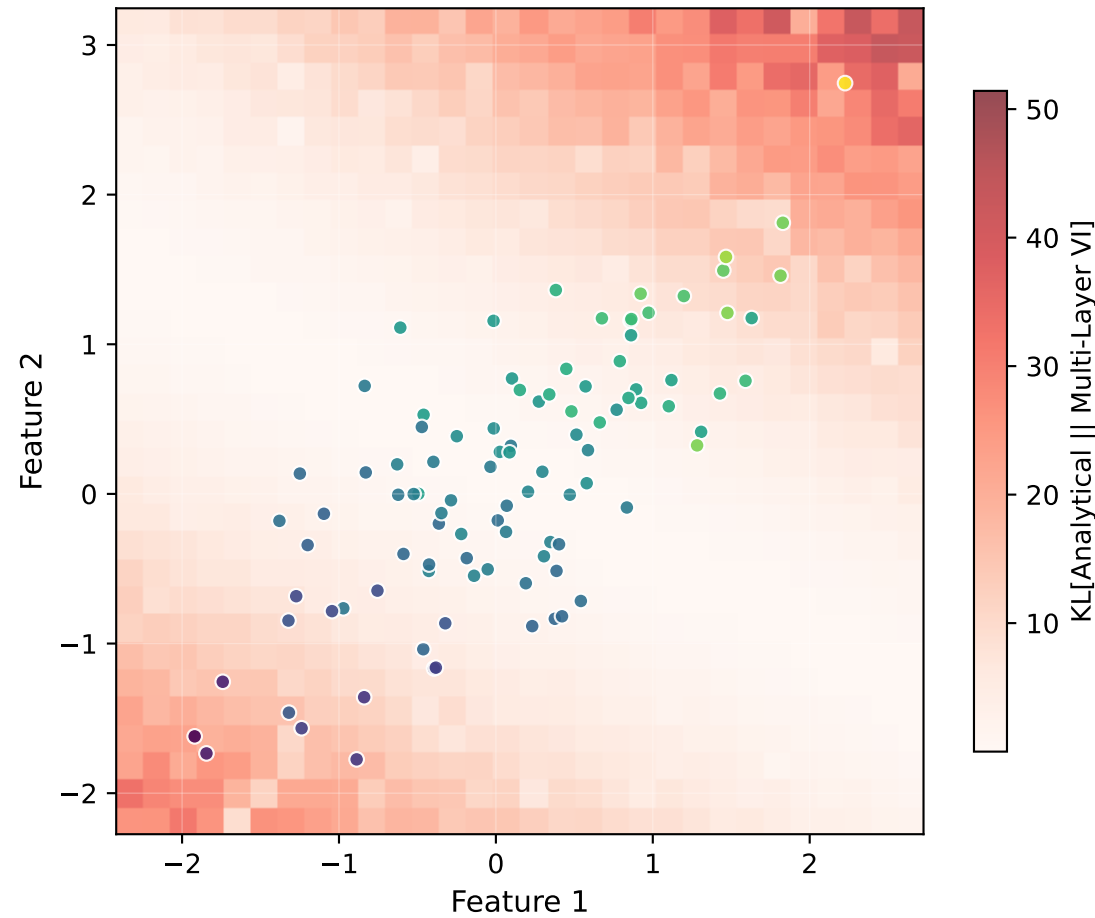


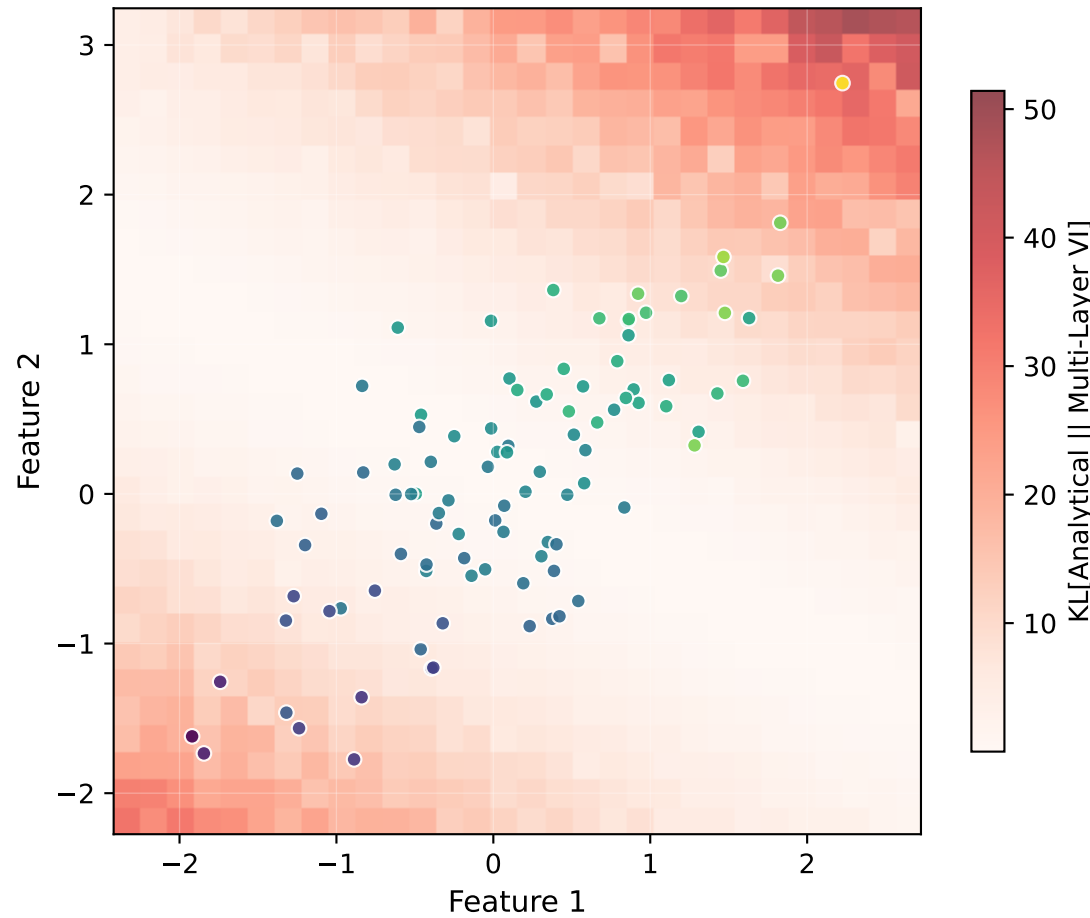
Multi-Layer KL Divergence Heatmaps vs Analytical Single-Layer

Architecture: $3 \rightarrow 4 \rightarrow 4 \rightarrow 4 \rightarrow 4 \rightarrow 4 \rightarrow 4 \rightarrow 4 \rightarrow 4 \rightarrow 1$

**Standard VI
Standard**



**Cold VI
Cold (T=0.8)**



**Tempered VI
Tempered ($\lambda=0.8$)**

