

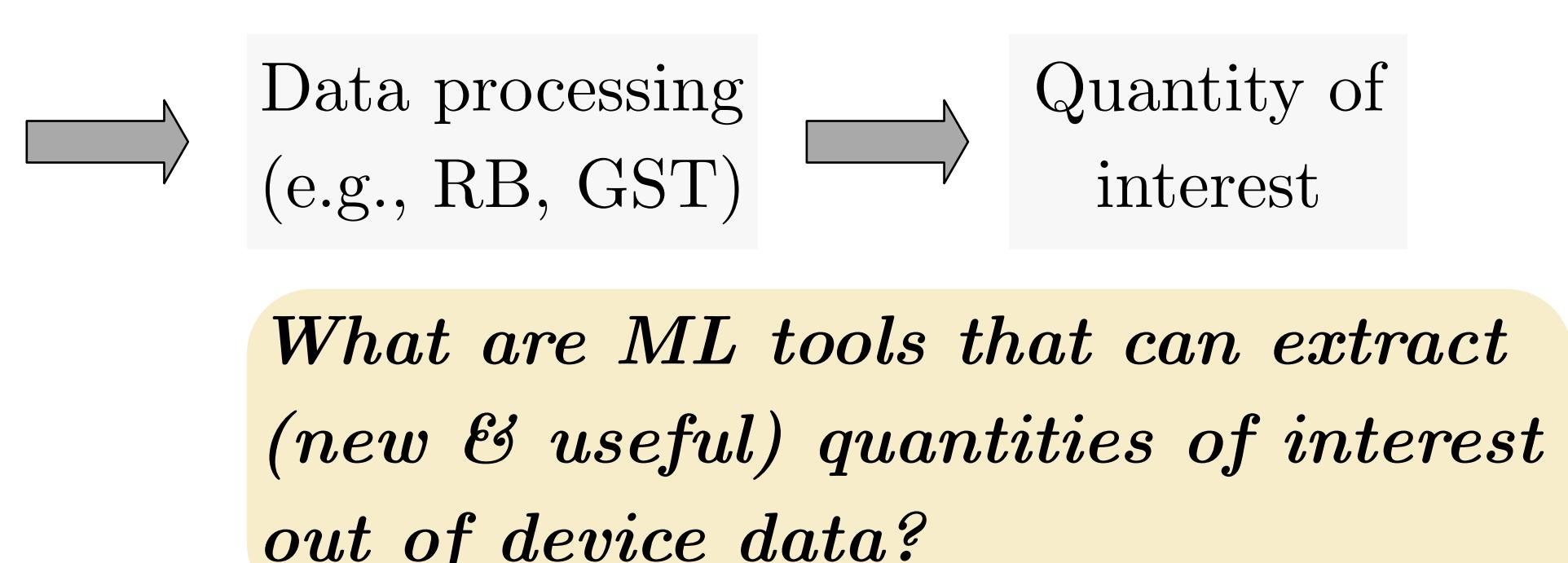
High-Accuracy Classification of Single-Qubit Noise via Machine Learning

Travis L Scholten, University of New Mexico & Sandia National Laboratories

Machine learning (ML) could help us develop new tomographic protocols.

Standard tomographic protocols may not scale as the number of qubits in a device increases, nor is the path to developing new ones always clear.

My Data Circuit(j)	PC(j)
{}	1
Gx	.44
Gy	.45
GxGx	.9
GxGxGx	.68
GyGyGy	.70



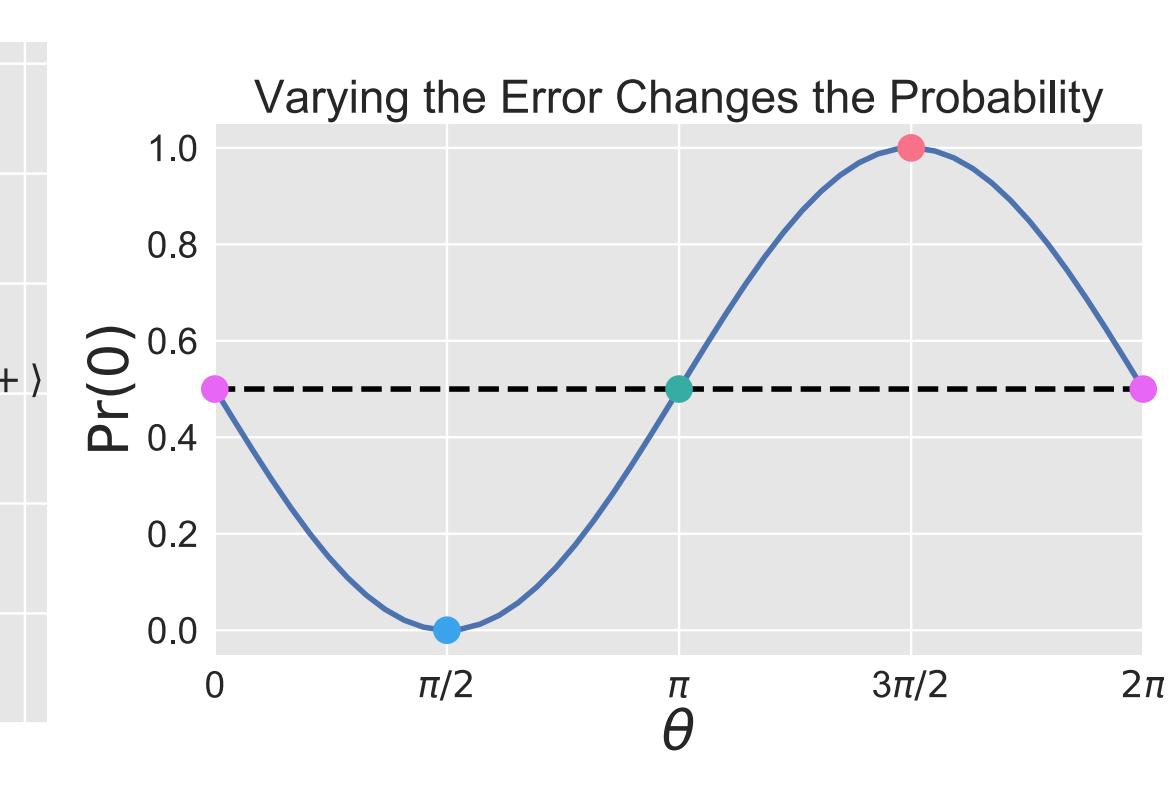
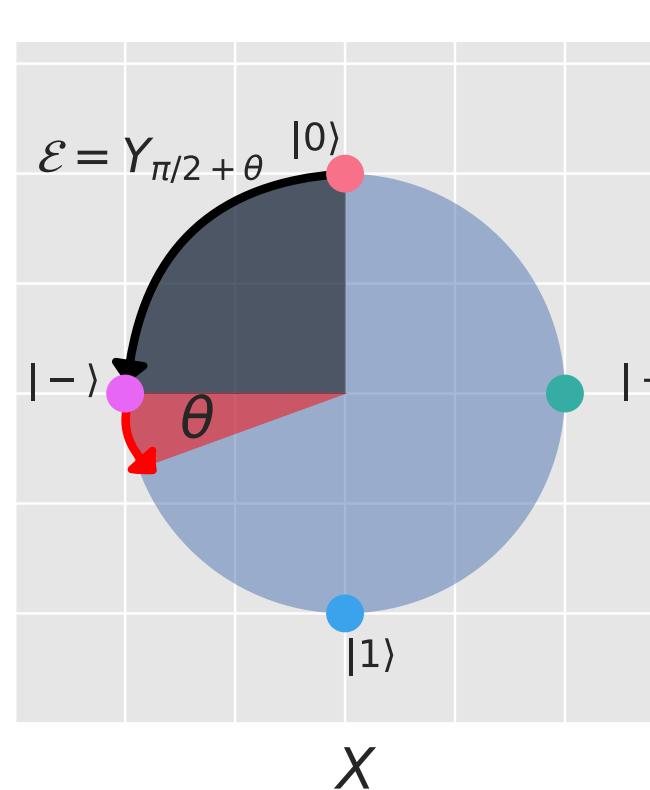
Noise affects circuits & manifests itself in their probabilities.

Noise affects circuit probabilities, and data sets (collections of circuits & their estimated probabilities) encode information about noise.

Example: over-rotation error of a single-qubit gate $|0\rangle \xrightarrow{Y_{\pi/2}} |-\rangle$ $\Pr(0) = \text{Tr}(|0\rangle\langle 0|\mathcal{E}(|0\rangle\langle 0|)) = \frac{1}{2}(1 - \sin \theta)$

(The circuit we write down)

(Noise affects outcome probability)



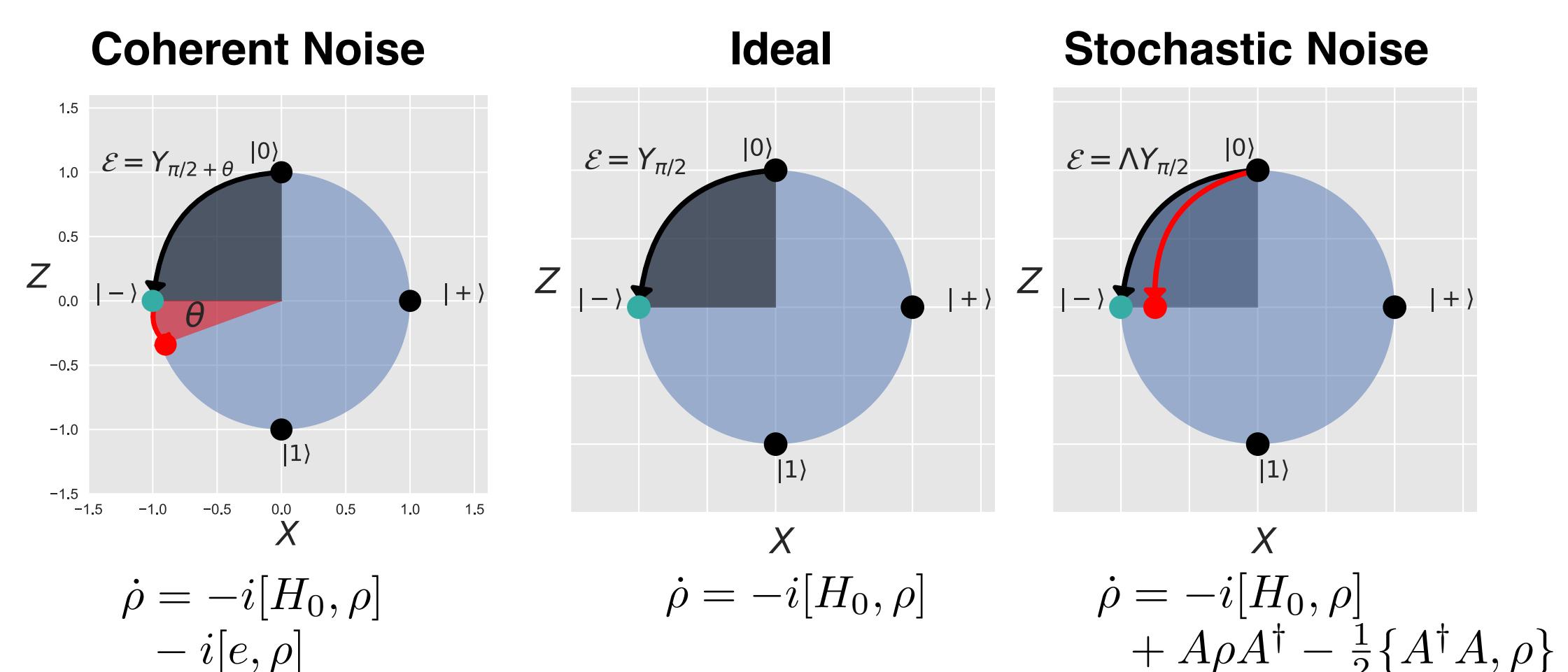
Gate set tomography (GST) data sets are useful in this regard, because the circuits in the data set amplify all possible noise.

Overview: machine learning (ML) to classify coherent & stochastic single-qubit noise

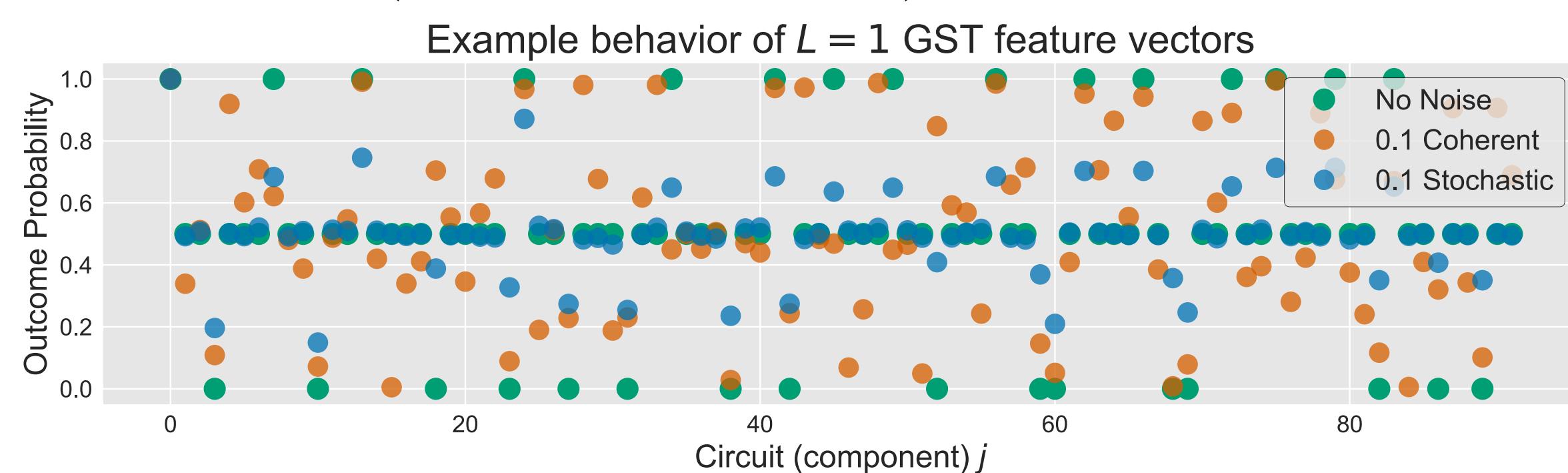
Telling the difference between coherent noise ("systematic" errors) and stochastic noise ("random" errors) would be useful in debugging a qubit.

As the circuits used for GST amplify all possible noise, GST data sets are useful for classifying these kinds of noise.

GST datasets (with counts normalized to [0, 1]) can be represented as feature vectors in a feature space.



The feature vectors encode information about noise in their components (circuit probabilities).

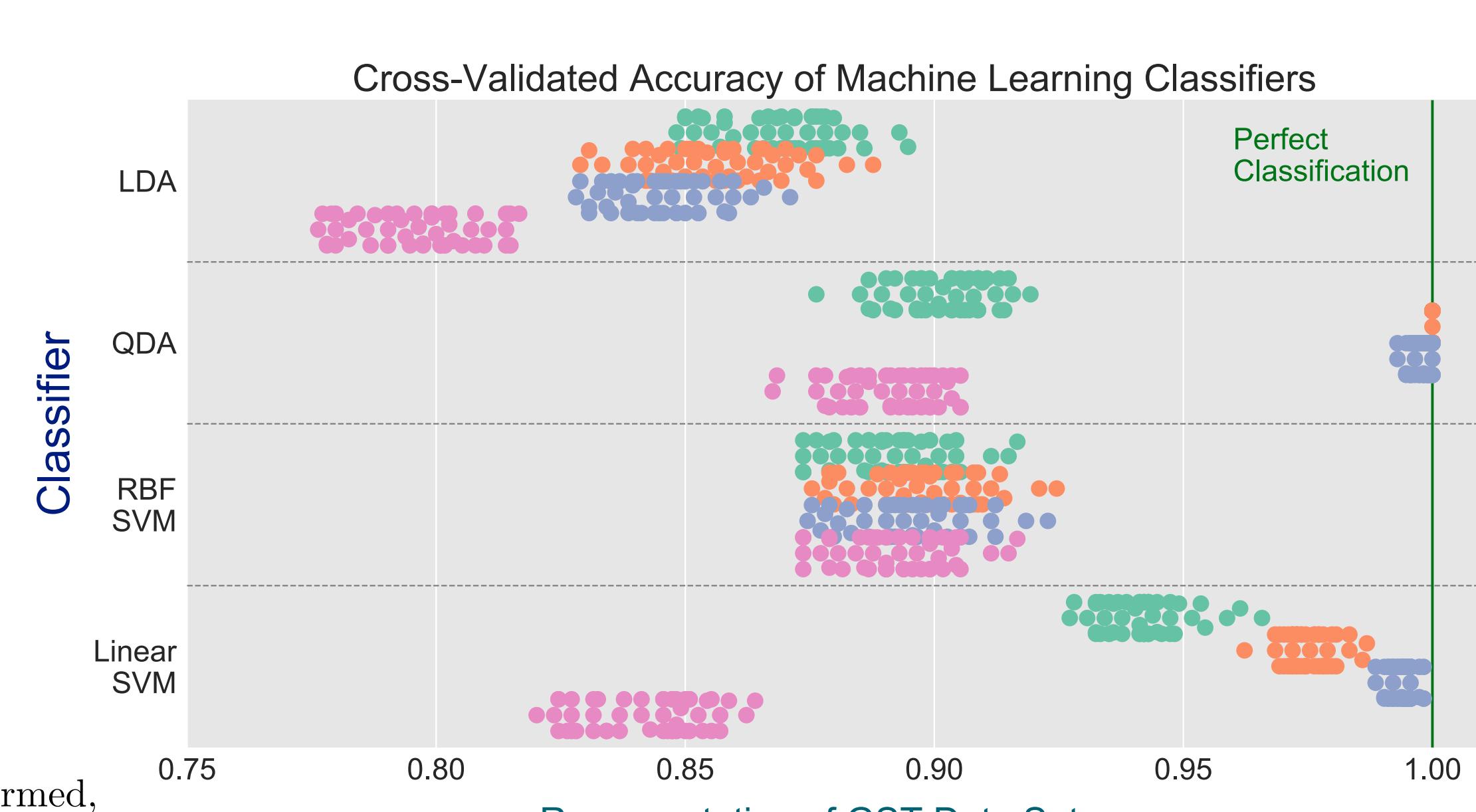


GST feature vectors are "quartically separable". Why?

We investigated 4 ML classifiers and 4 representations of GST data sets.

Classification accuracy depends on the **classifier** and the **representation** of the data set. (See "In Depth" below.)

A 50-fold shuffle-split cross-validation was performed, with 10% of the data used for testing. Data consists of 300 random instances for each of 2 noise types and 19 unique noise strengths in $[10^{-4}, .5]$.



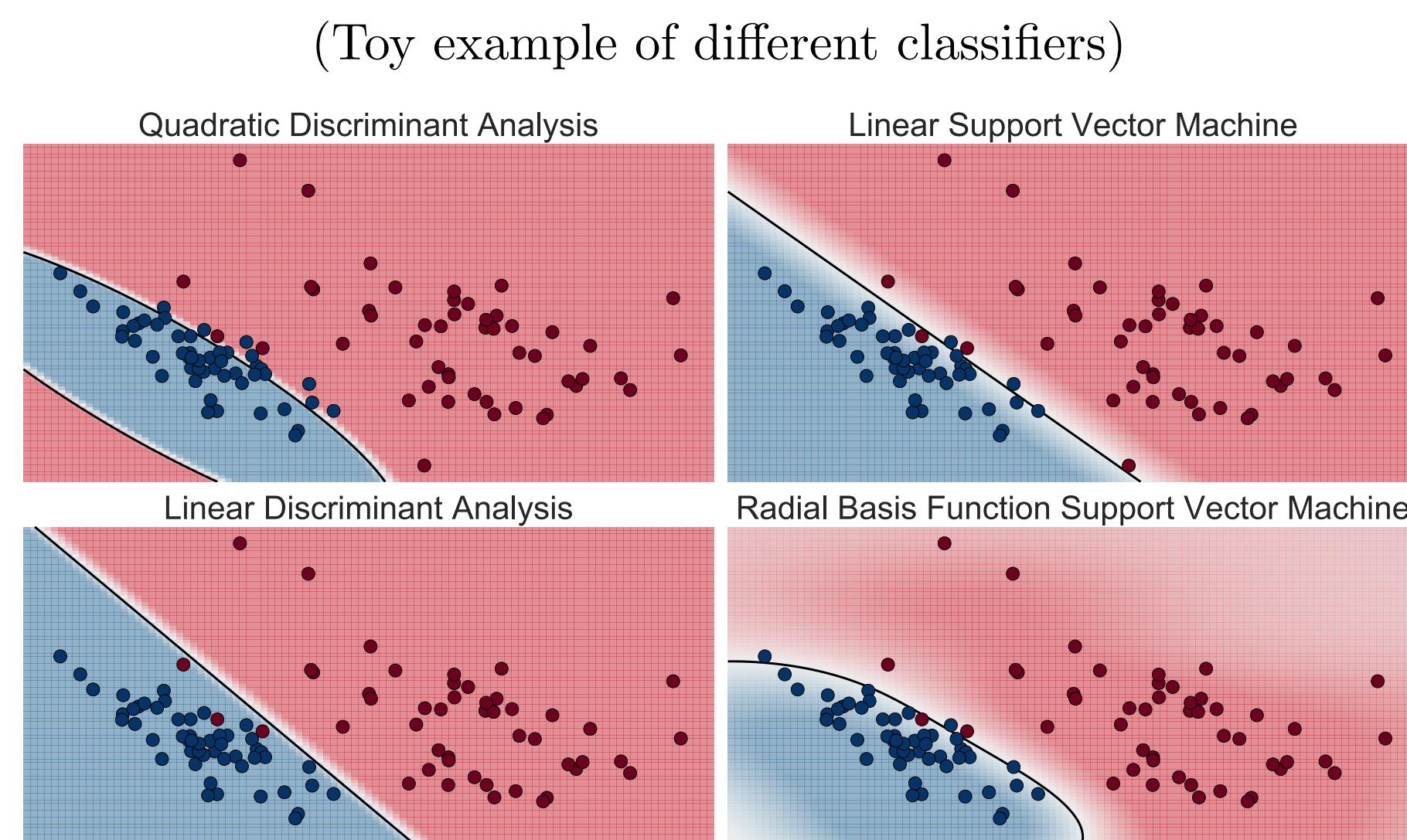
In depth: ML classifiers and representations of GST data sets

Quadratic Discriminant Analysis (QDA)

Bayesian solution for two-class classification assuming data for each class is Gaussian, with unequal covariance matrices.

Linear Discriminant Analysis (LDA)

Similar idea as QDA, but assumes covariances for each class are the same.



Linear Support Vector Machine (Linear SVM)

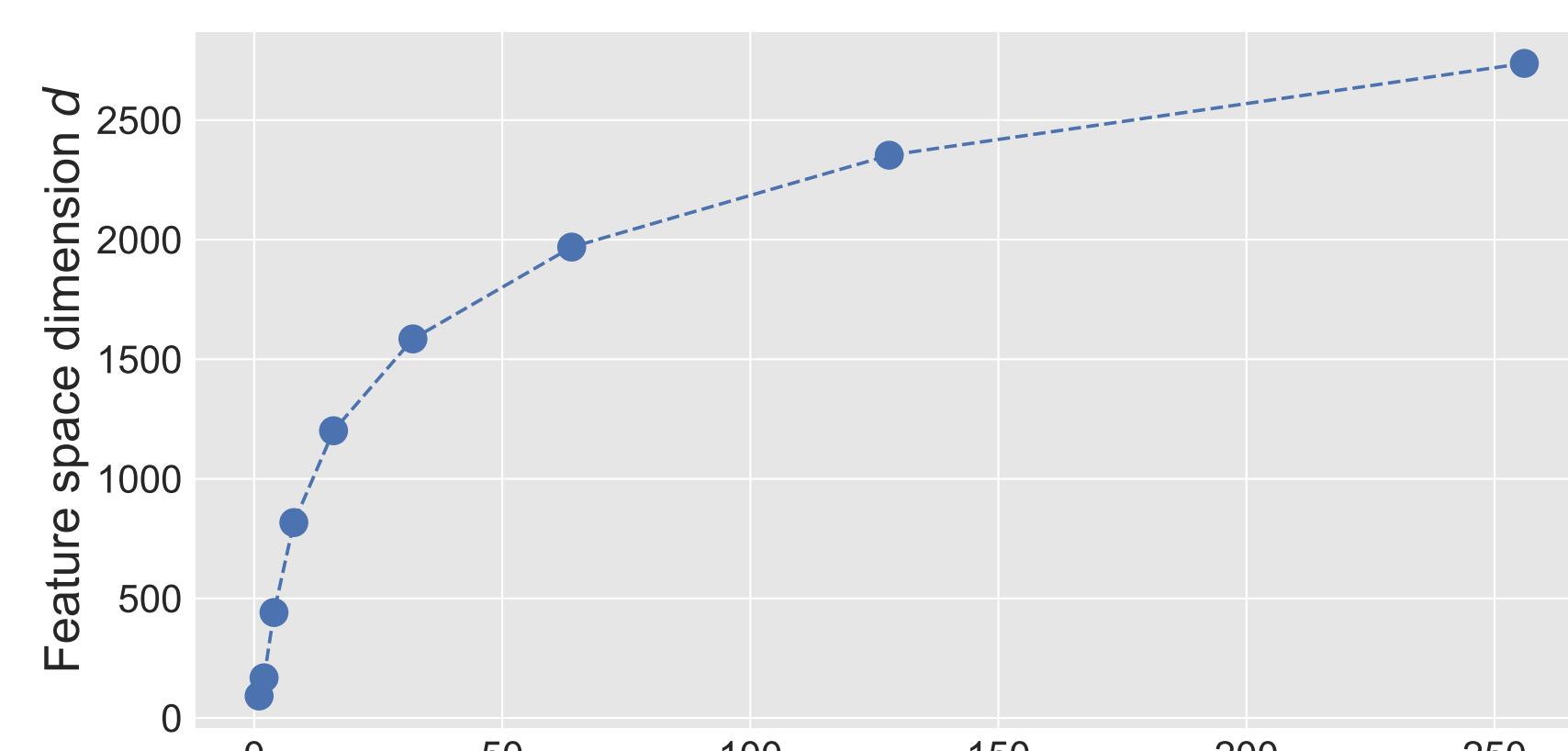
Finds a hyperplane by solving an optimization problem that trades off between accuracy on known data and accuracy on future data ("soft-margin").

Radial Basis Function Support Vector Machine (RBF SVM)

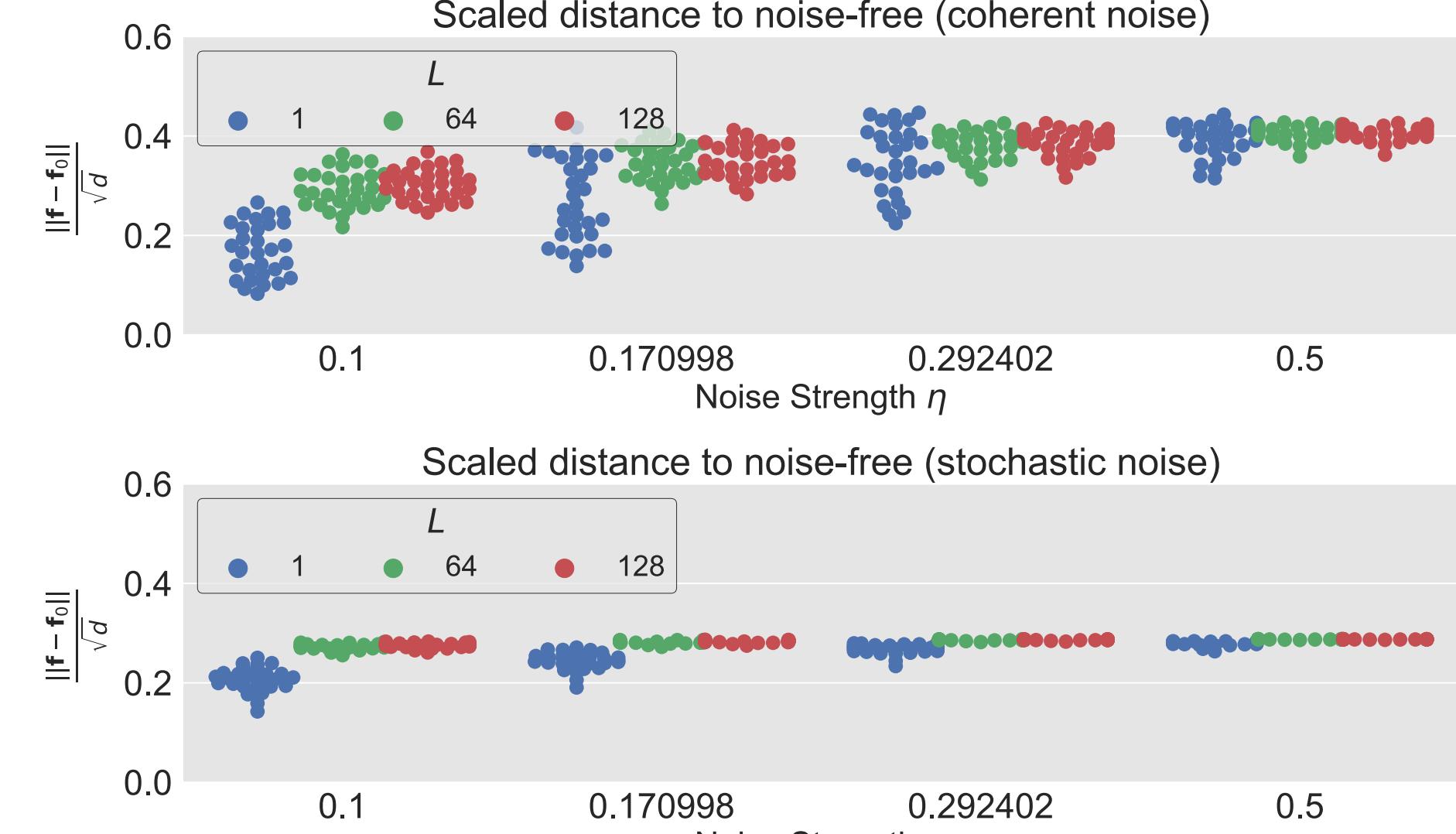
Similar idea as linear SVM, but classification depends on a Gaussian function of the data.

In depth: structure of (canonical) GST feature space

L controls the number of circuits in a GST data set, which in turn sets the dimension of the feature space, d .



As L increases, the noisy feature vectors "fill out" the feature space, moving away from the noise-free feature vector \mathbf{f}_0 .



Principal Component Analysis (PCA)

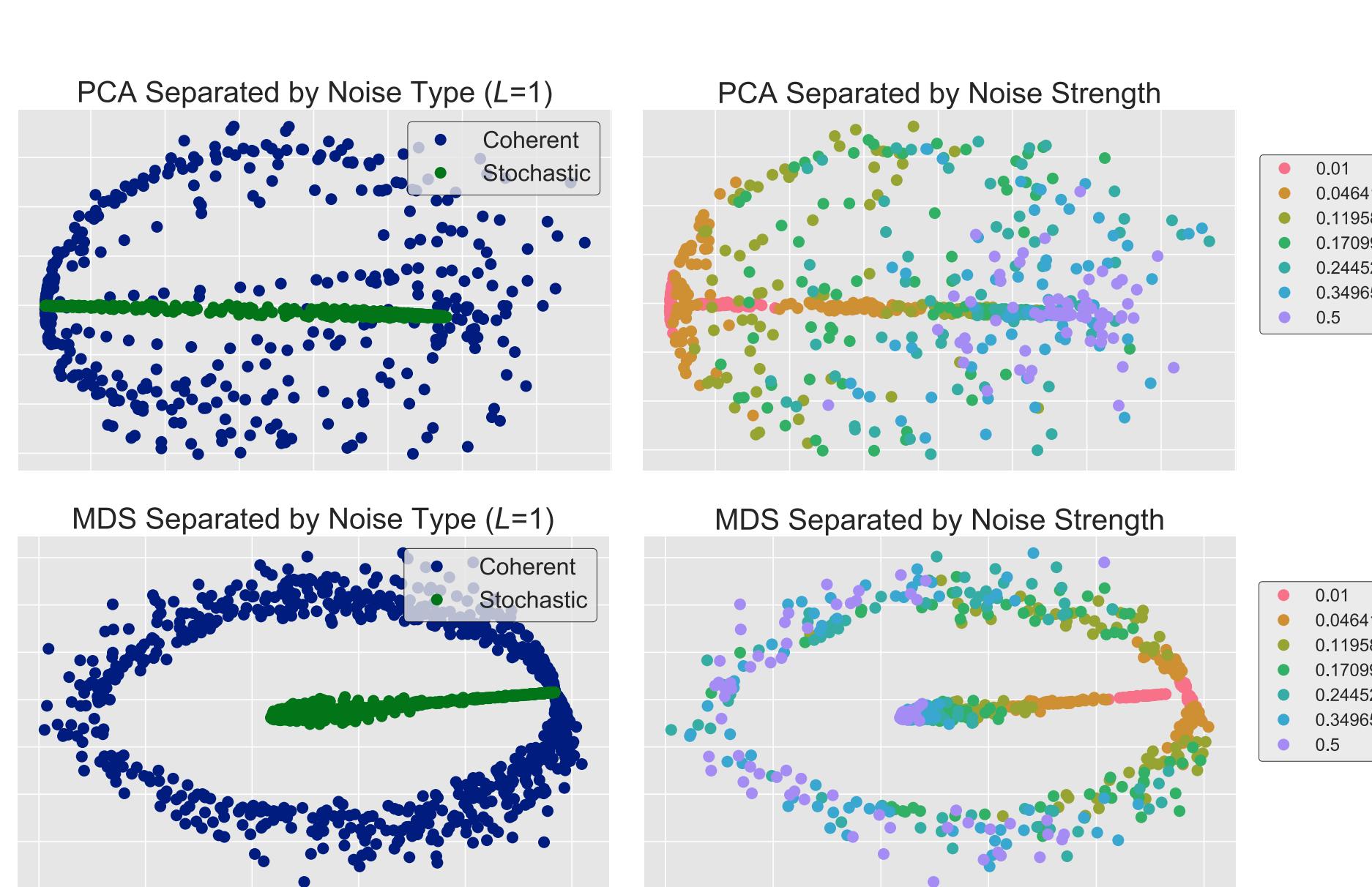
Representation captures maximal variance in the data.

Multidimensional scaling (MDS)

Representation maximally preserves pairwise distances.

Different representations of the data give rise to new feature spaces with their own structure.

- (Canonical) $\mathbf{f} = (f_1, f_2, \dots, f_d)$
- (Squares of all components; ALL_SQ) $\mathbf{f} \rightarrow \mathbf{f} \oplus \sum_j f_j^2$
- (All unique pairwise products; ALL_PP) $\mathbf{f} \rightarrow \mathbf{f} \oplus \sum_{j,k} f_j f_k$
- (All unique pairwise products & squares; P_QUART) $\mathbf{f} \rightarrow \mathbf{f} \oplus \sum_{j,k} f_j f_k + \sum_{j,k} f_j f_k$



Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

