1. $E_{x\sim p(x)} E_{v\sim p(v)} [\| v^T s(x;\theta)\|^2 + 2v^T \nabla_x (v^T s(x;\theta))]$

$= E_{x\sim p(x)} E_{v\sim p(v)} (\| v^T s(x;\theta)\|^2) + E_{x\sim p(x)} E_{v\sim p(v)} (2v^T \nabla_x (v^T s(x;\theta)))$ — ①

Since $E_{x\sim p(x)} E_{v\sim p(v)} (\| v^T s(x;\theta)\|^2)$

$= E_{x\sim p(x)} E_{v\sim p(v)} (v^T s(x;\theta))^2$

$= E_{x\sim p(x)} E_{v\sim p(v)} (s(x;\theta)^T (v v^T) s(x;\theta))$

By $E[vv^T]=I$ $= E_{x\sim p(x)} (s(x;\theta)^T s(x;\theta))$

$= E_{x\sim p(x)} \| s(x;\theta)\|^2$

Therefore, ① $= E_{x\sim p(x)} \| s(x;\theta)\|^2 + E_{x\sim p(x)} E_{v\sim p(v)} (2v^T \nabla_x (v^T s(x;\theta)))$
$= L_{SSM}(\theta)$

2. SDE is a type of differential equation that involves randomness

It can be written as

$$dX_t = \underbrace{f(x_t, t)\,dt}_{\text{drift}} + \underbrace{G(x_t, t)\,dW_t}_{\text{diffusion}}$$

and W_t is Wiener process (standard Brownian motion) defined by (1) $W_0 \equiv 0$

(2) $\Delta W = (t+\Delta t) - W(t) \sim N(0, \Delta t I)$

(3) $0 = t_0 < t_1 < t_2 \cdots < t_n = T$

$W_{t_1} - W_{t_0}, \cdots W_{t_n} - W_{t_{n-1}}$ are independent

(4) continuous path with probability 1

3. Still don't understand how to get $E[X(t)]$ and $Var[X(t)]$