University of Stuttgart
Institute for Theoretical Physics I

Application of machine learning to find exceptional points

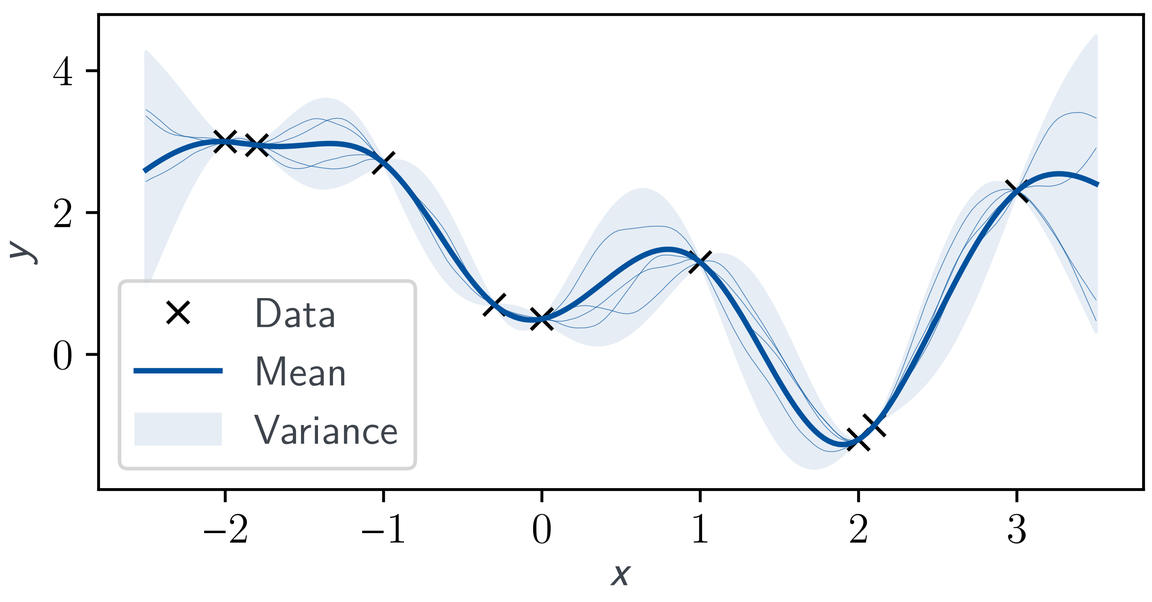
Patrick Egenlauf, Patric Rommel, and Jörg Main

Introduction

At an exceptional point (EP), both the eigenvalues and eigenvectors in non-hermitian Hamiltonians degenerate. EPs can occur in Rydberg systems in external electric and magnetic fields for specific field strengths. For excitons in cuprous oxide they exist at experimentally accessible field strengths. In general, such systems depend on a parameter $\kappa \in \mathbb{C}$, which corresponds to the fields ($\mathrm{Re}(\kappa) \equiv B$ and $\mathrm{Im}(\kappa) \equiv E$) in physical systems. The imaginary part of the complex eigenvalue describes the lifetime of the quasi-bound state (resonance). Due to the very time consuming calculations of eigenvalues for a physical system, a machine learning method is presented, which aims to minimize the number of diagonalizations by estimating the EP.

H. Cartarius, J. Main, and G. Wunner, PRL **99**, 173003 (2007)

Gaussian Process Regression (GPR)



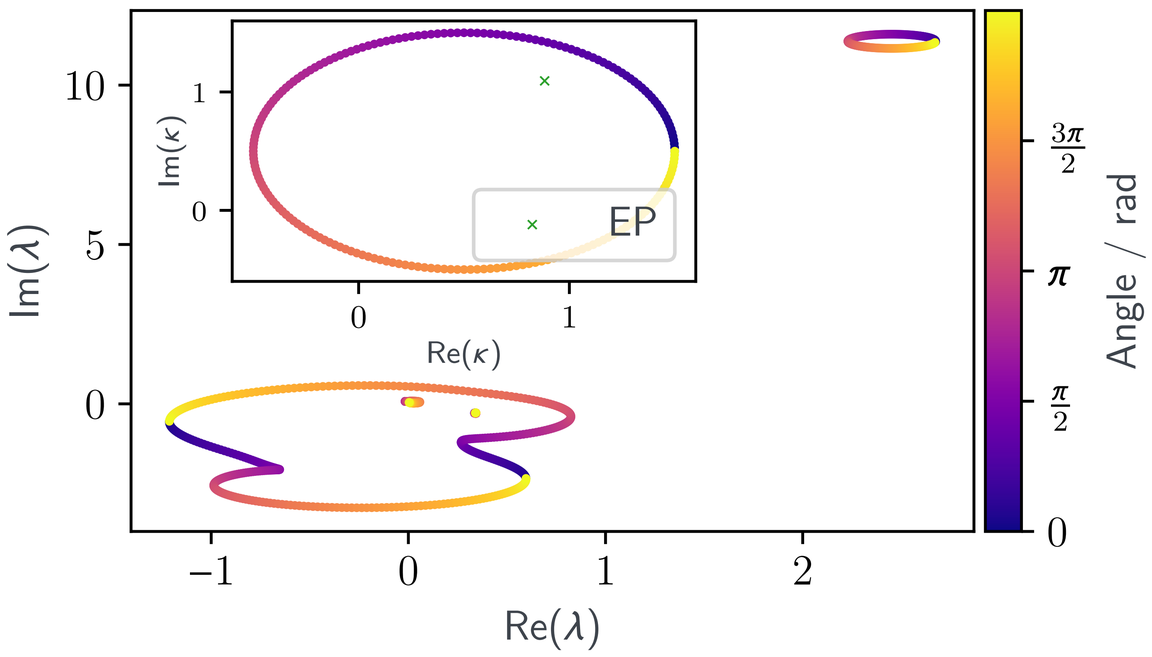
A GPR model describes a probability distribution over possible functions that fit a set of points. The regression function modeled by a multivariate Gaussian is given as $P(\mathbf{f}|X) = \mathcal{N}(\mathbf{f}|\boldsymbol{\mu}, \mathbf{K})$ with $X = [\mathbf{x}_1, \ldots, \mathbf{x}_n]$ the known data points, $\mathbf{f} = [f(\mathbf{x}_1), \ldots, f(\mathbf{x}_n)]$ the function values, $\boldsymbol{\mu} = [m(\mathbf{x}_1), \ldots, m(\mathbf{x}_n)]$ the mean functions, and $K_{ij} = k(\mathbf{x}_i, \mathbf{x}_j)$ the kernel matrix and function. This kernel or covariance function represents all possible functions through the data points and specifies their properties, such as smoothness. Predictions for new data points X_* can be made via

$$\mathbf{f}_*|X_*, X, \mathbf{f} \sim \mathcal{N}(\mathbf{K}(X_*, X)\mathbf{K}(X, X)^{-1}\mathbf{f}, \mathbf{K}(X_*, X_*) - \mathbf{K}(X_*, X)\mathbf{K}(X, X)^{-1}\mathbf{K}(X, X_*)),$$

not only predicting a value for $\mathbf{f}_*$ but also providing a variance (uncertainty) for that result.

C. E. Rasmussen and C. K. I. Williams, *Gaussian processes for machine learning*, (Nov. 23, 2005)

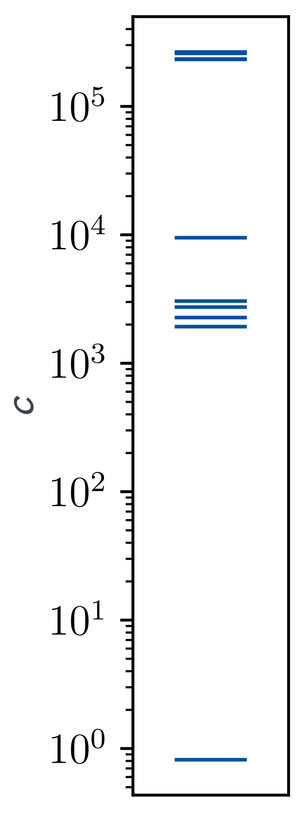
5×5 Matrix Model



A circulation in the parameter space ($\kappa \in \mathbb{C}$) around the EP results in an exchange behavior of the two complex eigenvalues belonging to the EP. These two eigenvalues can be used to train a GPR with κ as known data points and $p = (\lambda_1 - \lambda_2)^2$, $s = 1/2\,(\lambda_1 + \lambda_2)$ as function values at the respective κ-points. Due to the above mentioned degeneracy the EP should be at $p = 0$.

Iterative Searching Process

The first prediction is used to calculate the exact eigenvalues at this point. Depending on the system, the total number of eigenvalues is bigger than two, which is why the two eigenvalues belonging to the EP need to be selected.



Therefore p and s are calculated for all possible eigenvalue pairs and compared via a Gaussian

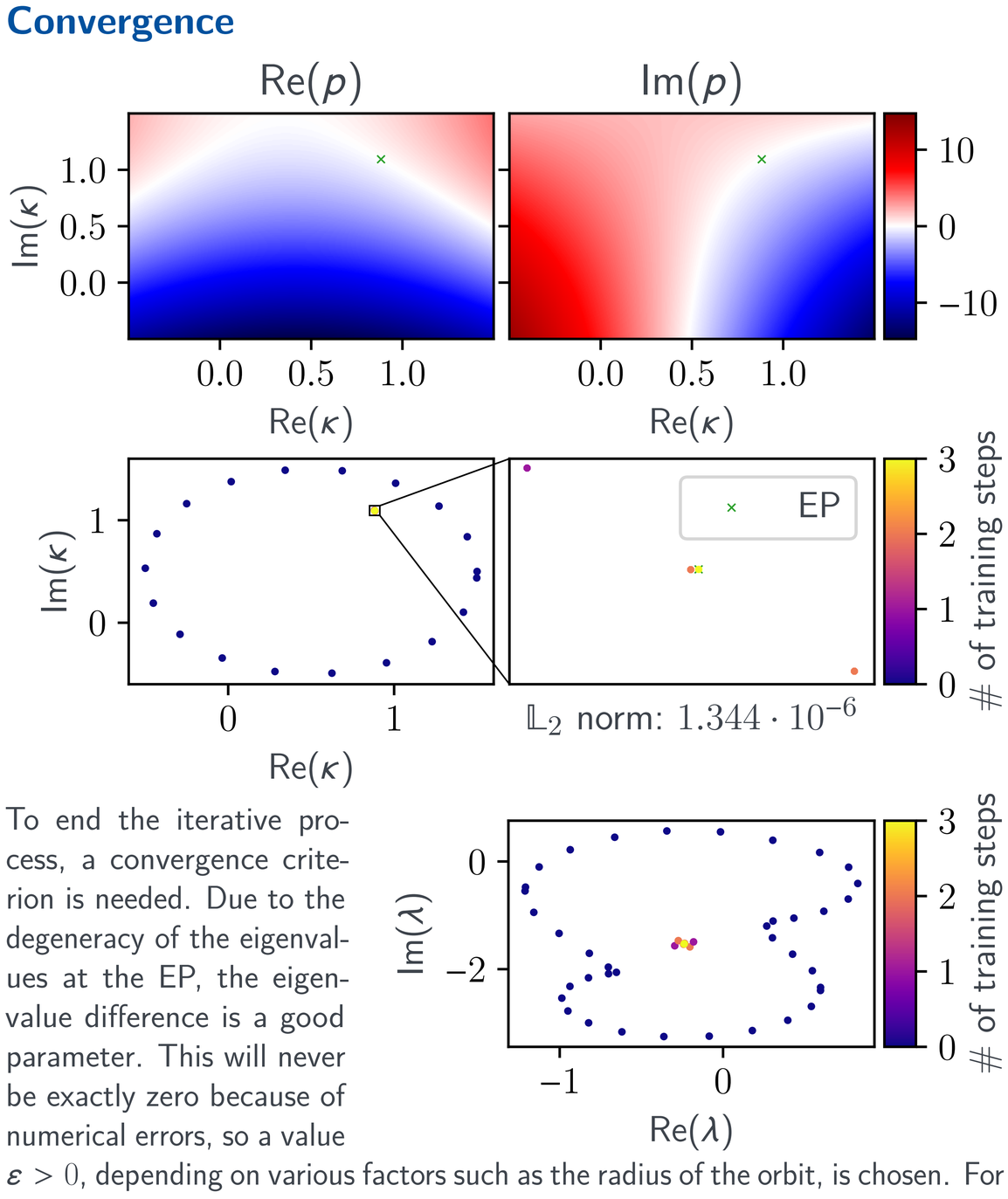
$$\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right),$$

where μ is the respective value predicted by the GPR model. This yields a similarity measure

$$c = \frac{(\mathrm{Re}(p) - \mathrm{Re}(p_\mathrm{m}))^2}{2\sigma^2_{\mathrm{p,Re}}} + \frac{(\mathrm{Im}(p) - \mathrm{Im}(p_\mathrm{m}))^2}{2\sigma^2_{\mathrm{p,Im}}} + \frac{(\mathrm{Re}(s) - \mathrm{Re}(s_\mathrm{m}))^2}{2\sigma^2_{\mathrm{s,Re}}} + \frac{(\mathrm{Im}(s) - \mathrm{Im}(s_\mathrm{m}))^2}{2\sigma^2_{\mathrm{s,Im}}},$$

with σ_i being the variance provided by the GPR model for the predicted point. The eigenvalue pair, where c is closest to 0, is most likely the one that belongs to the EP, and ideally there is a large gap between this pair and the next one. These two eigenvalues can be used as new additional training points.

Convergence



To end the iterative process, a convergence criterion is needed. Due to the degeneracy of the eigenvalues at the EP, the eigenvalue difference is a good parameter. This will never be exactly zero because of numerical errors, so a value $\varepsilon > 0$, depending on various factors such as the radius of the orbit, is chosen. For the 5×5 matrix model this process already converges after four diagonalizations (three training steps and one additional diagonalization to explore the energy plane). This method will be applied to excitons in external fields.

P. Egenlauf, "Application of machine learning to find exceptional points", Master thesis, in preparation (University of Stuttgart, 2023)

iRyd