Dict

BALD:Bayesian Active Learning by Disagreement; **BLR**:Bayesian Linear Reg; **BN**:Bayesian Network; **BNN**:Bayesian NN; **BO**:Bayesian Opt; **BP**:Belief Propagation; **CPD**:Cond Prob Dist; **DAG**:Directed Acyclic Graph; **DBE**:Detailed Balance Eq; **DDIM**:Denoising Diffusion Implicit Models; **DDPG**:Deep Deterministic PG; **DDPM**:Denoising Diffusion Prob Models; **DQN**:Deep Q-Net; **ECE**:Expected Calibration Error; **EI**:Expected Improvement; **ELBO**:Evidence Lower Bound; **GP**:Gaussian Process; **HMM**:Hidden Markov Model; **KF**:Kalman Filter; **KL**:Kullback-Leibler; **LDM**:Latent Diffusion; **LOTV**:Law of Total Var; **MALA**:Metropolis-Adjusted Langevin; **MAP**:Max A Posteriori; **MCMC**:Markov Chain MC; **MDP**:Markov Decision Process; **MH**:Metropolis-Hastings; **MI**:Mutual Info; **MLE**:Max Likelihood Est; **MPE**:Most Probable Explanation; **PF**:Particle Filter; **PI**:Prob of Improvement; **POMDP**:Partially Observable MDP; **RBF**:Radial Basis Fnc; **RFF**:Random Fourier Features; **SGLD**:Stoch Grad Langevin Dyn; **SWAG**:Stoch Weight Avg Gaussian; **TD**:Temporal Diff; **UCB**:Upper Confidence Bound; **VE**:Var Elimination; **VI**:Variational Inference; $k_{XX'} := k(X, X')$; $K_y := K_{XX} + \sigma_n^2 I$

Probability Fundamentals

Axioms: $\mathbb{P}(\Omega) = 1$; $\mathbb{P}(A) \ge 0$; Disjoint: $\mathbb{P}\left(\bigcup_i A_i\right) = \sum_i \mathbb{P}(A_i)$ **Product**: $\mathbb{P}(X_{1:n}) = \mathbb{P}(X_1) \prod_{i=2}^n \mathbb{P}(X_i | X_{1:i-1})$ **Sum**: $\mathbb{P}(X) = \sum_y \mathbb{P}(X, y)$ **Bayes**: $\mathbb{P}(X|Y) = \frac{\mathbb{P}(Y|X)\mathbb{P}(X)}{\mathbb{P}(Y)}$ **Cond Indep**: $X \perp Y | Z \leftrightarrow \mathbb{P}(X, Y|Z) = \mathbb{P}(X|Z)\mathbb{P}(Y|Z)$

Gaussian $\mathcal{N}(\mu, \Sigma)$: $\mathcal{N}(x; \mu, \Sigma) = \frac{1}{\sqrt{(2\pi)^d \det(\Sigma)}} \exp\left(-\frac{1}{2}(x-\mu)^\top \Sigma^{-1}(x-\mu)\right)$ **Marginal**: $X_A \sim \mathcal{N}(\mu_A, \Sigma_{AA})$ **Conditional**: $X_A | X_B \sim \mathcal{N}\left(\mu_{A|B}, \Sigma_{A|B}\right)$ $\mu_{A|B} = \mu_A + \Sigma_{AB}\Sigma_{BB}^{-1}(x_B - \mu_B)$ $\Sigma_{A|B} = \Sigma_{AA} - \Sigma_{AB}\Sigma_{BB}^{-1}\Sigma_{BA}$ **Linear**: $Y = MX \sim \mathcal{N}(M\mu, M\Sigma M^\top)$ **Sum**: indep $X + X' \sim \mathcal{N}(\mu + \mu', \Sigma + \Sigma')$

$\mathbb{E}$, Var, Cov: $\mathbb{E}[AX + b] = A\mathbb{E}[X] + b$; **Tower**: $\mathbb{E}_Y\left[\mathbb{E}_{X|X|Y}\right] = \mathbb{E}[X]$ $\mathrm{Var}[X] = \mathbb{E}[(X - \mathbb{E}[X])^2]$; $\mathrm{Cov}[X, Y] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$ $\mathrm{Var}[X+Y] = \mathrm{Var}[X] + \mathrm{Var}[Y] + 2\mathrm{Cov}[X, Y]$ **LOTV**: $\mathrm{Var}[X] = \mathbb{E}[\mathrm{Var}[X|Y]] + \mathrm{Var}[\mathbb{E}[X|Y]]$

Info Theory: **Entropy**: $H[p] = -\mathbb{E}_{p[\log p(x)]}$; Gauss $H = \frac{1}{2}\log\left((2\pi e)^d \det \Sigma\right)$ **KL**: $\mathrm{KL}(p\|q) = \mathbb{E}_{p\left[\log \frac{p}{q}\right]} \ge 0$; need $\mathrm{supp}(q) \subseteq \mathrm{supp}(p)$ **Forward** $\mathrm{KL}(p\|q)$: mean-seeking 覆盖; **Reverse** $\mathrm{KL}(q\|p)$: mode-seeking 过confident **MI**: $I(X; Y) = H[X] - H[X|Y] = H[Y] - H[Y|X] \ge 0$, symmetric **Cond MI**: $I(X; Y|Z) = H[X|Z] - H[X|Y, Z]$ **Gauss MI**: $I[X; Y] = \frac{1}{2}\log\det(I + \sigma_n^{-2}\Sigma)$ for $Y = X + \varepsilon$
MLE: $\hat{\theta}_{\mathrm{MLE}} = \arg\max_\theta \sum_i \log p(y_i | x_i, \theta)$
MAP: $\hat{\theta}_{\mathrm{MAP}} = \arg\min_\theta \underbrace{-\log p(\theta)}_{\text{reg}} + \underbrace{\ell_{\mathrm{nll}}}_{\text{fit}}$ Gaussian prior→L2; Laplace prior→L1

Bayesian Linear Regression

Model: $y = w^\top x + \varepsilon$, $\varepsilon \sim \mathcal{N}(0, \sigma_n^2)$; Prior: $w \sim \mathcal{N}(0, \sigma_p^2 I)$

Posterior: $w|X, y \sim \mathcal{N}(\mu, \Sigma)$ where $\Sigma^{-1} = \sigma_n^{-2}X^\top X + \sigma_p^{-2}I$; $\mu = \sigma_n^{-2}\Sigma X^\top y$ Note: Σ只依赖X, 不依赖y

Prediction: $y_* | x_*, X, y \sim \mathcal{N}(x_*^\top \mu, x_*^\top \Sigma x_* + \sigma_n^2)$

Connection: `MAP`=Ridge with $\lambda = \frac{\sigma_n^2}{\sigma_p^2}$; Online update: $O(nd^2)$

Gaussian Processes

Def: $f \sim \mathcal{GP}(\mu, k)$: any finite subset jointly Gaussian. $f_A \sim \mathcal{N}(\mu_A, K_{AA})$, $[K_{AA}]_{ij} = k(x_i, x_j)$

GP Regression: $y \sim \mathcal{N}(0, K_{XX} + \sigma_n^2 I) = \mathcal{N}(0, K_y)$ **Mean**: $\mu_*(x) = k(x, X)K_y^{-1}y$ **Cov**: $k_*(x, x') = k(x, x') - k(x, X)K_y^{-1}k(X, x')$ **Predictive**: $y_* \sim \mathcal{N}(\mu_*, k_* + \sigma_n^2)$

Kernels: **Linear**: $k(x, x') = x^\top x' + \sigma_0^2$ **RBF**: $k = \exp\left(-\frac{\|x - x'\|^2}{2\ell^2}\right)$ smooth 无限可微 **Exponential**: $k = \exp\left(-\frac{\|x - x'\|}{\ell}\right)$ rough **Matérn**: $\nu = 0.5$→Exp, $\nu \to \infty$→RBF, ν控制smoothness **Periodic**: $k = \sigma^2 \exp\left(-\frac{2}{\ell^2}\sin^2\left(\pi|x - x'\frac{1}{p}|\right)\right)$ **Closure**: $k_1 + k_2$, $k_1 \cdot k_2$, $c \cdot k$, $\exp(k)$仍valid kernel **Stationary**: $k(x, x') = k(x - x')$; **Isotropic**: $k = k(\|x - x'\|)$

Marginal Lik: $\log p(y|X) = -\frac{1}{2}y^\top K_y^{-1}y - \frac{1}{2}\log\det(K_y) + C$ Balance: Datafit(前) vs Complexity(后)

Approx $O(n^3)$ → lower: **RFF**: $k(x - x') \approx \varphi(x)^\top \varphi(x')$, $O(nm^2 + m^3)$ Bochner: stationary kernel ↔ Fourier of non-neg measure **Inducing Pts**: subset $m \ll n$ points for approx

Variational Inference

Goal: Approx $p(\theta|D)$ with $q(\theta|\lambda)$ by min $\mathrm{KL}(q\|p)$

ELBO: $\mathcal{L} = \mathbb{E}_{q[\log p(y|\theta)]} - \mathrm{KL}(q(\theta)\|p(\theta))$ $\log p(y) = \mathcal{L} + \mathrm{KL}(q\|p(\cdot|D)) \ge \mathcal{L}$ Max ELBO ⟺ Min KL to posterior **Derivation**: Jensen's $\log \mathbb{E}_{q\left[\frac{p}{q}\right]} \ge \mathbb{E}_{q\left[\log \frac{p}{q}\right]}$

Gaussian KL: $\mathrm{KL}(\mathcal{N}_p\|\mathcal{N}_q) = \frac{1}{2}\left[\mathrm{tr}(\Sigma_q^{-1}\Sigma_p) + (\mu_p - \mu_q)^\top \Sigma_q^{-1}(\mu_p - \mu_q) - d + \log\left(\det \frac{\Sigma_q}{\det}\Sigma_p\right)\right]$
Product: $\mathrm{KL}(Q_X Q_Y\|P_X P_Y) = \mathrm{KL}(Q_X\|P_X) + \mathrm{KL}(Q_Y\|P_Y)$

Reparam Trick: $\theta = g(\varepsilon; \lambda)$, $\varepsilon \sim \varphi$ $\mathbb{E}_{\theta\sim q}[f(\theta)] = \mathbb{E}_{\varepsilon[f(g(\varepsilon;\lambda))]}$ **Gaussian**: $\theta = \mu + \sigma \odot \varepsilon$, $\varepsilon \sim \mathcal{N}(0, I)$ Enable gradient: $\nabla_\lambda \mathbb{E}_{q[f]} = \mathbb{E}_{\varphi[\nabla_\lambda f(g(\varepsilon;\lambda))]}$

Laplace Approx: $q(\theta) = \mathcal{N}\left(\hat{\theta}, \Lambda^{-1}\right)$ $\hat{\theta}$ = MAP; $\Lambda = -\nabla^2 \log p\left(\hat{\theta}|D\right)$ (Hessian) Good at mode, overconfident elsewhere

Markov Chains & MCMC

MC basics: **Markov**: $X_{t+1} \perp X_{1:t-1} | X_t$ **Stationary** π: $\pi = \pi P$ **Irreducible**: all states reachable from any state **Aperiodic**: $\gcd\{t : P^{t(x,x)} > 0\} = 1$ **Ergodic**=Irreducible+Aperiodic: unique $\pi > 0$, $\lim_{t\to\infty} q_t = \pi$

DBE: $\pi(x)P(x'|x) = \pi(x')P(x|x')$ If satisfied→π stationary, chain **reversible** **Proof**: sum over x'得 $\pi(x) = \sum_{x'} \pi(x')P(x|x')$

Ergodic Thrm: $\frac{1}{n}\sum_{i=1}^n f(X_i) \xrightarrow{\text{a.s.}} \mathbb{E}_{x\sim\pi}[f(x)]$ Hoeffding: error prob decays $\exp(-n)$

MH Algorithm: Propose $x' \sim R(x'|x)$. Accept w.p.: $\alpha(x'|x) = \min\left\{1, \frac{q(x')R(x|x')}{q(x)R(x'|x)}\right\}$ Stationary: $p(x) \propto q(x)$ (unnormalized OK) Satisfies DBE→correct stationary dist

Gibbs Sampling: Iterate: $x_i^{(t+1)} \sim p\left(x_i | x_{-i}^{(t)}\right)$ Special MH with acceptance=1 **Practical**: 顺序scan all vars, sample each from conditional

Langevin & SGLD: **Langevin**: $R(x'|x) = \mathcal{N}(x'; x - \eta\nabla f(x), 2\eta I)$ where $p \propto e^{-f}$ **MALA**: MH-corrected Langevin, poly-time for log-concave **SGLD**: $\theta_{t+1} = \theta_t + \varepsilon_{t(\nabla \log p(\theta) + \nabla \log p(D|\theta))} + \sqrt{2\varepsilon_t}\xi$ Converge: $\sum_t \varepsilon_t = \infty$, $\sum_t \varepsilon_t^2 < \infty$; 常用$\varepsilon_t \in \Theta\left(t^{-\frac{1}{3}}\right)$

Gibbs Distribution: $p(x) = \frac{1}{Z}\exp(-f(x))$, f=energy function Posterior always interpretable as Gibbs

Bayesian Neural Networks

Model: Prior: $\theta \sim \mathcal{N}\left(0, \sigma_p^2 I\right)$ **Homoscedastic**: $y|x, \theta \sim \mathcal{N}\left(f(x; \theta), \sigma^2\right)$ fixed noise **Heteroscedastic**: $y \sim \mathcal{N}\left(f_\mu(x; \theta), \exp(f_\sigma(x; \theta))\right)$ input-dependent noise

Hetero NLL: $-\log p(y|x, \theta) = C + \frac{1}{2}\left[\log \sigma^2(x) + \frac{(y - \mu(x))^2}{\sigma^2(x)}\right]$ Model can "blame" noise but pays $\log \sigma$ penalty 防collapse

MAP for BNN: $\hat{\theta}_{\mathrm{MAP}} = \arg\min \frac{1}{2\sigma_p^2}\|\theta\|^2 + \frac{1}{2\sigma_n^2}\sum_{i(y_i - f(x_i;\theta))}^2$ Weight decay = Gaussian prior

Prediction: $p(y_*|x_*, D) \approx \frac{1}{m}\sum_{j=1}^m p(y_*|x_*, \theta^{(j)})$, $\theta^{(\{j\})} \sim q$ MC approx of posterior predictive

Uncertainty Decomp: **Total Var=Aleatoric+Epistemic** Aleatoric(data noise): $\frac{1}{m}\sum_j \sigma^2(x_*, \theta^{(j)})$ Epistemic(model uncertainty): $\frac{1}{m}\sum_{j(\mu(x_*,\theta^{(j)}) - |(\mu))}^2$ where $|(\mu) = \frac{1}{m}\sum_j \mu(x_*, \theta^{(j)})$

MC Dropout: $q_{j(\theta_j)} = p\delta_0(\theta_j) + (1 - p)\delta_{\lambda_j}(\theta_j)$ Test时keep dropout→multiple forward passes→uncertainty

SWAG: Store running avg of SGD iterates: μ, Σ Space: $O(d^2)$ covariance vs $O(Td)$ all models

Calibration: Goal: Confidence ≈ Accuracy **ECE**: $\sum \frac{|B_m|}{n}|\mathrm{acc}(B_m) - \mathrm{conf}(B_m)|$ **Temp Scaling**: $\frac{z}{T}$ on logits; $T > 1$→less confident

Active Learning

Objective: $I(S) = I(f_S; y_S) = H[f_S] - H[f_S|y_S]$ NP-hard; Greedy gives $\left(1 - \frac{1}{e}\right)$-approx (submodular, monotone)

Strategies: **Uncertainty Sampling**: $x = \arg\max H[y_x|D]$ Cannot distinguish aleatoric vs epistemic **BALD**: $x = \arg\max I(\theta; y_x|D) = H[y_x|D] - \mathbb{E}_{\theta[H[y_x|\theta]]}$ Finds where models **disagree** about y_x **Hetero**: $x = \arg\max \frac{\sigma^2_{\text{epistemic}}}{\sigma^2_{\text{aleatoric}}}$

Submodular: $F(A \cup \{x\}) - F(A) \ge F(B \cup \{x\}) - F(B)$ for $A \subseteq B$ Diminishing returns; MI is submodular

Bayesian Optimization

Regret: $R_T = \sum_{t=1}^T (f^* - f(x_t))$ Goal: sublinear $\frac{R_T}{T} \to 0$

Acquisition Fncs: **UCB**: $x_{t+1} = \arg\max\left[\mu_{t(x)} + \beta_t \sigma_{t(x)}\right]$ $\beta_t = 0$: pure exploit; $\beta_t \to \infty$: uncertainty sampling Regret: $R_T = O\left(\sqrt{T\gamma_T}\right)$ **PI**: $\mathrm{PI}(x) = \Phi\left(\frac{\mu(x) - f(x^+)}{\sigma(x)}\right)$ **EI**: $\mathrm{EI}(x) = (\mu - f^+)\Phi(Z) + \sigma\varphi(Z)$, $Z = \frac{\mu - f^+}{\sigma}$ **Thompson**: Sample $\tilde{f} \sim p(f|D_t)$, pick $\arg\max \tilde{f}$

Info Gain γ_T: Linear: $\gamma_T = O(d \log T)$ RBF: $\gamma_T = O\left((\log T)^{d+1}\right)$ Matérn($\nu > \frac{1}{2}$): $\gamma_T = O\left(T^{\frac{d}{2\nu+d}}(\log T)^{2\frac{\nu}{2\nu+d}}\right)$

MDP & RL Foundations

MDP: $(\mathcal{S}, \mathcal{A}, P, R, \gamma)$: states, actions, $P(s'|s, a)$, reward, discount **Value**: $V^{\pi(s)} = \mathbb{E}\left[\sum_{t\ge0}\gamma^t R_t | s_0 = s, \pi\right]$ **Q-fnc**: $Q^{\pi(s,a)} = R(s, a) + \gamma\sum_{s'} P(s'|s, a)V^{\pi(s')}$

Bellman Eqs: **Expectation**: $V^{\pi(s)} = R(s, \pi(s)) + \gamma\sum_{s'} P(s'|s, \pi(s))V^{\pi(s')}$ **Optimality**: $V^*(s) = \max_{a\left[R(s,a) + \gamma\sum_{s'} P(s'|s,a)V^*(s')\right]}$ $Q^*(s, a) = R(s, a) + \gamma\sum_{s'} P(s'|s, a)\max_{a'} Q^*(s', a')$ **Matrix**: $\boldsymbol{v}^\pi = (\boldsymbol{I} - \gamma\boldsymbol{P}^\pi)^{-1}\boldsymbol{r}^\pi$

Bellman's Thrm: π^* optimal iff greedy w.r.t. own V^π: $\pi^*(s) = \arg\max_a Q^*(s, a)$

PI & VI: **Policy Iter**: (1)Eval V^π exactly(solve LSE), (2) π →greedy. Fewer iters, $O(n^3)$/iter. **Value Iter**: $V \to \max_{a[r+\gamma PV]}$. More iters, $O(n^2 m)$/iter. Both converge to optimal; VI gives ε-optimal

POMDP: **Belief**: $b_{t(x)} = P(X_t = x|y_{1:t}, a_{1:t-1})$ **Bayes Filter**: $b_{t+1}(x) \propto o(y_{t+1}|x)\sum_{x'} P(x|x', a_t)b_{t(x')}$ Belief-state MDP: reward $\rho(b, a) = \mathbb{E}_{x\sim b}[r(x, a)]$

Tabular RL

Model-based: $\hat{P}(x'|x, a) = \frac{N(x'|x,a)}{N(a|x)}$, $\hat{r}(x, a)$ =avg rewards Converges but needs many samples

Q-Learning (Off-policy): $Q(s, a) \leftarrow Q(s, a) + \alpha(r + \gamma\max_{a'} Q(s', a') - Q(s, a))$ Uses max (ideal best a'); off-policy

SARSA (On-policy): $Q(s, a) \leftarrow Q(s, a) + \alpha(r + \gamma Q(s', a') - Q(s, a))$ Uses actual a' from policy; on-policy

TD Learning: $V(s) \leftarrow V(s) + \alpha(r + \gamma V(s') - V(s))$ **As SGD**: $\ell = \frac{1}{2}(V(s) - (r + \gamma V(s')))^2$ Converges if Robbins-Monro: $\sum \alpha_t = \infty$, $\sum \alpha_t^2 < \infty$

Exploration: ε**-greedy**: prob ε random, else best **Optimistic Init**: $Q = \frac{R_{\max}}{1-\gamma}$ **Rmax**: unknown$(s, a) \to R_{\max}$, PAC guarantee

Deep RL

DQN: $\mathcal{L} = \left(r + \gamma\max_{a'} Q_{\theta^-}(s', a') - Q_{\theta(s,a)}\right)^2$ **Target Net** θ^-: stabilize; **Experience Replay**: break correlation **Double DQN**: selection θ, eval θ^-; reduces overestimation

Policy Gradient: $\nabla_\theta J = \mathbb{E}_{\tau\sim\pi_\theta}\left[\sum_t \nabla \log \pi_{\theta(a_t|s_t)} G_t\right]$ $\nabla \log P(\tau) = \sum_t \nabla \log \pi(a_t|s_t)$ (dynamics cancel!) **REINFORCE**: MC estimate, high variance **Baseline**: $G_t - b(s_t)$, $b = V(s)$ optimal; unbiased

Actor-Critic: **Actor**: $\pi_{\theta(a|s)}$; **Critic**: $V_{\varphi(s)}$ or $Q_{\varphi(s,a)}$ $\nabla J \approx \mathbb{E}[\nabla \log \pi(a|s)(Q(s, a) - V(s))]$ Critic bootstrap 减variance 但引入bias

Advanced: **TRPO**: $\max \mathbb{E}\left[\left(\frac{\pi_\theta}{\pi_{\text{old}}}\right)A^{\pi_{\text{old}}}\right]$ s.t. $\mathrm{KL} \le \delta$ **DDPG**: continuous actions, deterministic $\mu_{\theta(s)}$ **Adv Fnc**: $A^{\pi(s,a)} = Q^{\pi(s,a)} - V^{\pi(s)}$

Bayesian Networks

Def: DAG G + CPDs $P\left(X_v|\mathrm{Pa}_{X_v}\right)$ **Joint**: $P(X_{1:n}) = \prod_i P(X_i|\mathrm{Pa}_i)$ Variable order 重要 for compact representation

D-Separation: $X \perp Y|Z$ iff all paths blocked by Z **Active trails** (path 通): Chain $X \to Y \to Z$: Y **not** observed Fork $X \leftarrow Y \to Z$: Y **not** observed Collider $X \to Y \leftarrow Z$: Y or descendant **is** observed

Inference: **VE**: Sum out non-query vars; complexity = treewidth **BP (Sum-Product)**: $\mu_{v\to u} \propto \prod \mu_{u'\to v}$ $\mu_{u\to v} \propto \sum f_u \prod \mu_{v'\to u}$ Exact on trees; loopy 可能不 converge **Max-Product**: replace $\sum$ with max for MPE/MAP

Approx Inference: **Rejection Sampling**: discard samples 不符 evidence, inefficient if rare **Likelihood Weighting**: weight by evidence prob **Gibbs**: sample each var from conditional given rest

Learning: **Params**: $\hat{\theta}_{X_i|\mathrm{Pa}_i} = \frac{\mathrm{count}(X_i,\mathrm{Pa}_i)}{\mathrm{count}(\mathrm{Pa}_i)}$ (MLE) **Structure**: Score-based, MLE score 偏好 fully connected **BIC**: $S_{\mathrm{BIC}} = \sum \hat{I}(X_i;\mathrm{Pa}_i) - \frac{\log N}{2N}|G|$ **Chow-Liu**: max spanning tree on MI weights⟶optimal tree BN

Diffusion Models

Setup: **Forward**: data⟶noise (fixed, no learning) **Backward**: noise⟶data (learned generation) Latent var model: $x_{1:T}$ are latents, x_0 is data

Forward Process: $q(x_t|x_{t-1}) = \mathcal{N}\left(x_t; \sqrt{1-\beta_t}x_{t-1}, \beta_t I\right)$ $x_t = \sqrt{1-\beta_t}x_{t-1} + \sqrt{\beta_t}\varepsilon_t$ Schedule: $\beta_t \in (0,1)$单调增, $\beta_1 \approx 10^{-4}$, $\beta_T \approx 0.02$

Closed-Form Marginal ⭐: Define: $\alpha_t = 1 - \beta_t$, $|(\alpha)_t = \prod_{s=1}^t \alpha_s$ $q(x_t|x_0) = \mathcal{N}\left(\sqrt{|(\alpha)_t}x_0, (1-|(\alpha)_t)I\right)$ **Reparam**: $x_t = \sqrt{|(\alpha)_t}x_0 + \sqrt{1-|(\alpha)_t}\varepsilon$, $\varepsilon \sim \mathcal{N}(0,I)$ As $t \to T$: $|(\alpha)_T \to 0$, $x_T \sim \mathcal{N}(0,I)$ indep of x_0

Reverse Process: $p_{\lambda(x_{t-1}|x_t)} = \mathcal{N}\left(\mu_{\lambda(x_t,t)}, \sigma_t^2 I\right)$ Prior: $p(x_T) = \mathcal{N}(0,I)$ Generate: sample x_T, iteratively denoise to x_0

Forward Posterior: $q(x_{t-1}|x_t,x_0) = \mathcal{N}\left(\tilde{\mu}_t, \tilde{\beta}_t I\right)$ $\tilde{\mu}_t = \frac{\sqrt{|(\alpha)_{t-1}}\beta_t}{1-|(\alpha)_t}x_0 + \frac{\sqrt{\alpha_t}(1-|(\alpha)_{t-1})}{1-|(\alpha)_t}x_t$ $\tilde{\beta}_t = \frac{(1-|(\alpha)_{t-1})\beta_t}{1-|(\alpha)_t}$ Key: given x_0, x_t, forward posterior is Gaussian (tractable)

ELBO & Loss: $\mathcal{L} = \mathrm{const} - \sum_{t=2}^T \underbrace{\mathrm{KL}\left(q(x_{t-1}|x_t,x_0)\|p_{\lambda(x_{t-1}|x_t)}\right)}_{L_t}$ Two Gaussians same var: $\mathrm{KL} \propto \|\mu_1 - \mu_2\|^2$

⭐**Noise Prediction**: Predict ε instead of μ (more stable): From $x_t = \sqrt{|(\alpha)_t}x_0 + \sqrt{1-|(\alpha)_t}\varepsilon$: $\tilde{\mu}_t = \frac{1}{\sqrt{\alpha_t}}\left(x_t - \frac{\beta_t}{\sqrt{1-|(\alpha)_t}}\varepsilon\right)$ **Simple Loss**: $L_{\mathrm{simple}} = \mathbb{E}_{t,x_0,\varepsilon}\left[\|\varepsilon - \varepsilon_{\lambda(x_t,t)}\|^2\right]$

Training Algo: Repeat: sample $x_0 \sim p_{\mathrm{data}}$, $t \sim \mathrm{Unif}\{1,...,T\}$, $\varepsilon \sim \mathcal{N}(0,I)$ $x_t = \sqrt{|(\alpha)_t}x_0 + \sqrt{1-|(\alpha)_t}\varepsilon$ $\nabla_\lambda\|\varepsilon - \varepsilon_{\lambda(x_t,t)}\|^2$

Sampling Algo: $x_T \sim \mathcal{N}(0,I)$ For $t = T,...,1$: $z \sim \mathcal{N}(0,I)$ if $t > 1$ else $z = 0$ $x_{t-1} = \frac{1}{\sqrt{\alpha_t}}\left(x_t - \frac{\beta_t}{\sqrt{1-|(\alpha)_t}}\varepsilon_{\lambda(x_t,t)}\right) + \sigma_t z$

Connection: $\varepsilon_{\lambda(x_t,t)} \approx -\sqrt{1-|(\alpha)_t}\nabla_{x_t}\log q(x_t)$ **Denoising = Score matching**

Variants: **LDM**: diffusion in VAE latent space, more efficient **DDIM**: deterministic sampling, fewer steps **Cond Gen**: $\varepsilon_{\lambda(x_t,t,c)}$, Classifier-Free Guidance: $\tilde{\varepsilon} = (1+w)\varepsilon_{\lambda(x_t,t,c)} - w\varepsilon_{\lambda(x_t,t)}$

QuickCheck:

- **VI**: Approx posterior via ELBO. Laplace `MAP`, Reparam for grad.
- **MCMC**: Sample posterior. MH accept/reject, Gibbs coordinate, Langevin uses ∇.
- **GP**: Prior over fncs, closed-form posterior. RBF smooth, Matérn tunable.
- **BNN**: Prior on weights, MC predictive. Aleatoric=data noise, Epistemic=model.
- **Active**: Max MI, BALD for disagreement, submodular⟶greedy$\left(1-\frac{1}{e}\right)$.
- **BO**: UCB balance explore/exploit, EI expected gain, Thompson sample.
- **BN**: DAG factorization, d-sep for indep, BP exact on trees.
- **KF**: Linear Gaussian, Kalman gain trades predict vs observe.
- **Diffusion**: Forward=noise, Backward=denoise, train predict ε.
- **On/Off**: On=SARSA,REINFORCE,PPO; Off=Q-learn,DQN,SAC
- **Bellman**: $V = R + \gamma PV$;