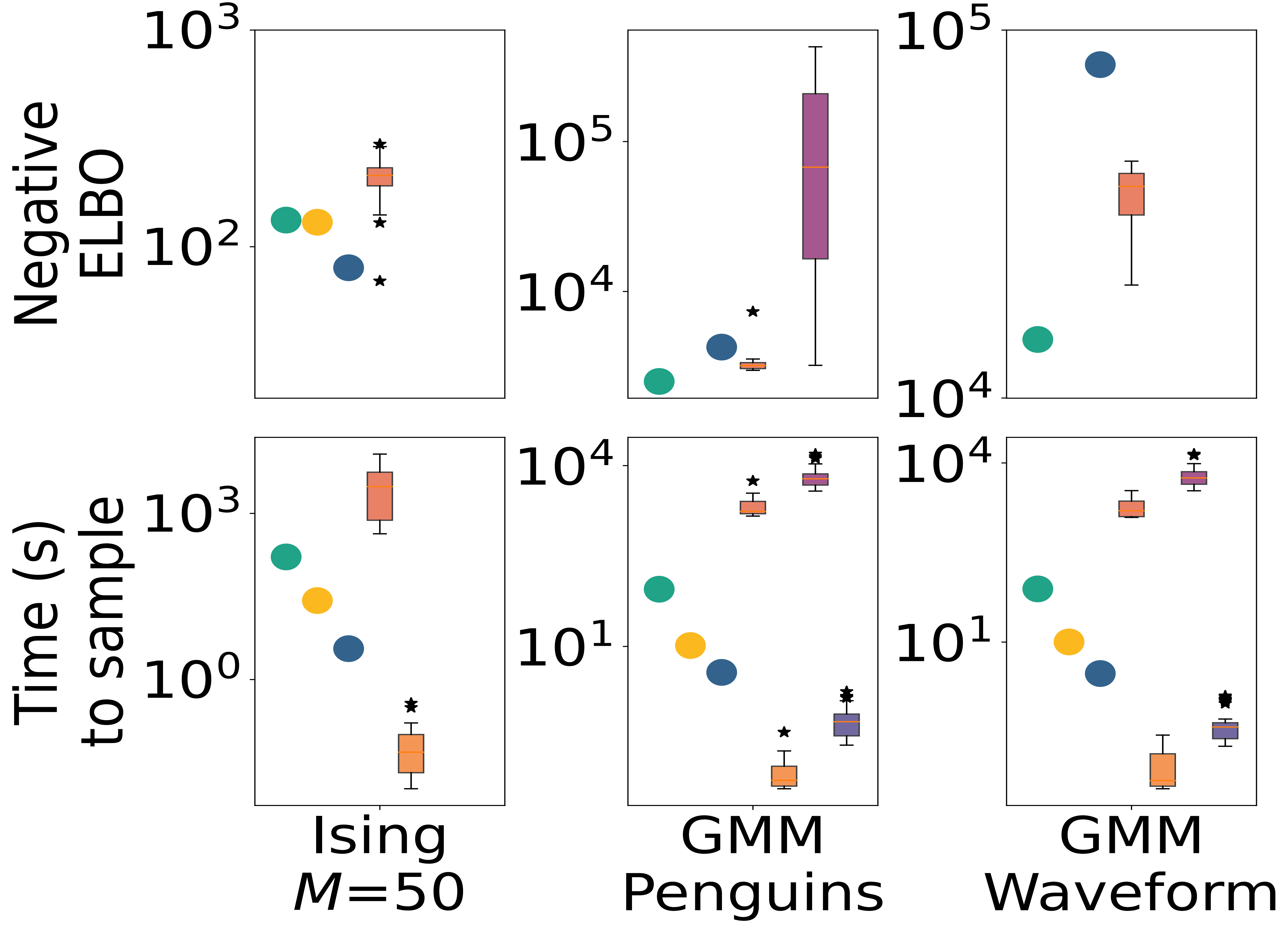
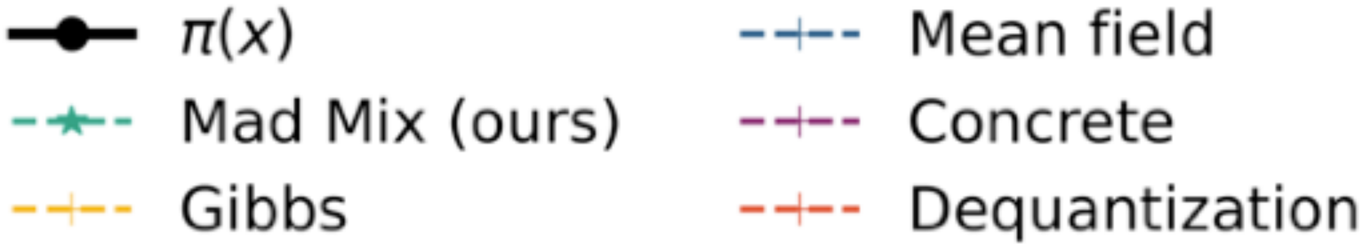
Quantitative results, real-world data

9



Negative ELBO
Time (s) to sample
10^3
10^2
10^5
10^4
10^5
10^4
10^3
10^0
10^4
10^1
10^4
10^1
Ising M=50
GMM Penguins
GMM Waveform


π(x)
Mad Mix (ours)
Gibbs
Mean field
Concrete
Dequantization

lower is better

"distance" to p

CPU time

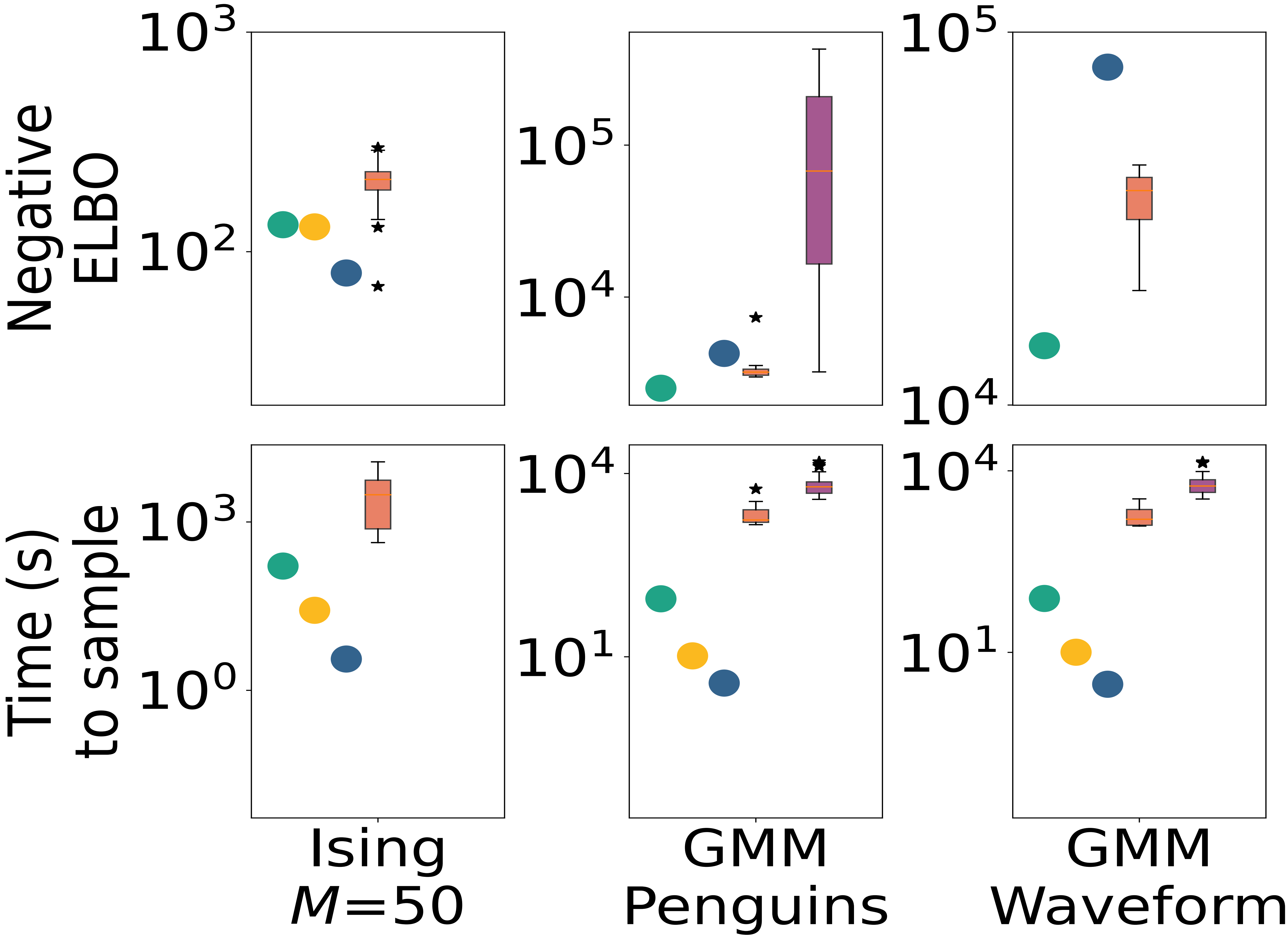
purely discrete

synthetic ($d=2^{50}$)

Gaussian mixture model

Palmer penguins (d=1,044) and waveform PCA (d=918) data sets

could not fit Concrete in this case!



Gaussian mixture model

Palmer penguins (d=1,044) and waveform PCA (d=918) data sets

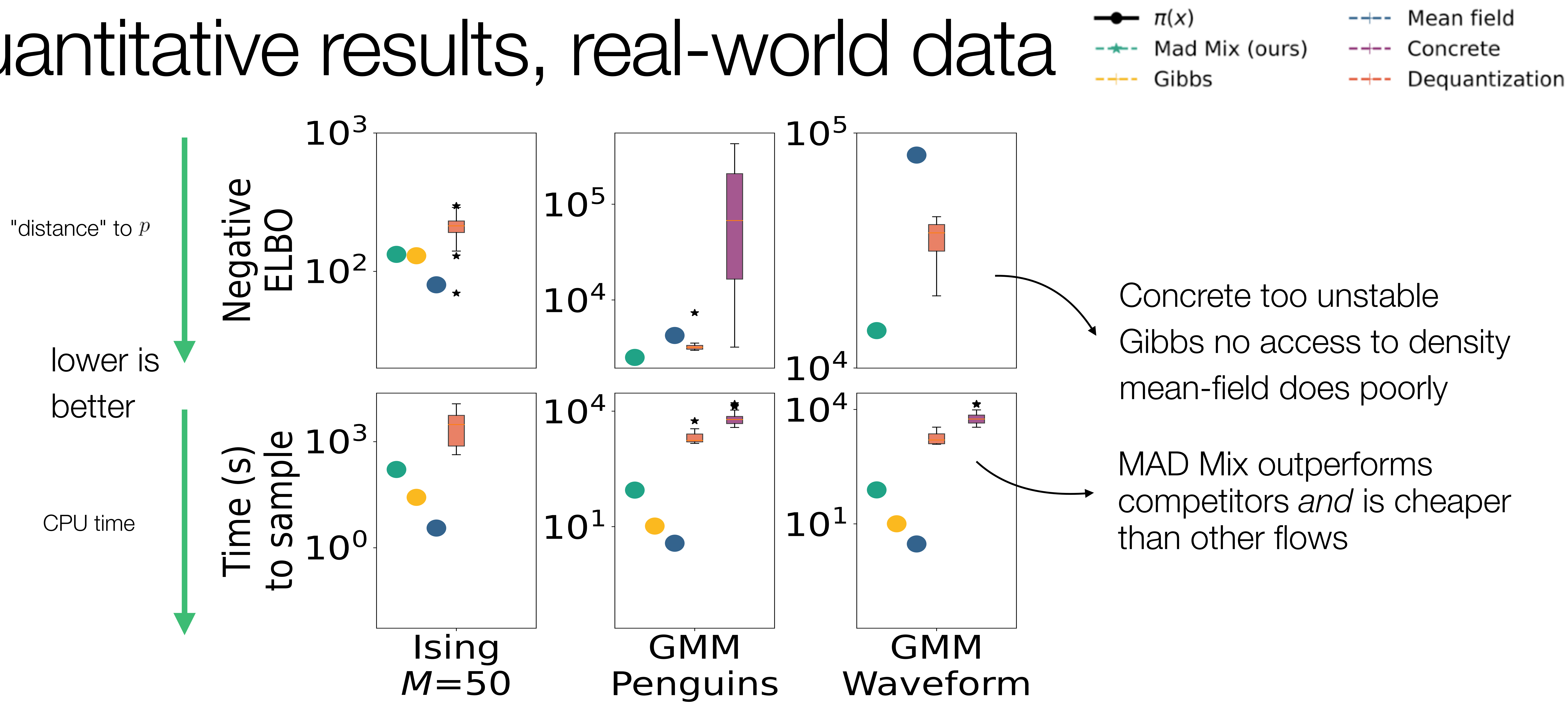
Concrete too unstable

Gibbs no access to density

mean-field does poorly

MAD Mix outperforms competitors *and* is cheaper than other flows setup: $N \approx 500$, $\xi = \pi/16$ for MAD Mix; 5K iterations for Gibbs (+20K burn-in); wide architecture search for continuous-embedding flows (concrete & dequantization)

Quantitative results, real-world data



Concrete too unstable
Gibbs no access to density
mean-field does poorly

MAD Mix outperforms competitors *and* is cheaper than other flows

purely discrete synthetic (d=2^{50})

Gaussian mixture model
Palmer penguins (d=1,044) and waveform PCA (d=918) data sets

setup: $N \approx 500$, $\xi = \pi/16$ for MAD Mix; 5K iterations for Gibbs (+20K burn-in); wide architecture search for continuous-embedding flows (concrete & dequantization)

Conclusion

Gian Carlo Diluvi
PhD student @ UBC
https://giankdiluvi.github.io

MAD Mix: measure-preserving and discrete MixFlows

- inference for discrete posteriors without continuous-embedding
- state-of-the-art performance with orders of magnitude less compute and tuning effort