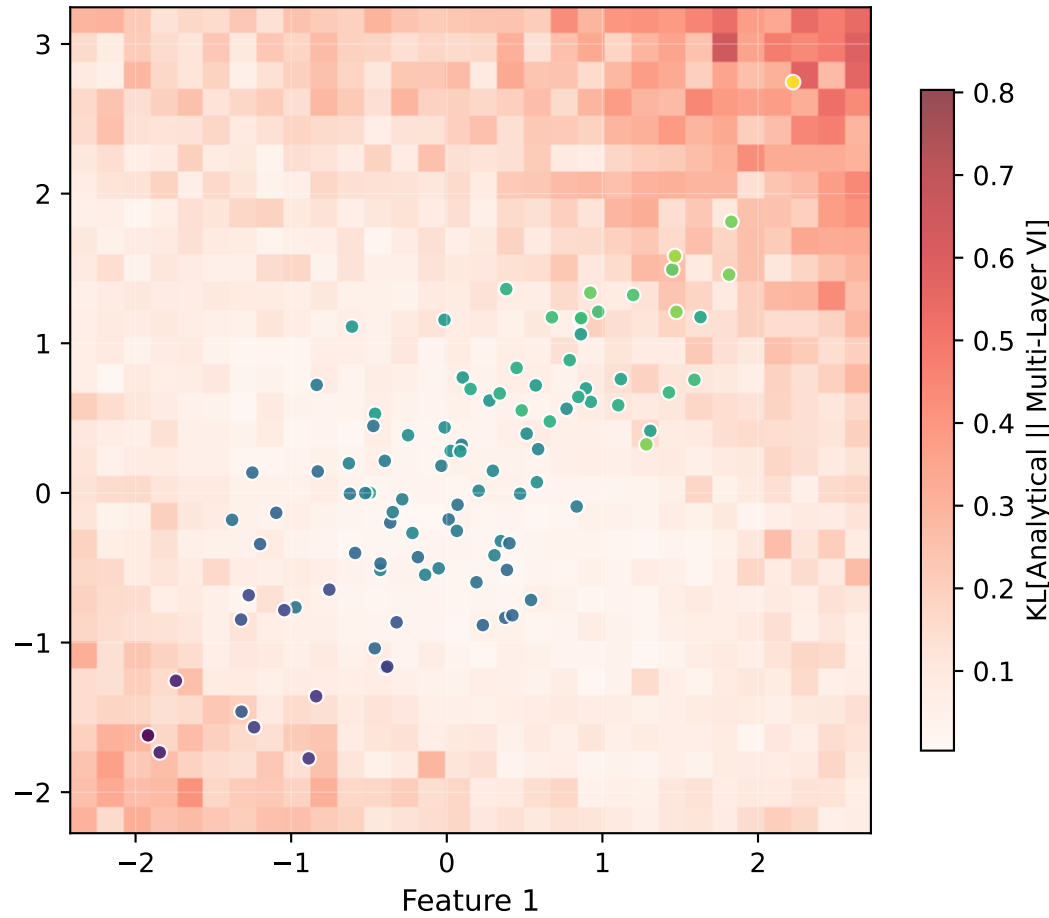


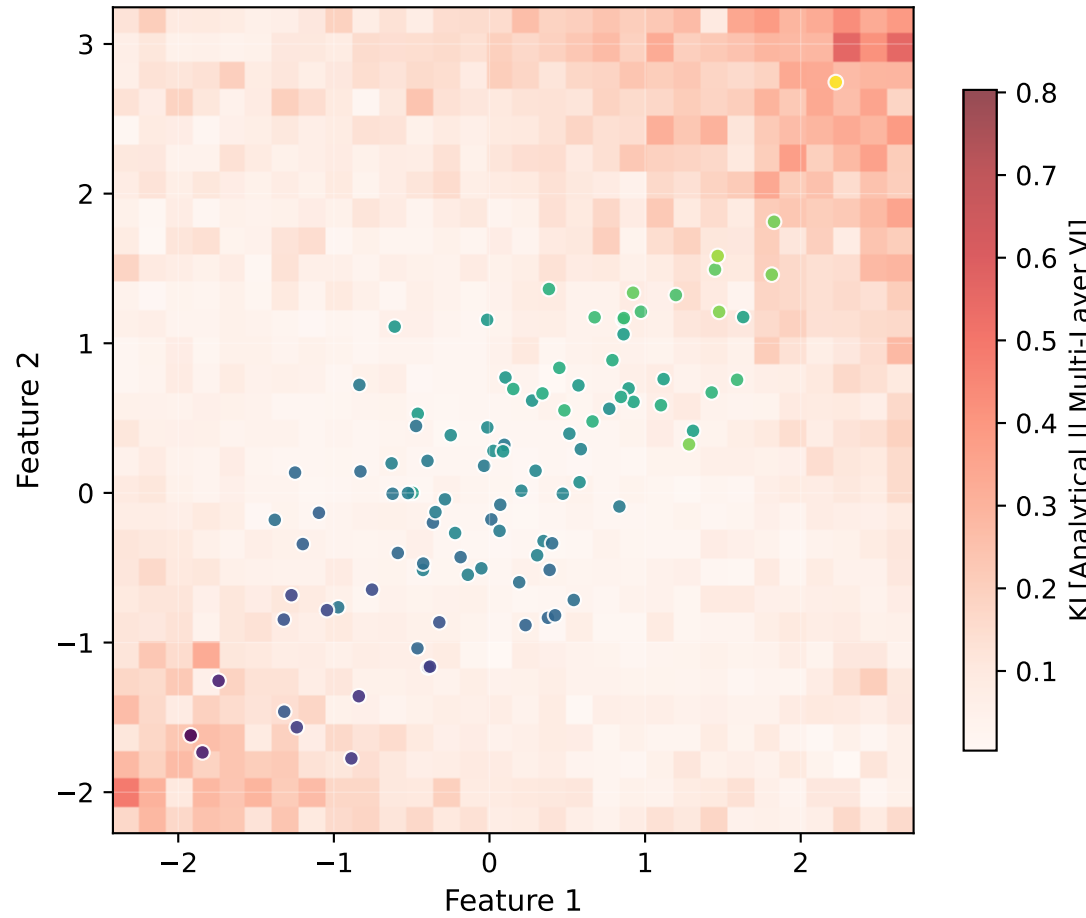
Multi-Layer KL Divergence Heatmaps vs Analytical Single-Layer

Architecture: 3 → 16 → 16 → 1

Standard VI
Standard



Cold VI
Cold (T=0.8)



Tempered VI
Tempered ($\lambda=0.8$)

