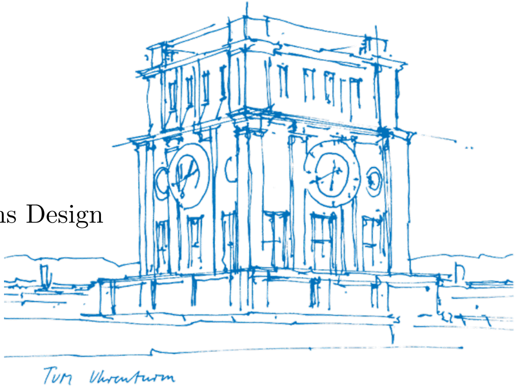
Quantum Capacity of a Deformed Bosonic Dephasing Channel

Shahram Dehdashti, Janis Nötzel, Peter van Loock

Emmy-Noether Group Theoretical Quantum Systems Design
Electrical Engineering Department
Technical University of Munich

TUM, Munich, 2nd Q.TOK workshop

May 07, 2023



Non-Linear Optics

Nonlinear optical refractive, in terms of a field-dependent index, is conveniently given by

$$n(\mathcal{E}) = n^2 + \tilde{n}_2|\mathcal{E}|^2 + \cdots$$

which $\mathcal{E}$ is a single monochromatic wave, and $\tilde{n}_2$ is proportion described by the third-order susceptibility $\tilde{\chi}^{(3)}$ [1]. The index of refraction on the intensity of the field is called the optical Kerr effect and the associated media with non-negligible values of $\tilde{n}_2$ are called Kerr media.

The quantum Hamiltonian which describes the first order of non-linearity is the starting point of our work. We consider instead of the Hamiltonian $\hat{H}_0 = \omega\hat{n}$ the Kerr Hamiltonian,

$$\hat{H}_{\omega,\lambda} = \omega\hat{n} + \Lambda\hat{n}^2/2,\ \Lambda \in \mathbb{R}$$

where Λ defines the power of the non-linearity[1].

[1] J. Garrison, R. Chiao, Quantum optics. OUP Oxford (2008); M. O. Scully, M. S. Zubairy, Quantum optics, American Association of Physics Teachers (1999).

Kerr medium and more

The Kerr Hamiltonian models many different quantum devices in different area. For example it can model

- Transmon Gate [J. J. Garcia-Ripoll, A. Ruiz-Chamorro, and E. Torrontegui, Physical Review Applied 14, 044035 (2020); M. Moein, A. Petrescu, and H. E. Tureci., Physical Review B 101.13, 134509 (2020).]
- Optical Kerr Cavities [N. Goldman, *et al.*, arXiv:2304.05865 (2023).]
- Kerr Cat Qubits [T. Aoki, *et al.* arXiv:2303.16622 (2023).]
- ...

and in fact any system basically is modeled by a Harmonic oscillator and includes the first order of non-linearity.

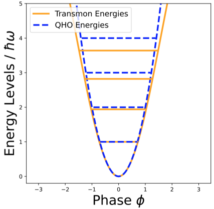


Figure: The spectrum of energy for anharmonic oscillator and harmonic oscillator.

In addition, mathematically gives the energy spectrum of

- a position-dependent quantum oscillator (the sing of the anharmonicity λ gives the constant rate of increase or decrease of the mass)[2].
- it models an oscillator confined in a finite or infinite well, respectively for negative and positive values of the anharmonicity[3].
- it describes a confined oscillator on a one-dimensional space with constant curvature, i.e., circle (negative) and hyperbolic (positive)[4].

[2] J. F. Carinena, *et al.*, J. Math. Phys. 52, 072104 (2011)
[3] M. B. Harouni, *et al.*, J. Phys. A, Math. Theor. 42, 045403 (2009); M. D. Darareh, *et al.*, Phys. Lett. B 374, 4099 (2010).
[4] J. F. Carinena, *et al.*, 45, 265303 (2012).

Motivation:

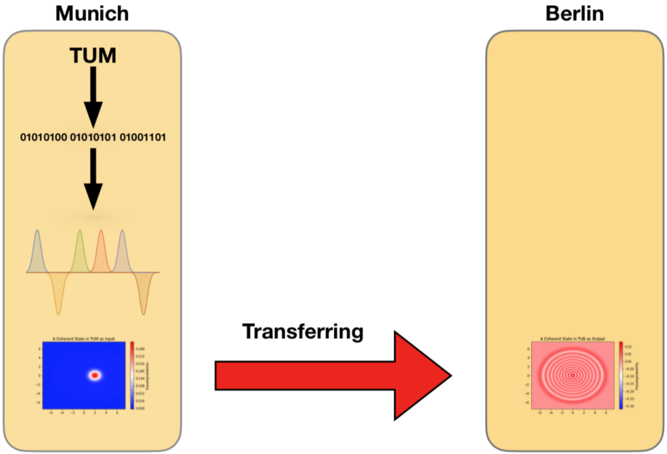
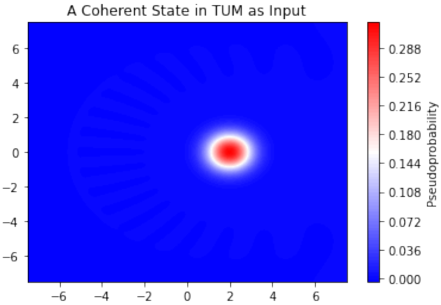
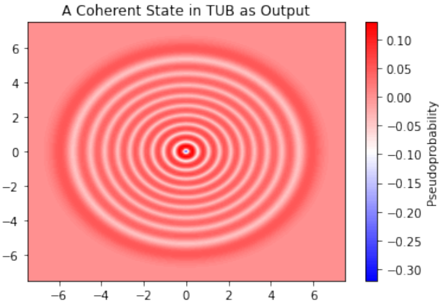


Figure: Transferring the word ”TUM” from Munich to Berlin, via a noisy channel.





The transition can be modeled by a bosonic dephasing channel:

$$\rho \mapsto \mathcal{N}_\gamma(\hat{\rho}) = \mathcal{N}_\gamma = \mathrm{Tr}_E\left[\hat{U}\left(\rho \otimes |0\rangle\langle 0|\right)\hat{U}^\dagger\right] = \sum_{n,m=0} e^{-\frac{1}{2}\gamma(m-n)^2}\rho_{m,n}\,|m\rangle\langle n|\,,$$

in which the unitary operator $\hat{U} = \exp\left[-i\sqrt{\gamma}\hat{a}^\dagger\hat{a}\left(\hat{b}+\hat{b}^\dagger\right)\right]$ defines the interaction between system and environment and is composed of annihilation and creation operators $\hat{a}$ and $\hat{a}^\dagger$ acting on the system Hilbert space and their corresponding counterparts $\hat{b}$ and $\hat{b}^\dagger$ acting on the Hilbert space of the environment, $\rho = \sum_{m,n}\rho_{m,n}\,|m\rangle\langle n| \in \mathcal{T}(\mathcal{H}_S)$ is an initial state of the system and $|0\rangle \in \mathcal{H}_E$ is a fixed initial state of the environment.

Additionally

- Decoherence by definition is a process in which a coherent superposition state is reduced to an incoherent probabilistic mixture of the states, i.e., $\sum_{n,m} c_n c_m^* |\psi_n\rangle\langle\psi_m| \rightarrow \sum_n |c_n|^2 |\psi_n\rangle\langle\psi_n|$.

  Preventing decoherence is one of the biggest challenges in the quantum domain.

  A simple model for decoherence we can consider a dynamical process by which the inputs are mapped to an output as follows,

$$\sum_{n,m} c_n c_m^* |n\rangle\langle m| \mapsto \sum_{n,m} e^{-\gamma(n-m)^2/2} c_n c_m^* |n\rangle\langle m| .$$

  in which $\gamma > 0$, so-called the decoherence parameter, is related to the strength of the decoherence.

Deformed Operators

- This question is the starting point of our work by considering the Kerr Hamiltonian,

$$\hat{H} = \Omega \hat{a}^\dagger \hat{a} + \frac{\lambda}{2} \hat{a}^{\dagger 2} \hat{a}^2,$$

where λ, the so-called anharmonicity, is related to the non-linear susceptibility of the Kerr medium and defines the power of the non-linearity. We define deformed annihilation and creation operators as

$$\hat{A}(\lambda, \omega) = \hat{a} f(\hat{n}) = f(\hat{n}+1)\hat{a}, \qquad \hat{A}^\dagger(\lambda, \omega) = f(\hat{n})\hat{a}^\dagger = a^\dagger f(\hat{n}+1),$$

in which $\omega := \Omega - \frac{\lambda}{2}$ and the deformation function is given by

$$f(\hat{n}) = \sqrt{1 + \frac{\lambda}{2\omega} \hat{n}}.$$

With this choice, we can rewrite the Kerr Hamiltonian as

$$\hat{H} = \omega \hat{A}^\dagger(\lambda, \omega) \hat{A}(\lambda, \omega).$$

Mathematical Properties

The sign of anharmonicity defines diffident mathematical structures:

$\lambda < 0$

- The algebra is generalization of $su(2)$-algebra.
- The Mandel Q-parameter of the associated coherent state is sub- Poissonian.
- Fubini–Study metric of of the associated coherent state is a sphere.

$\lambda = 0$

- The algebra is W-H-algebra.
- The Mandel Q-parameter of the associated is Poissonian.
- Fubini–Study metric of of the associated coherent state is a flat surface.

$\lambda > 0$

- The algebra is generalization of $su(1,1)$-algebra.
- The Mandel Q-parameter of the associated is super Poissonian.
- Fubini–Study metric of of the associated coherent state is a hyperbolic.

Deformed Creation and Annihilation Operators

Lemma

Let the operators $\hat{A} = \hat{a} f(\hat{n})$, $\hat{A}^\dagger = f(\hat{n})\hat{a}$, where $f(\hat{n}) = \sqrt{1 + \frac{\lambda}{2\omega}\hat{n}}$, then the commutation relations are given by

$$[\hat{A}, \hat{A}^\dagger] = 2\hat{K}_0, \qquad [\hat{K}_0, \hat{A}] = -\frac{\lambda}{2}\hat{A} \qquad [\hat{K}_0, \hat{A}^\dagger] = \frac{\lambda}{2}\hat{A}^\dagger.$$

Note that for $\lambda = 2$, $\lambda = 0$ and $\lambda = -2$ the above commutation relations is respectively identified by $su(1,1)$, W-H and $su(2)$ commutation relations.

Deformation Displacement Operator

Definition

For any $\beta \in \mathbb{C}$, we define a deformation displacement operator as

$$D(\beta, \lambda) = e^{\beta A^\dagger - \beta^* A}$$

Lemma

The Gaussian decomposition of the displacement operator $D(\beta, \lambda) = \exp\left[\beta A^\dagger - \beta^* A\right]$ *is given by*

$$D(\beta, \lambda) = e^{\beta A^\dagger - \beta^* A} = e^{\zeta A^\dagger} e^{\ln[\zeta_0] K_0} e^{-\zeta^* A}$$

where

$$\zeta = \frac{\beta}{|\beta|} \sqrt{\frac{2}{\lambda}} \tanh\left(\sqrt{\frac{\lambda}{2}} |\beta|\right), \qquad \zeta_0 = \cosh^{-4/\lambda}\left(\sqrt{\frac{\lambda}{2}} |\beta|\right).$$

By considering the negative values for λ*, by which* $\tanh(ix) = \tan(x)$ *and* $\cosh(ix) = \cos(x)$*, we can find the decomposition of the displacement operator for the negative values.*

Lemma

For $\lambda < 0$ the deformed displacement coherent state is obtained as

$$\left|\alpha;\lambda^-\right\rangle = \cos^{2\nu}\left(\sqrt{\frac{\lambda}{2}}|\mu|\right)\sum_{k=0}^{\lfloor 2\nu\rfloor} e^{-im\phi}\tan^k\left(\sqrt{\frac{\lambda}{2}}|\mu|\right)\sqrt{\frac{(2\nu)_k}{k!}}\,|k\rangle$$

where $\nu := |\lambda|^{-1} + 1/2$, $(x)_q = \Gamma(x+1)/\Gamma(x-q+1)$ is the falling factorial (alternatively, $(x)_q = \prod_{i=0}^{q}(x-i)$) and $\lfloor 2\nu\rfloor$ is the greatest integer less than or equal to 2ν

Lemma

For $\lambda > 0$, the deformed displacement coherent state is given by

$$\left|\alpha;\lambda^+\right\rangle \quad = \quad \cosh^{-2\nu}\left(\sqrt{\frac{\lambda}{2}}\mu\right)\sum_{k=0}^{\infty} e^{-im\phi}\tanh^k\left(\sqrt{\frac{\lambda}{2}}|\mu|\right)\times\sqrt{\frac{(2\nu)^k}{k!}}\,|k\rangle$$

where $(x)^q = \Gamma(x+q)/\Gamma(x)$ is the Pochhammer symbol (rising factorial).

Deformed Quantum Dephasing Channel

Similar to the definition of the dephasing channel, we can define a deformed quantum dephasing channel as follows:

Definition

Given $\lambda \in \mathbb{R}$ and $\gamma, \omega > 0$ the deformed quantum dephasing channel is defined as

$$\mathcal{N}_{\gamma}(\rho) := \mathrm{Tr}_E \left[U \left(\hat{\rho} \otimes \hat{\sigma} \right) U^{\dagger} \right]$$

in which the unitary operator $\hat{U}$ is

$$\hat{U} = \exp \left[-i \sqrt{\gamma} \omega \hat{A}^{\dagger} \hat{A} \left(\hat{B} + \hat{B}^{\dagger} \right) \right] \in \mathcal{L}(\mathcal{F} \otimes \mathcal{F})$$

where $\hat{A} = \hat{a} f(\hat{n})$ and $\hat{B} = \hat{b} f(\hat{n})$ are the deformed annihilation operators of the system and environment, respectively.

Theorem

Let $\rho = \sum_{m,n=0}^{\infty} \rho_{nm} |m\rangle\langle n| \in \mathcal{S}(\mathcal{F})$. *Then the deformed dephasing bosonic channel is given by*

$$\mathcal{N}_\gamma(\rho) = \sum_{n,m=0}^{\infty} \mathcal{K}^p_{n,m} \rho_{nm} |n\rangle \langle m|$$

By setting $\tau_n := \sqrt{\gamma \frac{\lambda}{2}}\, \omega\, n(1 + n\frac{\lambda}{2\omega})$*, and* $\nu := 1 + 2\omega/\lambda$*, we have*

$$\mathcal{K}^p_{m,n} = \langle -i\tau_n; \lambda | -i\tau_m; \lambda\rangle = \begin{cases} \cosh^{-2\nu}[\tau_n - \tau_m] & \lambda > 0 \\ e^{-\frac{\gamma}{2}(m-n)^2} & \lambda = 0 \\ \cos^{2\nu}[\tau_n - \tau_m] & \lambda < 0 \end{cases}$$

where $p = (\gamma, \lambda, \omega)$ *is the set of parameters defining the action of the channel.*

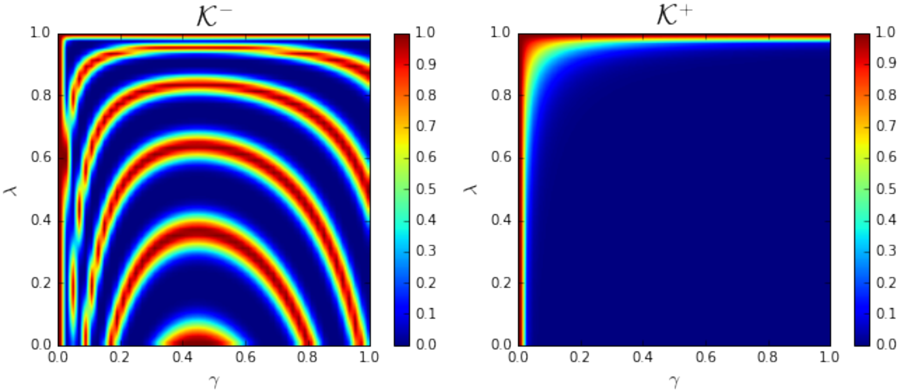


Figure: Density plot of functions $\mathcal{K}_{m,n}^{(\gamma,-\lambda,\omega)}$ and $\mathcal{K}_{m,n}^{(\gamma,\lambda,\omega)}$ as a function of parameters λ and γ for values $m, n = 1, 2$ and $\omega = 1$.

If $\lambda \geq 0$, the coefficient $\mathcal{K}^p$, approaches zero exponentially when $m \neq n$, when $\gamma \to \infty$, which means the off-diagonal elements of the density matrix map to zero in this channel, for $\gamma \gg 1$. In the case $\lambda < 0$, due to the periodic nature of the function $\mathcal{K}^p$, we are able to suppress the decoherence.

Definition

The quantum capacity of the bosonic dephasing channel is defined as

$$\mathcal{Q}(\mathcal{N}_\gamma) = \max_{\hat{\rho}} J(\hat{\rho}, \mathcal{N}_\gamma),$$

where

$$J(\hat{\rho}, \mathcal{N}_\gamma) = S(\mathcal{N}(\hat{\rho})) - S(\mathcal{N}^c(\hat{\rho}))$$

and $S(\hat{\rho}) = -\operatorname{Tr}[\hat{\rho}\log_2 \hat{\rho}]$ is the von Neumann entropy and the complementary channel $\mathcal{N}_\gamma^c$ is

$$\mathcal{N}^c(\hat{\rho}) = \operatorname{Tr}_S\left[U\left(\hat{\rho}\otimes\hat{\sigma}\right)U^\dagger\right].$$

It is possible to show that

Lemma

The input state achieving the maximum in (1) *is diagonal in the photon number basis* $\{|0\rangle, |1\rangle, \ldots\}$.

Theorem

For each dephasing channel in which the coefficient $\mathcal{K}_{n,m}$ can be written as a Toeplitz matrix, the capacity is given by[5]

$$Q(\mathcal{N}_\gamma) = \log_2(2\pi) - h(p)$$

in which the differential entropy $h(p)$ is defined:

$$h(p) = -\int d\phi p(\phi) \log_2 p(\phi).$$

Here, $p(\phi)$ is the Fourier kernel for which $\mathcal{K}_{n,m} = \int_{-\pi}^{\pi} d\phi p(\phi) e^{\mathbb{i}\phi(n-m)}$.

[5] L. Lami, M.M. Wilde, Exact solution for the quantum and private capacities of bosonic dephasing channels. Nat. Photon. (2023).

By considering $\lambda \ll 1$, we can write $\tau_n = \sqrt{\gamma}\omega n$. Hence, we are able to rewrite $\mathcal{K}_{n,m}$ as a symmetric Toeplitz matrix, i.e., matrices which the elements are only proportion with on the absolute of the difference $|n-m|$.

Lemma

For the negative case λ, the distribution $p(\phi)$ is given by

$$p(\phi) = 1 + 2\sum_{n=0}^{\lfloor 2\nu \rfloor} \cos^{2d}(\sqrt{\gamma}\omega n)\cos(\phi n)$$

where $\nu := 1 + 2\omega/\lambda$.

Lemma

For the positive λ, $p(\phi)$ can be written as

$$p(\phi) = \sum_{-\infty}^{\infty} e^{in\phi} \cosh^{-2\nu}(\sqrt{\gamma}\omega n) = 1 + 2\sum_{n=1}^{\infty} \cos(n\phi) \cosh^{-2\nu}(\sqrt{\gamma}\omega n)$$

By considering $\gamma \ll 1$, we can rewrite it as

$$p(\phi) = 1 + 2\sum_{n=1}^{\infty} \frac{\cos(\phi n)}{1 + 2\nu\gamma\omega n^2} = \frac{\pi}{\sigma} \frac{\cosh\left[\frac{\pi-\phi}{\sigma}\right]}{\sinh\left[\sigma^{-1}\right]}$$

where $\sigma = \sqrt{2\nu\omega\gamma}$.
For $\gamma \gg 1$, we have

$$p(\phi) = 1 + 2\sum_{n=1}^{\infty} \frac{\cos(\phi n)}{\omega^{\nu}\gamma^{\nu} n^{2\nu}} = 1 + (\omega\gamma)^{\nu} \frac{(2\pi)^{2\nu}}{(2\nu)!} B_{2\nu}\left(\frac{\phi}{2\pi}\right)$$

in which $B_{2\nu}(x)$ is the Bernoulli polynomials.

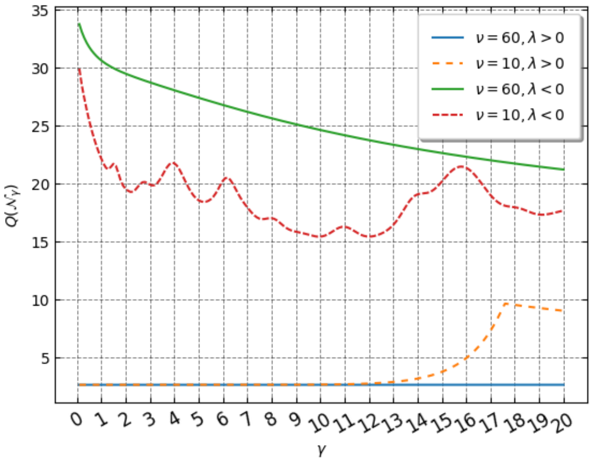


Figure: The quantum capacity $Q(\mathcal{N}_\gamma)$; the maximum of values resulting from the approximations in $\gamma \ll 0$ and $\gamma \gg 0$ versus dephasing parameter γ, for different values of λ and ν.

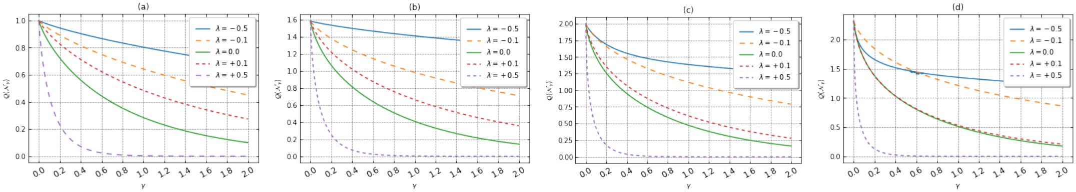


Figure: The optimal von Neumann entropy versus γ, for different values of λ, for $N = 1, 2, 3.$ and $N = 4$, where $N + 1$ is the dimension of Hilbert space,in plots (a), (b), (c) and (d), respectively.

For details please check the paper:

Dehdashti, Shahram, Janis Notzel, and Peter van Loock. ”Quantum capacity of a deformed bosonic dephasing channel.” arXiv:2211.09012 (2022).

Non-Linear Cat States

Definition

Cat qubits are encoded in a quantum harmonic oscillator within the two-dimensional subspace defined by two coherent states of opposite phase

$$|0\rangle_c = \frac{1}{\sqrt{2}}\left[|C, \lambda\rangle + \left|C^-, \lambda\right\rangle\right], \; |1\rangle_c = \frac{1}{\sqrt{2}}\left[\left|C^+, \lambda\right\rangle - \left|C^-, \lambda\right\rangle\right]$$

in which two-component deformed cat states are defined as the even and odd superposition of deformed coherent states with opposite displacement, i.e.,

$$\left|C^{\pm}, \lambda\right\rangle = \frac{1}{N_{\alpha}^{\pm}}\left(|\alpha; \lambda\rangle \pm |-\alpha; \lambda\rangle\right)$$

in which $N_{\alpha}^{\pm}$ is the normalized constant. Simple calculations respectively give the normalized constant with negative and positive values of the anharmonicity as

$$N_{\alpha}^{\pm} = \sqrt{2\left[1 \pm \cosh^{-2\nu}\left(\sqrt{2|\lambda|}|\alpha|\right)\right]}, \; N_{\alpha}^{\pm} = \sqrt{2\left[1 \pm \cos^{2\nu}\left(\sqrt{2|\lambda|}|\alpha|\right)\right]}$$

Lindblad Master Equation

The master equation the system can be modeled by (Gorini–Kossakowski–Sudarshan–) Lindblad equation

$$\partial_t \hat{\rho} = \mathcal{L}\hat{\rho}$$

where the Liouvillian superoperator is defined by

$$\mathcal{L}\hat{\rho} = \kappa_l \mathcal{D}[\hat{A}]\hat{\rho} + \kappa_d \mathcal{D}[\hat{A}^\dagger \hat{A}]\hat{\rho}$$

where

$$\mathcal{D}[\hat{C}]\hat{\rho} = \hat{C}^\dagger \hat{\rho}\hat{C} - \frac{1}{2}\{\hat{C}^\dagger \hat{C}, \hat{\rho}\}$$

for an operator C. The dissipator superoperators $\mathcal{D}[\hat{A}]$ and $\mathcal{D}[\hat{A}^\dagger \hat{A}]$ describe particle loss and dephasing processes occurring with rates κ_l and κ_d, respectively.

Deformed Loss-dephasing Channel

The noise channel gives the evolution of an initial state $\hat{\rho}$ for the interval time τ,

$$\mathcal{N}_{\kappa_l,\kappa_d}(\hat{\rho}(0)) = e^{\mathcal{L}\tau}\hat{\rho}(0) = \sum_{j}^{\infty} \hat{K}_j \hat{\rho}(0) \hat{K}_j^{\dagger}$$

where $\hat{K}_j$s are Kraus operators. To leading order in $\kappa_{l,d}\tau \ll 1$, the channel can be written as

$$\mathcal{N}_{\kappa_l,\kappa_d}(\hat{\rho}(0)) \approx \sum_{j=0}^{2} \hat{K}_j \hat{\rho}(0) \hat{K}_j^{\dagger}$$

where

$$\begin{aligned}
K_0 &= \mathbb{I} - \frac{\kappa_l \tau}{2} \hat{A}^{\dagger}\hat{A} - \frac{\kappa_d \tau}{2} \left(\hat{A}^{\dagger}\hat{A}\right)^2 \\
K_1 &= \sqrt{\kappa_d \tau}\,\hat{A}^{\dagger}\hat{A} \\
K_2 &= \sqrt{\kappa_d \tau}\,\hat{A}
\end{aligned}$$

K-L Condition

We define the following set of the errors, E^m:

$$\{E^m\} = \{\mathbb{I}, \hat{A}, \hat{A}^\dagger \hat{A}, (\hat{A}^\dagger \hat{A})^2\},$$

The Knill–Laflamme (KL) conditions, characterizes the errors:

$$0 \leq \frac{|\langle\psi_i| \hat{E}_m^\dagger \hat{E}_{m'} |\psi_j\rangle|}{\sqrt{\langle\psi_i| \hat{E}_m^\dagger \hat{E}_{m'} \hat{E}_m^\dagger \hat{E}_{m'} |\psi_j\rangle}} \leq 1$$

where zero corresponds to perfect correction of errors if $i \neq j$, and 1 corresponds to a maximum violation of KL conditions.

Wigner Function

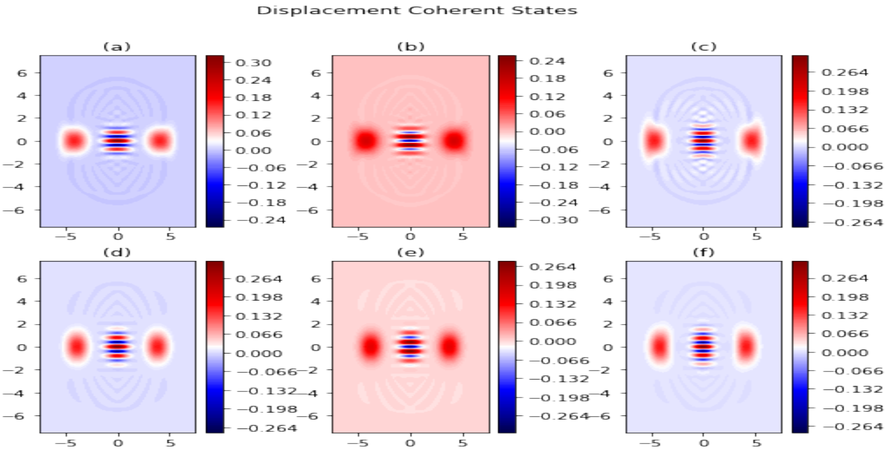


Figure: Wigner quasi-probability distribution function $W(X, Y)$ of the codewords the cat states, particle loss $\hat{A}\left|C^{+}\right\rangle$ and dephasing $\hat{A}^{\dagger}\hat{A}\left|C^{+}\right\rangle$, are depicted for the state $\left|\mathcal{C}^{+},\lambda\right\rangle$, for $\lambda = 0.1$ in the left plots (a)-(c), respectively. Plots (d)-(f) are illustrated the same with $\lambda = -0.1$.

Cat states suppress the dephasing error, as seen from the KL conditions. In particular, for large λ the KL conditions for the set of errors $\{\mathbb{I}, \hat{A}^\dagger \hat{A}\}$, we have

$$\left\langle \mathrm{C}^{\pm}, \lambda \right| \hat{A}^\dagger \hat{A} \left| \mathrm{C}^{\mp}, \lambda \right\rangle = 0$$

and

$$\langle \mathrm{C}^{\pm}, \lambda | \hat{A}^\dagger \hat{A} | \mathrm{C}^{\pm}, \lambda \rangle = \tfrac{\lambda}{2} \left[\tfrac{\mathrm{N}_\alpha^{\mp}}{\mathrm{N}_\alpha^{\pm}} \right]^2 \mu^2(\lambda)$$

where $\mu(\lambda)$ is defined as

$$\mu = \begin{cases} \tanh \left[\sqrt{\frac{\lambda}{2}} |\alpha| \right] & \lambda > 0 \\ \tan \left[\sqrt{\frac{|\lambda|}{2}} |\alpha| \right] & \lambda < 0 \end{cases}$$

For the positive values of anharmonisity, μ approaches $\lambda/2$, when $\lambda \ll 1$. However, in the case of negative one, when $\sqrt{|\lambda|}|\alpha| = (2n+1)\pi/\sqrt{8}$, with $n \in \mathbb{N}$, we have $\langle \mathrm{C}^{\pm}, \lambda | \hat{A}^\dagger \hat{A} | \mathrm{C}^{\pm}, \lambda \rangle = \lambda/\sqrt{8}$.

Recovery Channel

- We consider the channel fidelity approach in which we measure and recover the quantum state periodically at time intervals τ.
- We define a set of recovery maps $\mathcal{R}$, i.e.,

$$\hat{\rho}(t+\tau) = \mathcal{R}(e^{\mathcal{L}\tau}\hat{\rho}(t)) = \mathcal{R} \circ \mathcal{N}_{\kappa_l,\kappa_d}(\hat{\rho}(t))$$

- The whole dynamical process can be written by the Kraus operators:

$$\mathcal{E}(\hat{\rho}(\tau)) = \sum_{j=0}^{\infty} \hat{S}_j \hat{\rho}(0) \hat{S}_j^{\dagger}$$

where

$$\hat{S}_j = \sum_{l,r=0}^{\infty} R_l K_r$$

where K_r are the Kraus operators associated with the noise channel $\mathcal{N}_{\kappa_l,\kappa_d}$ and R_l are the Kraus operators associated with the recovery $\mathcal{R}$.

The average channel fidelity, defined as

$$\mathcal{F}_{avg} = \frac{1}{d^2} \sum_{j=0}^{\infty} | \operatorname{Tr}\left(\hat{S}_j\right)|^2 = \frac{1}{d^2} \sum_{l,k=0}^{\infty} | \operatorname{Tr}\left(\hat{R}_l \hat{K}_r\right)|^2$$

quantifies the average performance of a certain recovery operation $\mathcal{R}$.

We should find the best recovery operation, i.e. the $\mathcal{R}$ which maximizes the average channel fidelity.

To optimize the recovery operation $\mathcal{R}$, we can express the operators $\hat{R}_l$ in terms of a set of basis operators $\hat{R}_i = \sum x_{il} \hat{B}_l$, in which the set of basis operators $\{\hat{B}_i\}$ is orthogonal, i.e., $\operatorname{Tr}\left[\hat{B}_i \hat{B}_j\right] = \delta_{ij}$.

We define the non-orthogonal subspaces spanned by the states $\left|\psi_i^{\pm}\right\rangle$ as

$$\left|\psi_i^{\pm}\right\rangle = \hat{E}_m \left|\mathcal{C}_f^{\pm}\right\rangle, \quad \hat{E}_m = \{\mathbb{I}, \hat{A}, \hat{A}^{\dagger}\hat{A}, (\hat{A}^{\dagger}\hat{A})^2\}$$

The recovery operations, $\hat{B} = P_m^{(n)}$, $m, n = 0, 1, 2, 3$, are defined as the following:

$$\begin{aligned}
P_m^{(0)} &= \left|\mathcal{C}_f^{+}\right\rangle \left\langle\psi_m^{+}\right| + \left|\mathcal{C}_f^{-}\right\rangle \left\langle\psi_m^{+}\right| \\
P_m^{(1)} &= \left|\mathcal{C}_f^{+}\right\rangle \left\langle\psi_m^{-}\right| + \left|\mathcal{C}_f^{-}\right\rangle \left\langle\psi_m^{+}\right| \\
P_m^{(2)} &= \left|\mathcal{C}_f^{+}\right\rangle \left\langle\psi_m^{-}\right| - i \left|\mathcal{C}_f^{-}\right\rangle \left\langle\psi_m^{+}\right| \\
P_m^{(3)} &= \left|\mathcal{C}_f^{+}\right\rangle \left\langle\psi_m^{+}\right| + \left|\mathcal{C}_f^{-}\right\rangle \left\langle\psi_m^{-}\right|
\end{aligned}$$

They are mutually orthogonal and describe an arbitrary action of the recovery operation projecting $\left|\psi_m^{\pm}\right\rangle$ back onto the code space.

Optimization

The optimization procedure aims to find the coefficients x_{il} which maximize the channel fidelity $\mathcal{F}_{avg}$, for a given set $\{\hat{B}_i\}$. In the other word, we should maximize

$$\frac{1}{d^2}\sum_{ij} X_{ij} W_{ij} = \frac{1}{d^2}\,\mathrm{Tr}\,[XW]$$

where

$$W_{ij} = \sum_{l} \mathrm{Tr}\{B_i K_l\}\,\mathrm{Tr}\{B_j K_l\}^*$$

subject to

$$\mathbf{X} \succcurlyeq 0$$

$$\sum_{ij} X_{ij} B_i^\dagger B_j = \mathbb{I}$$

which assert positive semidefinite of the recovery matrix and trace preservation, respectively.

Optimization

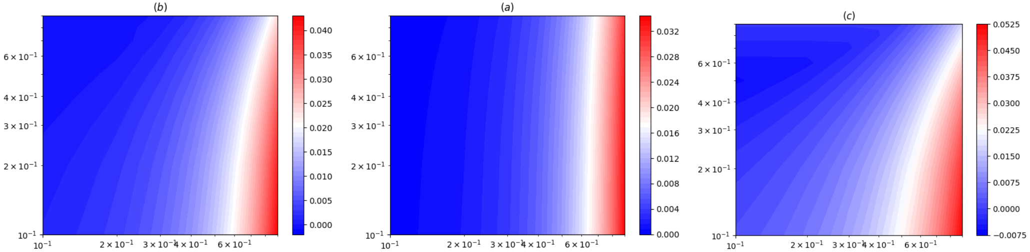


Figure: Color-map plot of the channel infidelity $1 - \mathcal{F}$ computed for the optimal SC code as a function of the particle loss parameter $\kappa_l \tau$ and dephasing parameter $\kappa_d \tau$ for various values of anharmonisity $\lambda = 0.1, 0$ and -0.1 in plots (a), (b) and (c), respectively.

Thanks

Looking forward to a productive interaction!

Quantum Support Vector Machines

- Better Decision Boundaries
- Descriptive feature space using the Kernel trick
- Potential Quantum advantage
- Promising Results

Better Decision Boundaries

Logistic Regression (LR)

The LR model is given by

$$f_w(x) = w^T x,$$
$$h(x) = \frac{1}{1 + e^{-f_w(x)}},$$
$$\mathcal{L}(w) = \frac{1}{m}\sum_{i=1}^{m}(-y_i \log(h(x_i)) + (1 - y_i)\log(1 - h(x_i)).$$

where $x_1, \ldots, x_m$ are the training samples, w are the weights of the model, and $y_1, \ldots, y_m \in \{0, 1\}$ indicate which class a data point belongs to. The goal of the training phase is to minimize $\mathcal{L}(w)$ given training data $(x_1, y_1), \ldots, (x_m, y_m)$.

A choice of weights w for which $\mathcal{L}(w) = \min_w \mathcal{L}(w)$ is kept as a result of the training phase.

Later on, a new data point x' is classified as $y' = 0$ if $h(x') < 1/2$ and as $y' = 1$, else.

LR Decision Boundary

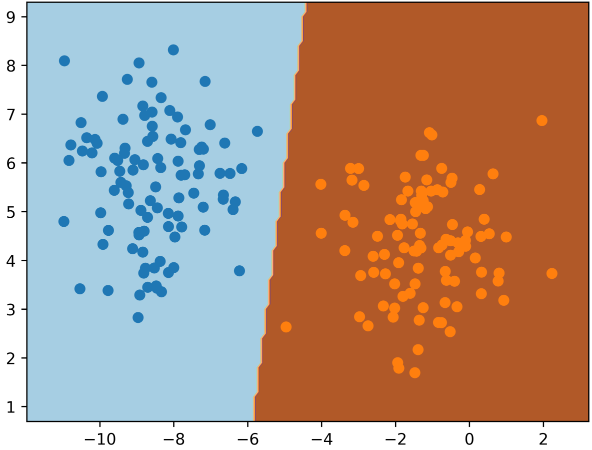


Figure: Decision boundary of the Logistic Regression model

Support Vector Machines

The Support Vector Machines (SVM) model in its dual form is given by

$$\max_{\lambda_1,\ldots,\lambda_m} \left[\sum_{i=1}^{m} \lambda_i - \frac{1}{2} \sum_{i,j=1}^{m} \lambda_i \lambda_j y_i y_j \, \mathrm{K}(x_i, x_j) \right],$$

$$\text{subject to } 0 \leqslant \lambda_i \leqslant C,$$

$$\sum_{i=1}^{m} \lambda_i y_i = 0.$$

where λ is the Lagrange multiplier, $(\mathrm{K}(x_i, x_j))_{i,j=1}^{m}$ is the Gram matrix, C is a penalty parameter, and $y_1, \ldots, y_m$ are the labels.

After a solution has been obtained, a new data point x' is classified as "0" if $\sum_{i=1}^{m} \lambda_i y_i K(x_i, x') < 0$ and as "1", else.

Support Vector Machines

Finally, $(\mathrm{K}(x_i, x_j))_{i,j=1}^{m}$ is a matrix of the shape $m \times m$ where m is the number of samples. This matrix is positive semi-definite since its eigenvalues are non-negative. The elements of this matrix correspond to the inner product between two feature vectors, or simply, it measures the similarity between them. It's also known as the kernel matrix. In short, the above equation indicates that the SVM is a convex optimization problem that can be used to deal with both linear separable and non-linear separable datasets.

SVM Decision Boundary

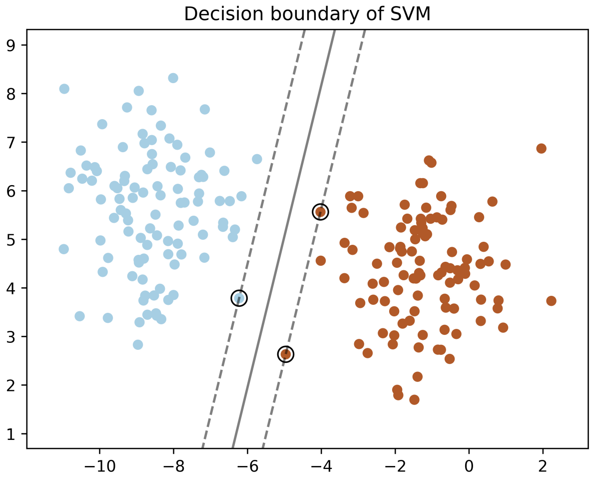


Figure: Decision boundary of the SVM model

Descriptive feature space using the Kernel trick

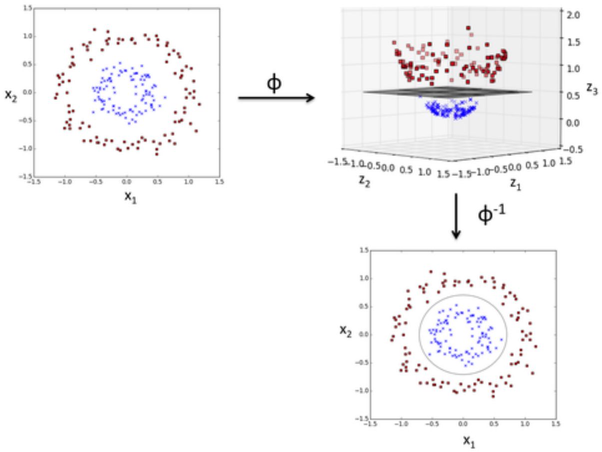


Figure: The effect of the kernel trick

Types of Classical Kernels

type of SVM	$\mathrm{K}(\mathrm{x}, \mathrm{y})$	Comments
Polynomial learning machine	$\left(\mathbf{x}^{\top}\mathbf{y}+1\right)^{\mathrm{p}}$	p: selected a priori
Radial basis function	$\exp\left(-\frac{1}{2\sigma^2}\|\mathbf{x}-\mathbf{y}\|^2\right)$	σ^2 : selected a priori
Two-layer perceptron	$\tanh\left(\beta_0\mathbf{x}^{\top}\mathbf{y}+\beta_1\right)$	only some $\beta_\circ$ and β_1 values are feasible.

Importance of SVM in Communications

- SVM can be used in Network Intrusion Detection systems.
- SVM can be used in Antenna allocation scenarios.
- SVM can be used in PSK/QAM constellation classification problems.
- SVM can be used in localization scenarios based on wireless sensor networks.

Quantum Advantage

Definition (Quantum kernel)

Let ϕ be a feature map over the input space of the dataset. A quantum kernel can be defined in terms of the inner product between two feature vectors x_i, x_j each represented by a quantum state operator $\rho(x_i), \rho(x_j)$ with $x_i, x_j \in \mathcal{X}$,

$$K(x_i, x_j) = \mathrm{tr}\{\rho(x_j)\rho(x_i)\} = |\langle \phi(x_j), \phi(x_i) \rangle|^2.$$

For generic quantum states σ and ρ, one can use instead of the scalar product the SWAP test measurement to calculate

$$(tr(\sqrt{\rho\sigma}))^2.$$

- Why do we need a quantum paradigm for such optimization problems?

Quantum Advantage

- Encoding classical data into the complex quantum Hilbert space gives us a richer feature space that makes the classification problem easier.
- Calculating the inner products between two quantum systems can be performed as shown below

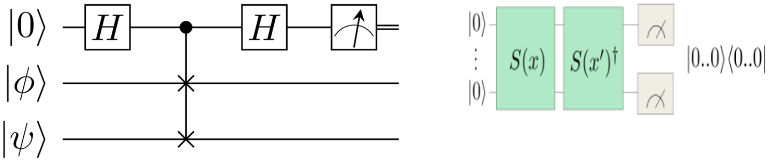


Figure: Two quantum protocols to measure the inner product between two samples. The left quantum circuit is the typical swap test between ρ and ϕ. The right circuit represents the Compute-Uncompute protocol between two samples x and x'.

Quantum Advantage

- From the previous figure, one can construct the Gram matrix using a quantum computer and then feed it into a classical SVM.
- It is also possible to perform the classification task, fully, in a quantum setting instead of the hybrid approach.

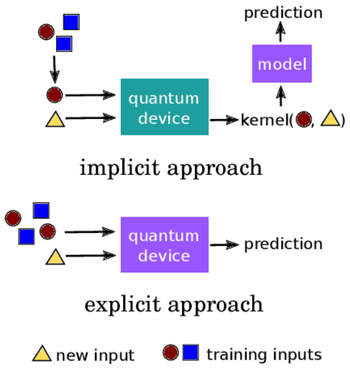


Figure: Illustration of the two approaches to using quantum feature maps for supervised learning.

Quantum Advantage

- There are two new proposed quantum kernels. These two kernels arise from the aforementioned properties of the nonlinear coherent states.
  - Negative Kernel

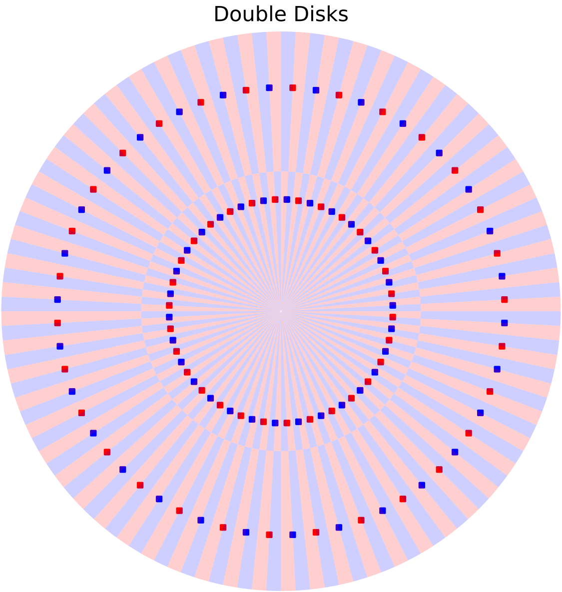
$$\langle \beta', \alpha | \beta, \alpha \rangle = \frac{\left[1 + e^{i(\phi-\phi')} \tan\left(\sqrt{\frac{\alpha}{2}}\beta\right) \tan\left(\sqrt{\frac{\alpha}{2}}\beta'\right)\right]^{2k}}{\left[1 + \tan^2\left(\sqrt{\frac{\alpha}{2}}\beta\right)\right]^{k} \left[1 + \tan^2\left(\sqrt{\frac{\alpha}{2}}\beta'\right)\right]^{k}}$$

  - Positive Kernel

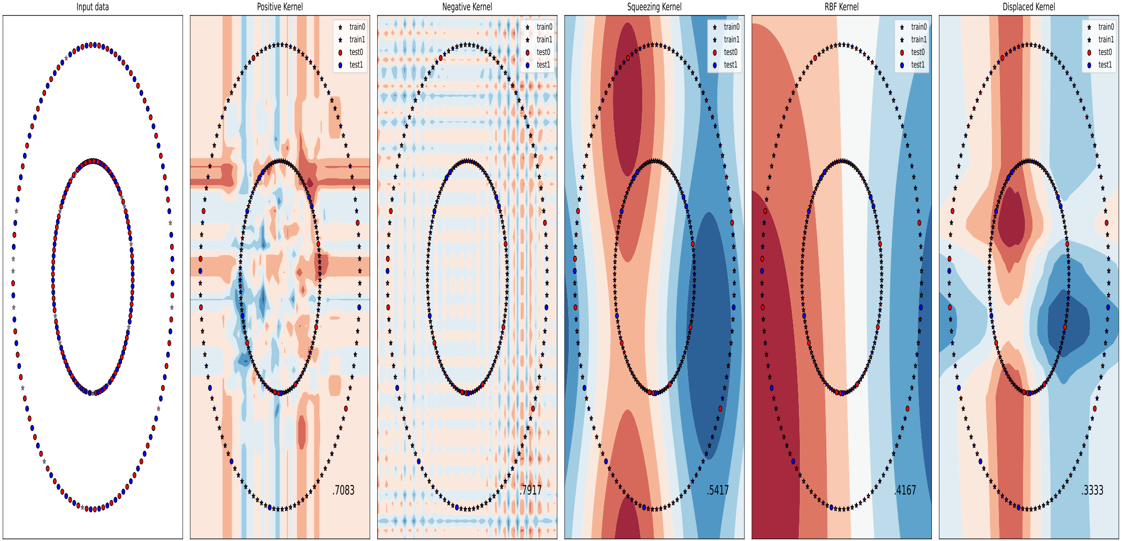
$$\langle \beta', \alpha | \beta, \alpha \rangle = \frac{\left[1 - \tanh^2\left(\sqrt{\frac{\alpha}{2}}\beta\right)\right]^{k} \left[1 - \tanh^2\left(\sqrt{\frac{\alpha}{2}}\beta'\right)\right]^{k}}{\left[1 - e^{i(\phi-\phi')} \tanh\left(\sqrt{\frac{\alpha}{2}}\beta\right) \tanh\left(\sqrt{\frac{\alpha}{2}}\beta'\right)\right]^{2k}}$$

  where $k, \alpha > 0$ and ϕ and ϕ' represent the data samples.
- The output from the two kernels is bounded $\in [0, 1]$.
- The data encoding scheme here is based on *Angle Encoding* since we are using the phase of each nonlinear coherent state to encode the data.
- This means that the quantum circuit scales linearly with the number of features per sample when the Compute-Uncompute approach is used.

Promising Results



Promising Results



Promising Results

- The previous figure shows the outstanding ability of the negative kernel to fit the data correctly despite its difficulty.
- Over multiple simulations the Displaced (Linear Coherent State) Kernel and the classical RBF had similar performance.
- The Squeezing kernel that uses a Squeezed state has a slightly better performance than the Coherent state kernel.
- The positive kernel and the negative kernel are better than all other kernels in terms of training and testing accuracy.
- The high performance is due to the fact that the k parameter can finetune the bandwidth of the Kernel and the β parameter can determine the shape of the decision surface as shown in the following figures.

Promising Results

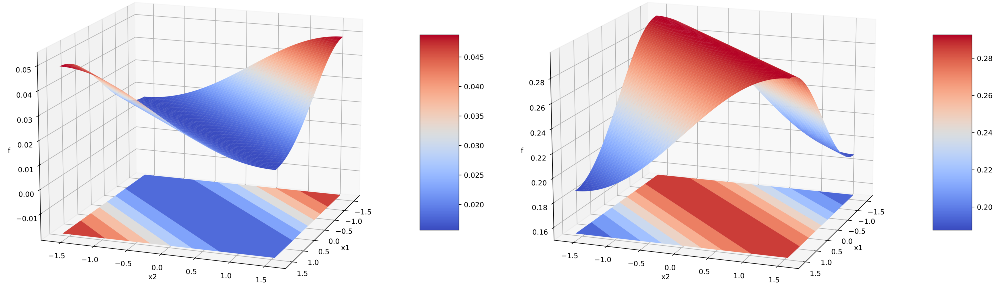
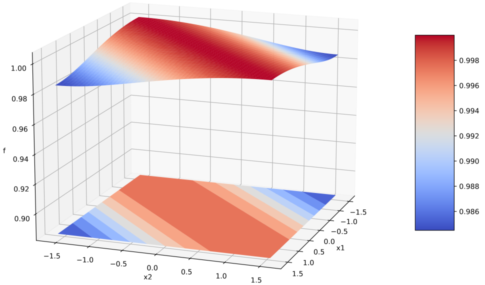


Figure: The decision surface of the negative kernel. The left figure is generated when $\beta = 0.1$ and $\beta' = \pi$. The right figure is generated when $\beta' = \pi/2$

Promising Results



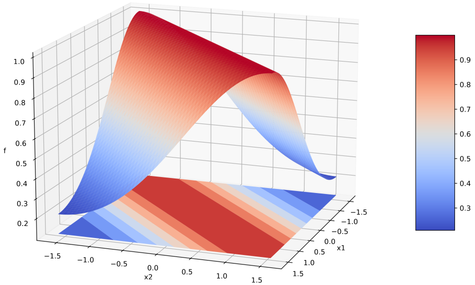


Figure: The decision surface of the negative kernel. The left figure is generated when $k = 0.01$. The right figure is generated when $k = 1$

Thanks

Looking forward to a productive interaction!

Addition Materials:

Lemma

Let the operators $\hat{A}$, $\hat{A}^\dagger$ and $\hat{K}_0$ satisfy the commutation relations

$$[\hat{A},\hat{A}^\dagger] = 2\hat{K}_0, \qquad [\hat{K}_0,\hat{A}] = -\frac{\lambda}{2}\hat{A} \qquad [\hat{K}_0,\hat{A}^\dagger] = \frac{\lambda}{2}\hat{A}^\dagger.$$

Then the Gaussian decomposition of the displacement operator $D(\beta,\lambda) = \exp\left[\beta A^\dagger - \beta^* A\right]$ *is given by*

$$D(\beta,\lambda) = e^{\beta A^\dagger - \beta^* A} = e^{\zeta A^\dagger} e^{\ln[\zeta_0] K_0} e^{-\zeta^* A}$$

where

$$\zeta = \tfrac{\beta}{|\beta|}\sqrt{\frac{2}{\lambda}}\tanh\left(\sqrt{\frac{\lambda}{2}}|\beta|\right), \qquad \zeta_0 = \cosh^{-4/\lambda}\left(\sqrt{\frac{\lambda}{2}}|\beta|\right).$$

For $\lambda < 0$ the deformed displacement coherent state is obtained as

$$\left|\alpha;\lambda^-\right\rangle = \cos^{2\nu}\left(\sqrt{\frac{\lambda}{2}}|\mu|\right)\sum_{k=0}^{\lfloor 2\nu\rfloor} e^{-im\phi}\tan^k\left(\sqrt{\frac{\lambda}{2}}|\mu|\right)\sqrt{\frac{(2\nu)_k}{k!}}\,|k\rangle$$

where $\nu := |\lambda|^{-1} + 1/2$, $(x)_q = \Gamma(x+1)/\Gamma(x-q+1)$ is the falling factorial (alternatively, $(x)_q = \prod_{i=0}^{q}(x-i)$) and $\lfloor 2\nu\rfloor$ is the greatest integer less than or equal to 2ν

For $\lambda > 0$, the deformed displacement coherent state is given by

$$\left|\alpha;\lambda^+\right\rangle \quad = \quad \cosh^{-2\nu}\left(\sqrt{\frac{\lambda}{2}}\mu\right)\sum_{k=0}^{\infty} e^{-im\phi}\tanh^k\left(\sqrt{\frac{\lambda}{2}}|\mu|\right)\times\sqrt{\frac{(2\nu)^k}{k!}}\,|k\rangle$$

where $(x)^q = \Gamma(x+q)/\Gamma(x)$ is the Pochhammer symbol.

We calculate the partial trace over the environment as

$$\mathcal{N}_\gamma(\rho) = \mathrm{Tr}_E\left[\hat{U}\rho \otimes |0\rangle\langle 0|\,\hat{U}^\dagger\right] = \sum_{m,n} \rho_{m,n}\,\mathrm{Tr}_E\left[U\,|m\rangle\langle n| \otimes |0\rangle\langle 0|\,U^\dagger\right],$$

where the sums can run from 0 to ∞ in case of $\lambda \geq 0$, or to a finite number d in case of $\lambda < 0$, i.e., a finite dimension d. The calculation of the partial trace in the above relation relies on the following:

$$\hat{U}\,|m\rangle\,|0\rangle = \sum_{k=0}^{\infty}(-i\sqrt{\gamma}A^\dagger A)^k \frac{(B+B^\dagger)^k}{k!}\,|m\rangle\,|0\rangle = |m\rangle \sum_{k=0}^{\infty} \frac{(-i\tau_m(B+B^\dagger))^k}{k!}\,|0\rangle$$

where $\tau_m = \sqrt{\gamma}\cdot\langle m|\,\hat{A}^\dagger\hat{A}\,|m\rangle$.