise LDPs

( which we know!

$$p = \max D_{KL} = e^{-\frac{1}{\beta} V},$$

which satisfies an LDP )

Natural cost on sample paths:

Freidlin-Wentzell 98

$= $ speed pa.

$\downarrow$

$$-\log p(x_t) = C_\beta \int_0^T \left( \dot{x}_t - \partial_t \mathbb{E} x_t \right)^2 dt$$

$F_t$

$:= C_\beta \|w\|_2^2$

$G_t$

$$-\log q(y_t) = C_\beta \int_0^T \left( \dot{y}_t - \partial_t \mathbb{E} y_t \right)^2 dt$$

$:= C_\beta \|v\|_2^2$

- set an additonal cost

$$\mathbb{E}_p x_t = \sigma \left( \mathbb{E}_q y_t \right)$$

Then as $\beta \to 0$ path of least is synchronisation

But recall: this requires sampler under COA

min $(DKL_{T1})$

Main result

Statement based on entanglement:

min FET to find the path stats

min $h$ to get synched.

Fin