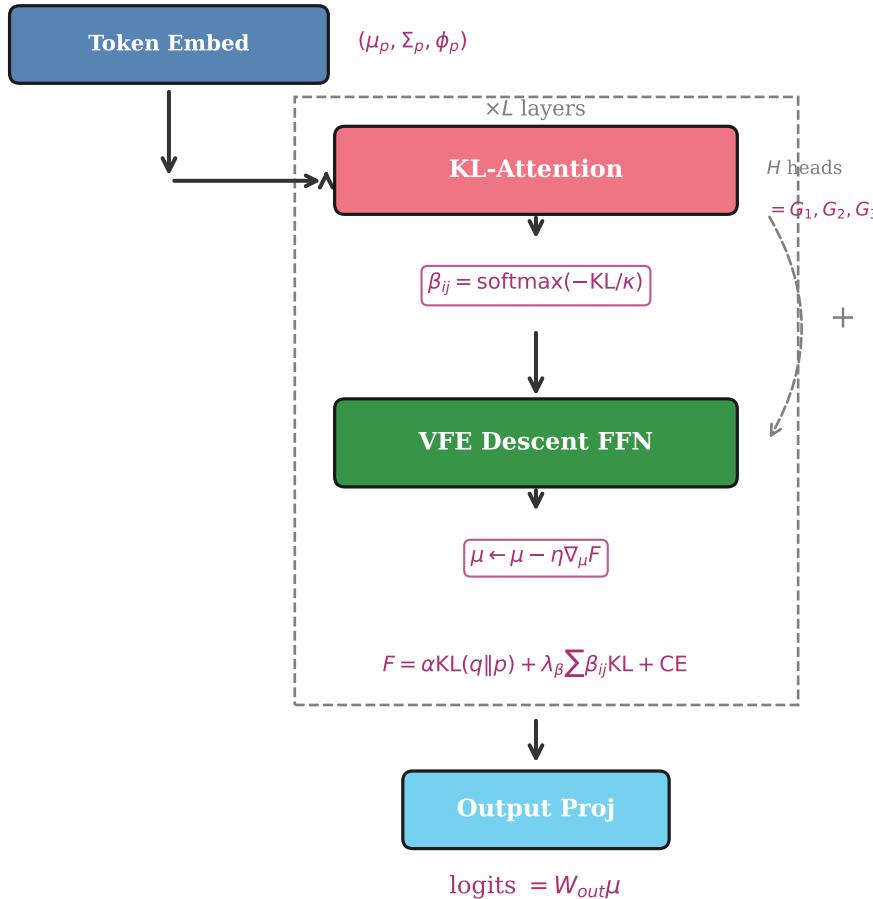
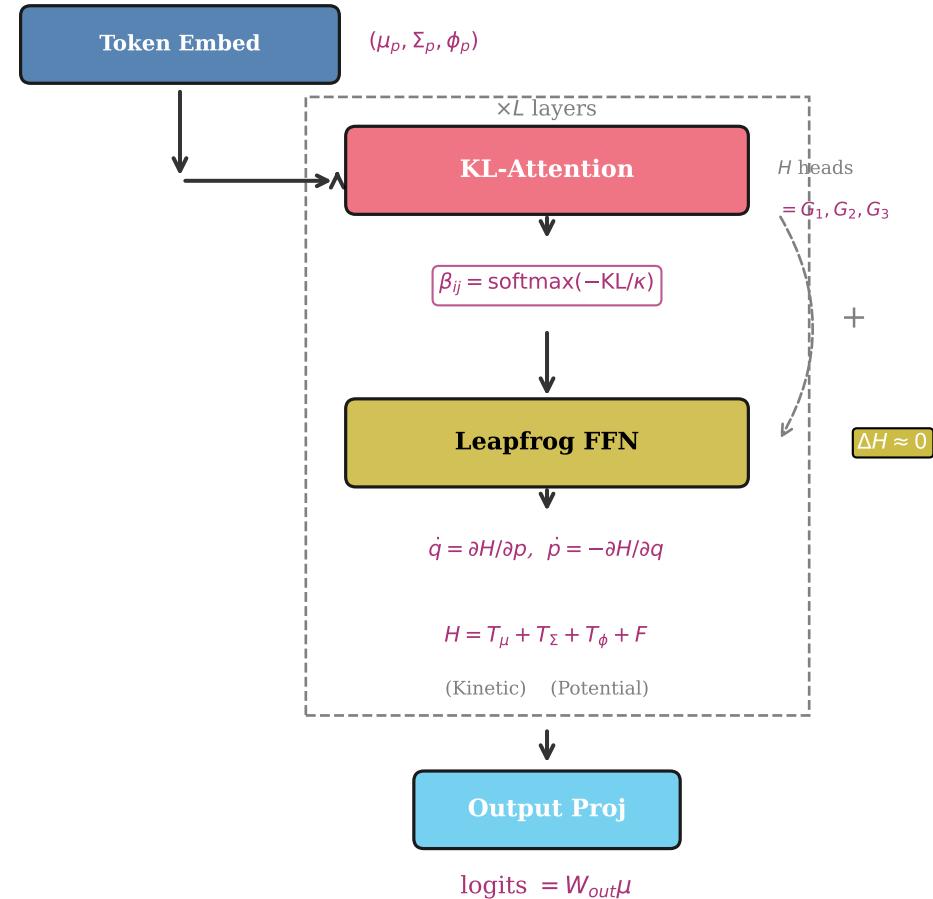


Gauge Transformer Architectures

(a) Gauge-VFE Transformer



(b) Gauge-Hamiltonian Transformer



Both architectures use KL-divergence attention with gauge-equivariant parallel transport.
VFE descent minimizes free energy via gradient flow. Hamiltonian dynamics conserves energy via symplectic integration.