# Exploring the Cost Landscape of Variational Quantum Algorithms

IAAS Research

University of Stuttgart

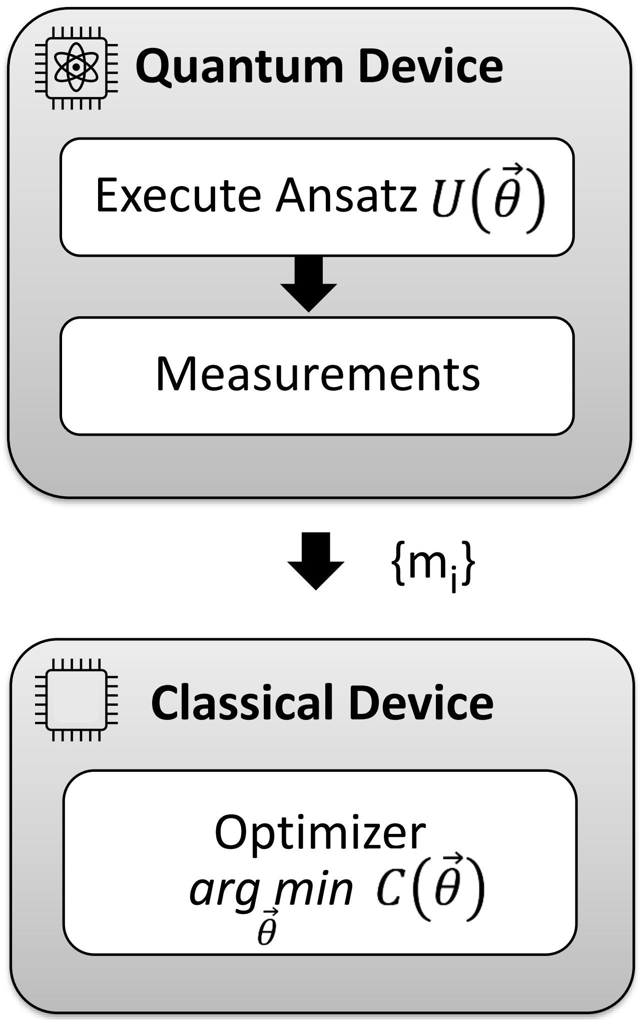
<u>**Lavinia Stiliadou**</u>**, Johanna Barzen, Frank Leymann,**

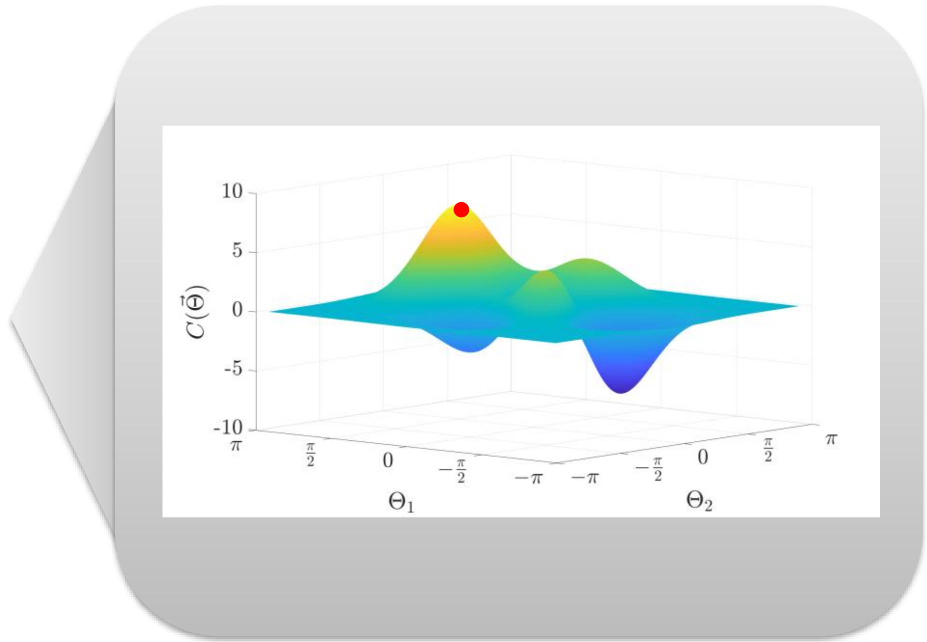
**Alexander Mandl, Benjamin Weder**

*lastname@iaas.uni-stuttgart.de*

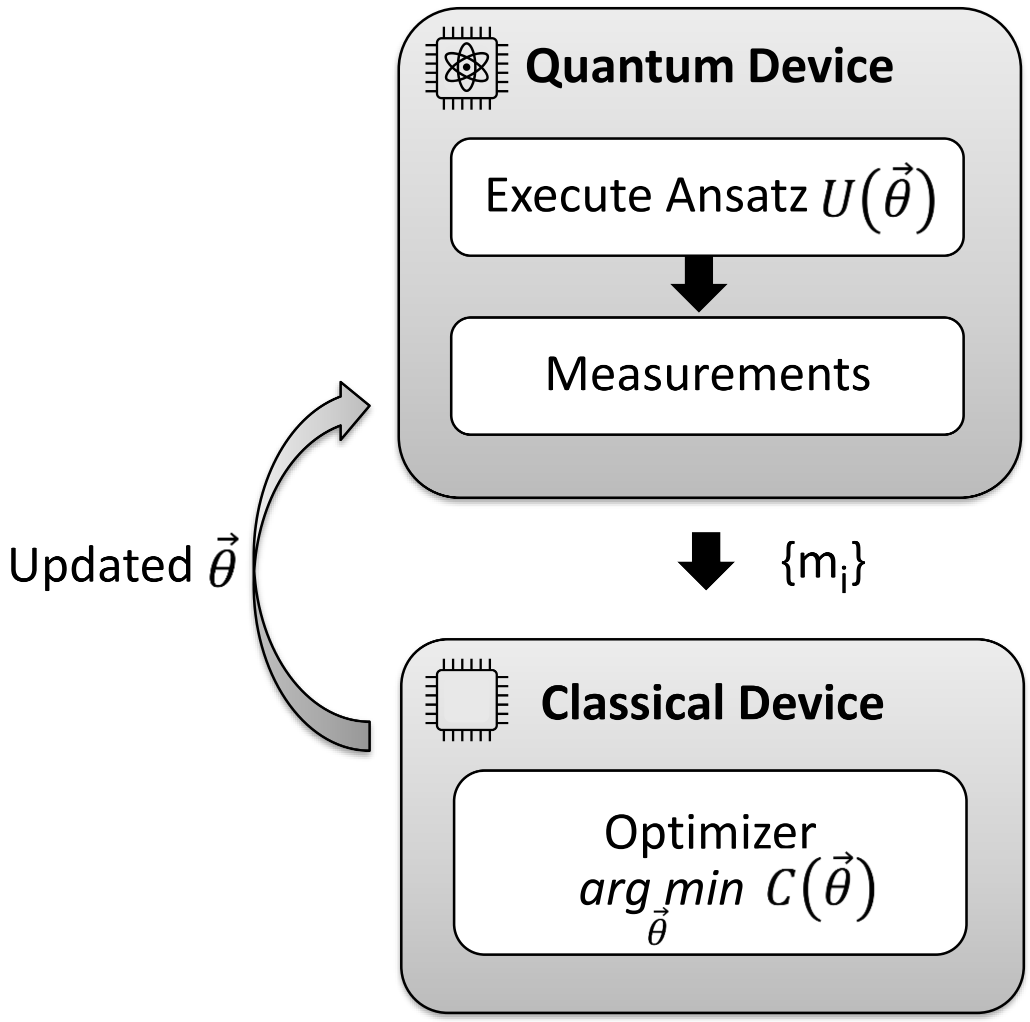
Institute of Architecture of Application Systems

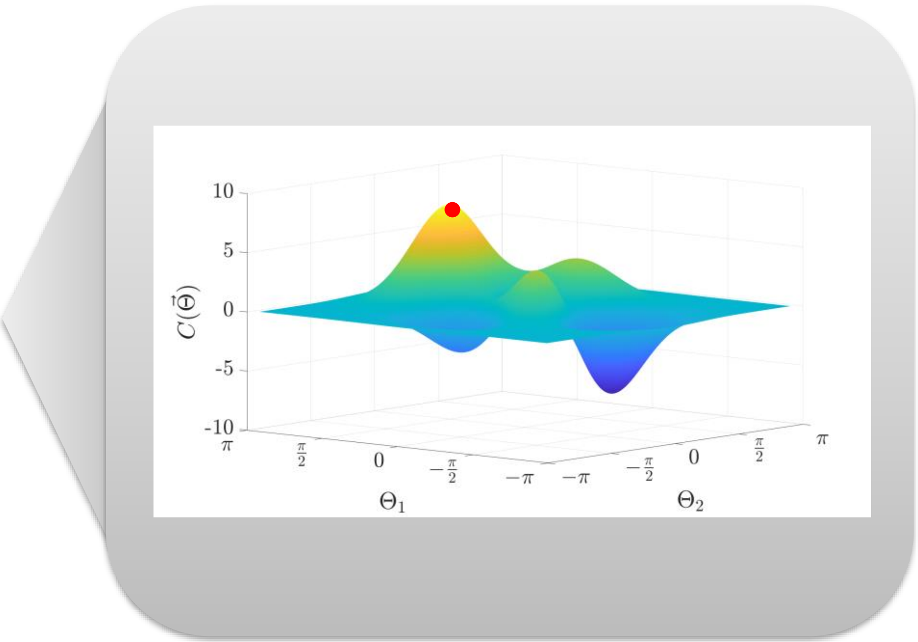
# Motivation


Quantum Device
Execute Ansatz $U(\vec{\theta})$
Measurements
$\{m_i\}$
Classical Device
Optimizer
$\arg\min_{\vec{\theta}} C(\vec{\theta})$


$C(\vec{\Theta})$
10
5
0
-5
-10
$\Theta_1$
$\Theta_2$
$\pi$
$\frac{\pi}{2}$
0
$-\frac{\pi}{2}$
$-\pi$

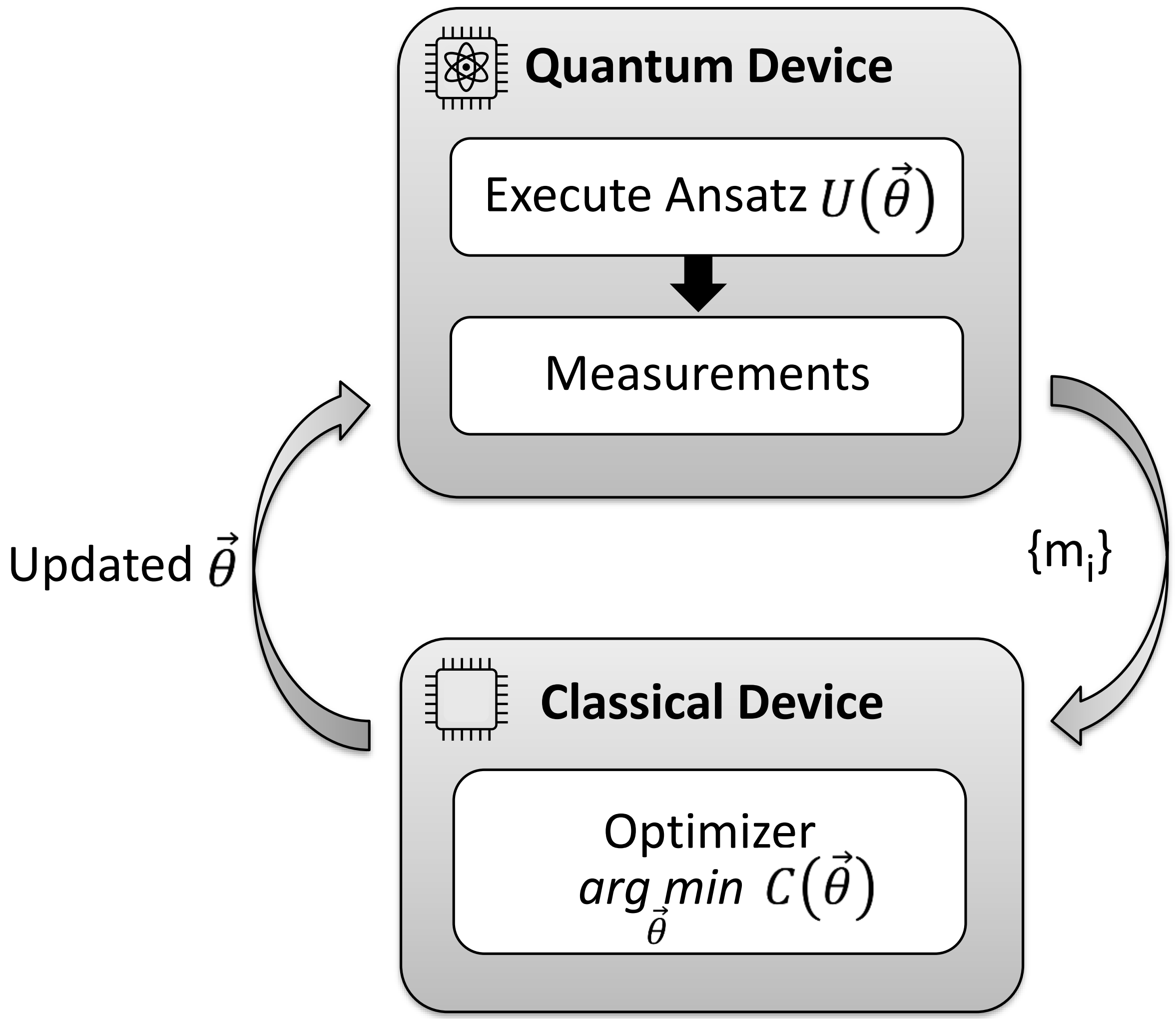
# Motivation



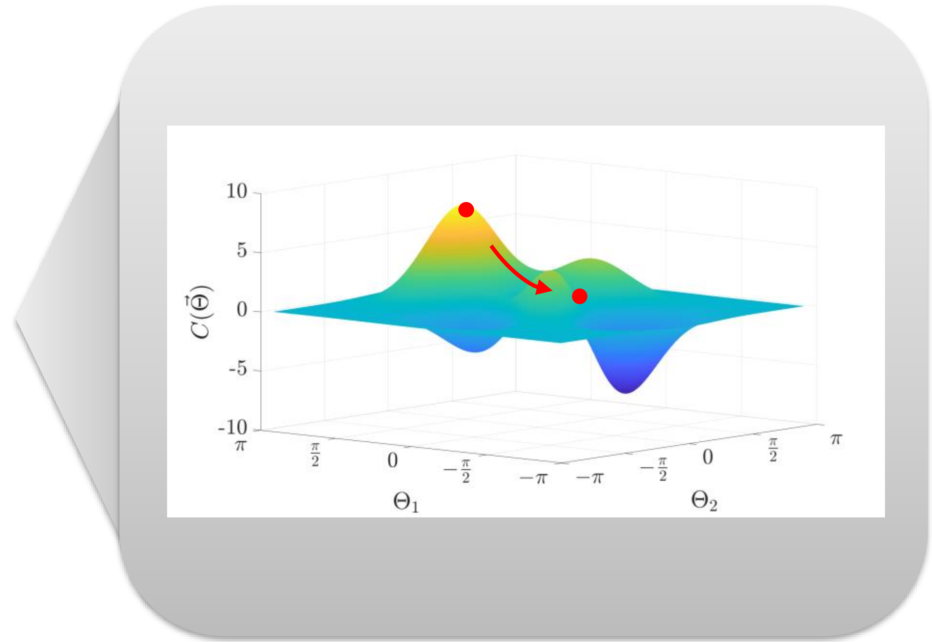




Research

# Motivation

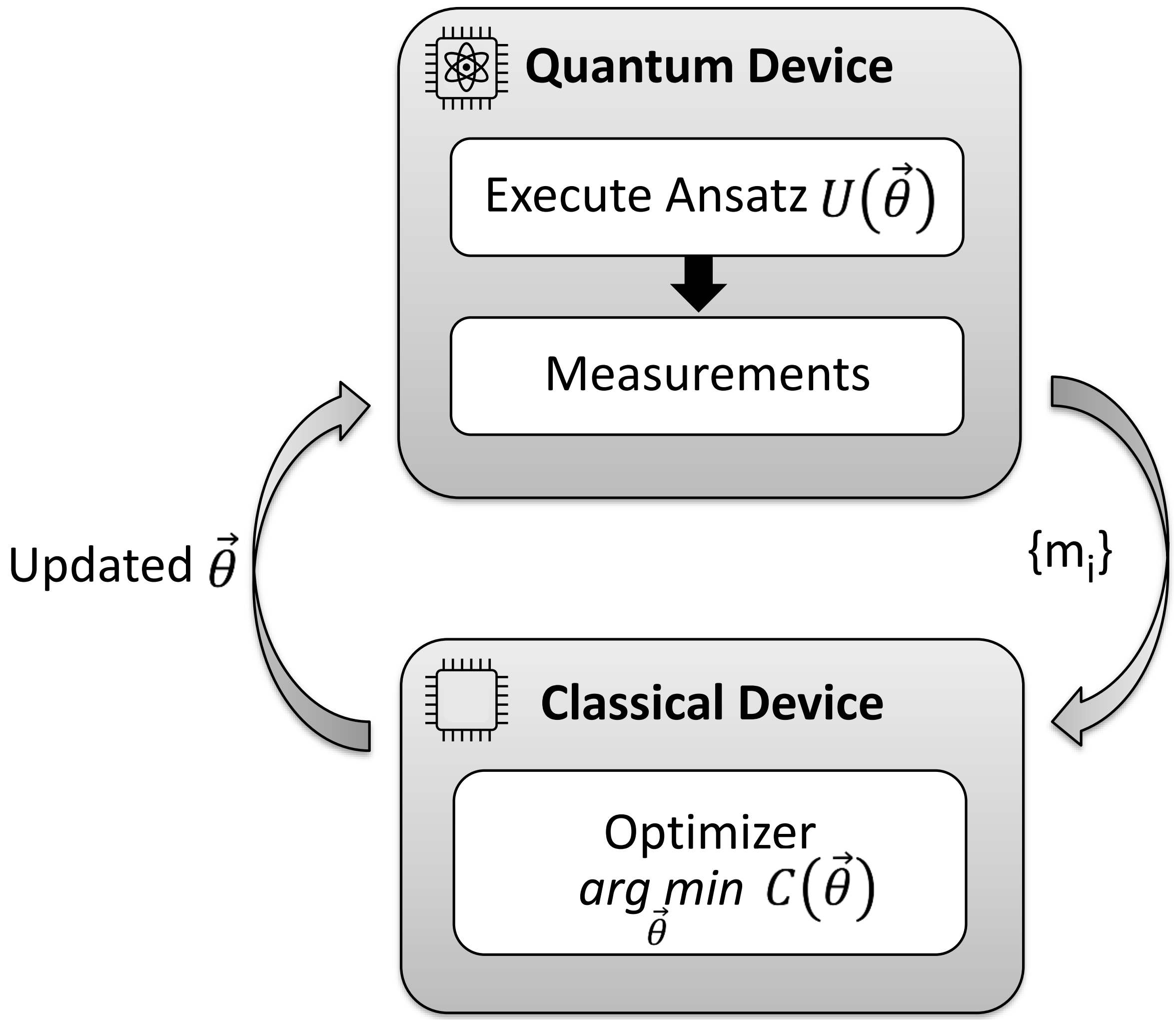


Quantum Device
Execute Ansatz $U(\vec{\theta})$
Measurements
Updated $\vec{\theta}$
$\{m_i\}$
Classical Device
Optimizer
$\arg\min_{\vec{\theta}} C(\vec{\theta})$

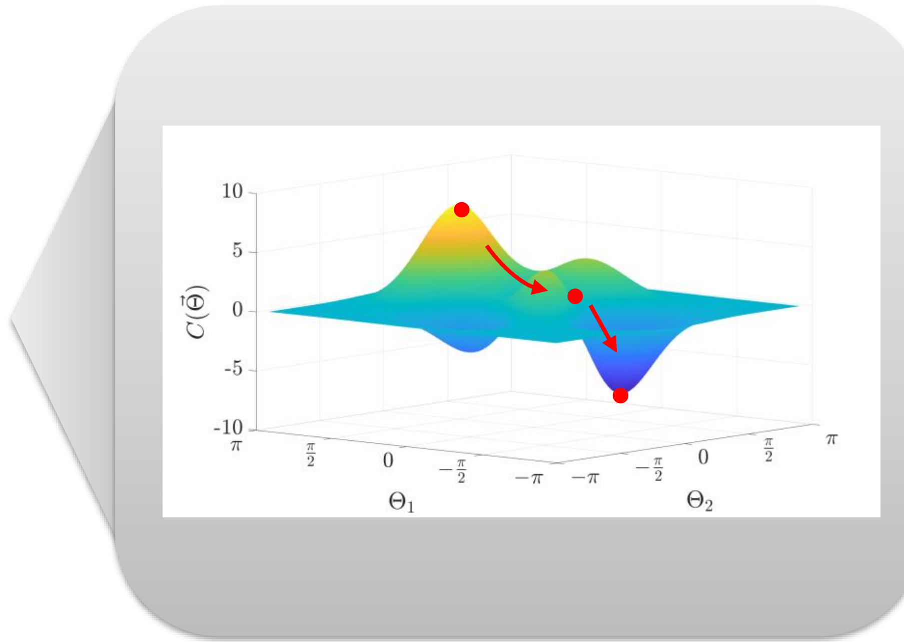



$C(\vec{\Theta})$
10
5
0
-5
-10
$\pi$
$\frac{\pi}{2}$
0
$-\frac{\pi}{2}$
$-\pi$
$\Theta_1$
$-\pi$
$-\frac{\pi}{2}$
0
$\frac{\pi}{2}$
$\pi$
$\Theta_2$

# Motivation



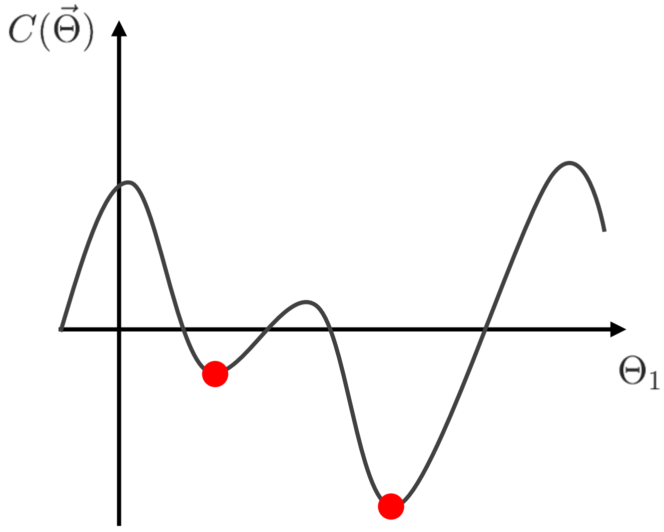
Quantum Device
Execute Ansatz $U(\vec{\theta})$
Measurements
Updated $\vec{\theta}$
$\{m_i\}$
Classical Device
Optimizer
$\arg\min_{\vec{\theta}} C(\vec{\theta})$

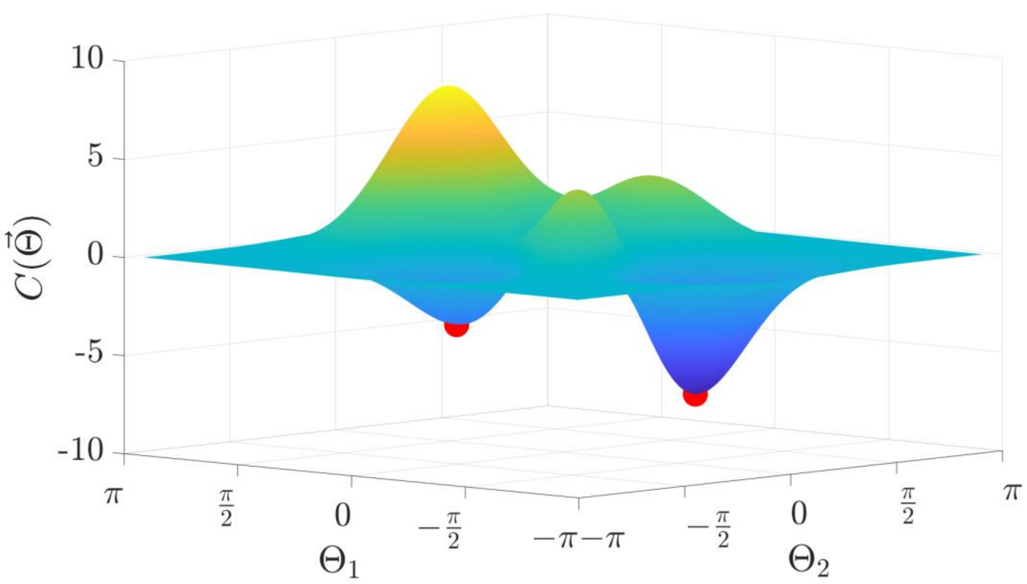



$C(\vec{\Theta})$
10
5
0
-5
-10
$\Theta_1$
$\Theta_2$

# Motivation

- Get stuck in a local minimum instead of a global minimum



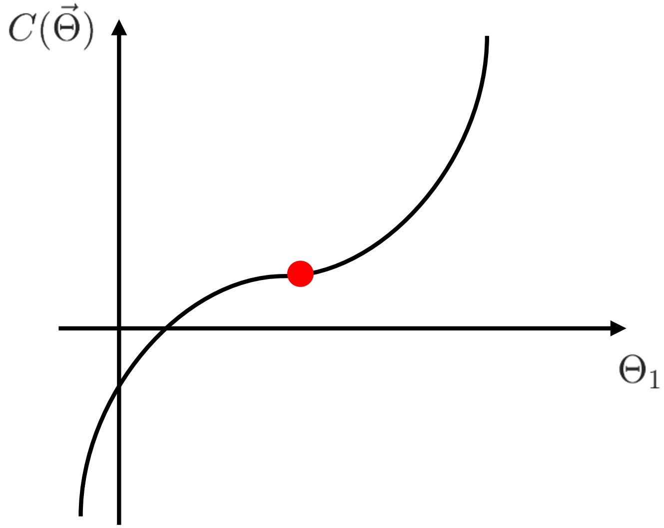
# Motivation

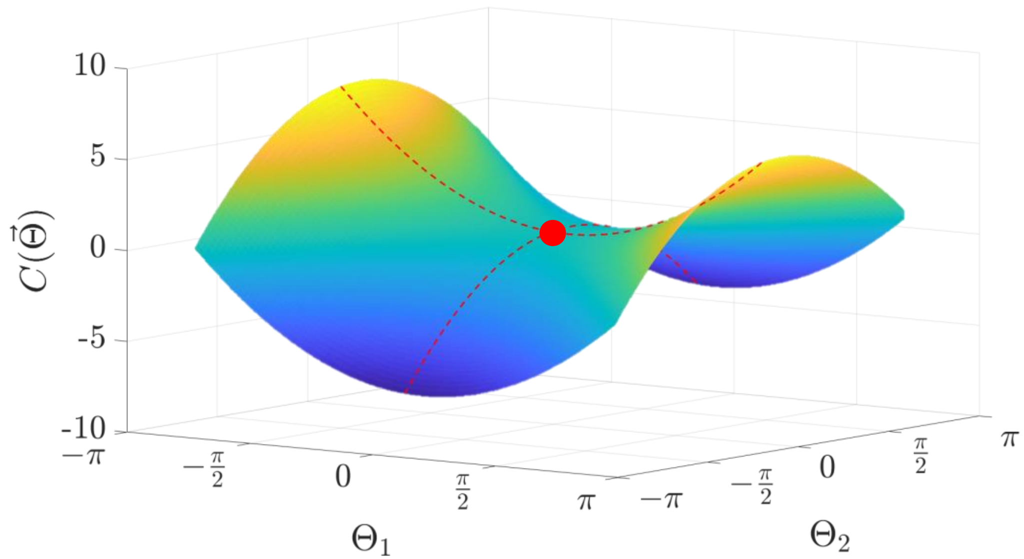
- **Features of the Cost Landscape:**

Local optima

# Motivation

- Likelihood of being a saddle point increases with the number of dimensions

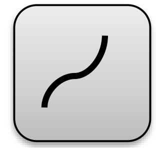


# Motivation

- **Features of the Cost Landscape:**

Local optima

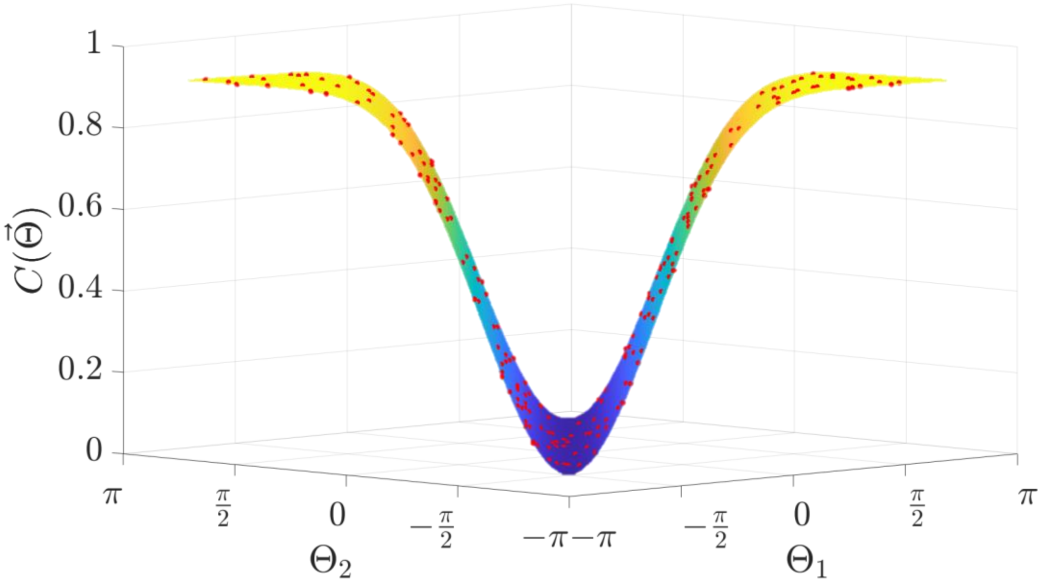


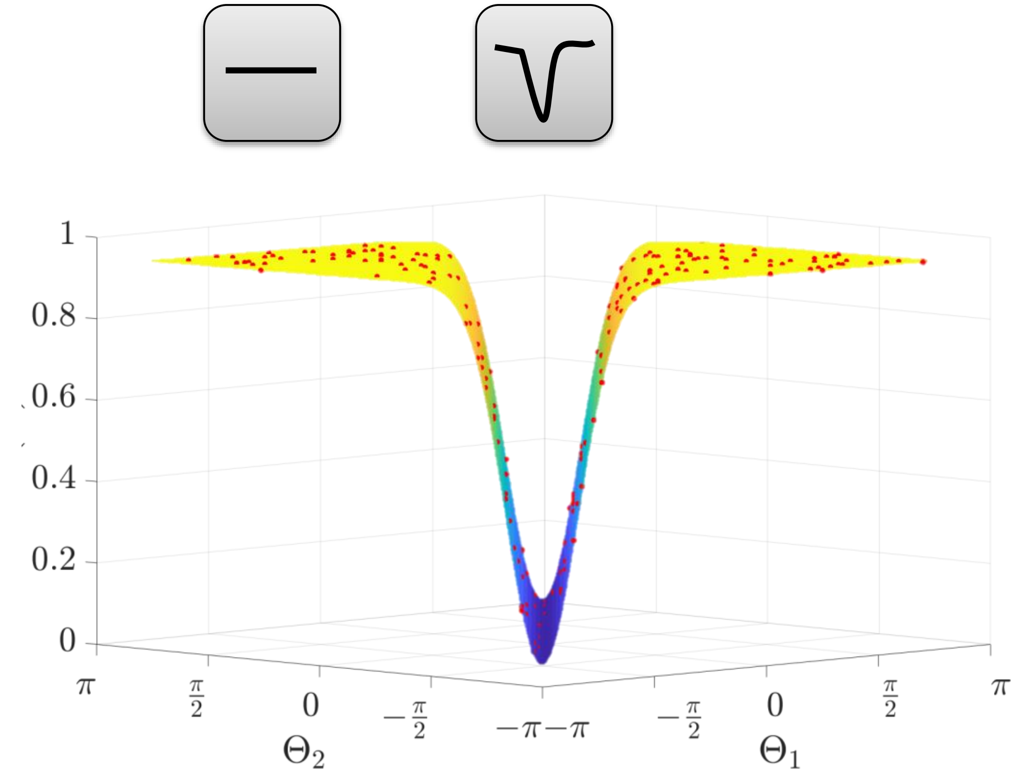
Saddle points

# Motivation

- Many points on the flat area, where the gradient is zero or close to zero, providing no guidance

→ Bad starting points for optimization

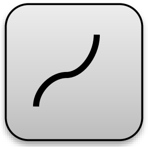


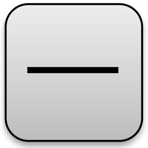
# Motivation

- **Features of the Cost Landscape:**

Local optima



Saddle points

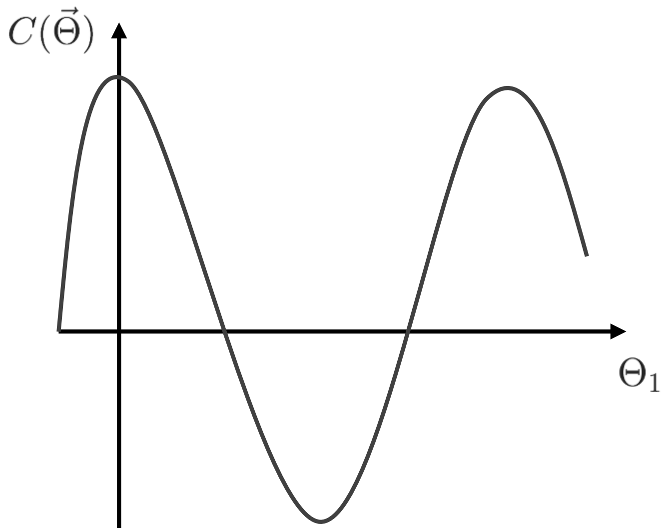


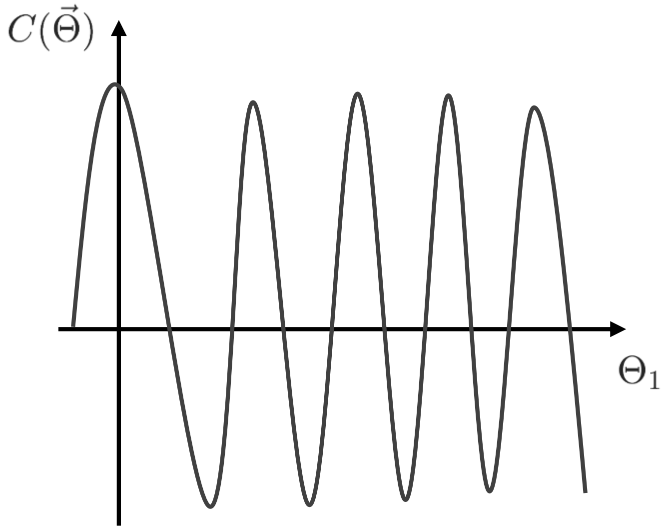
Barren plateaus

Narrow gorges

# Background – Frequency

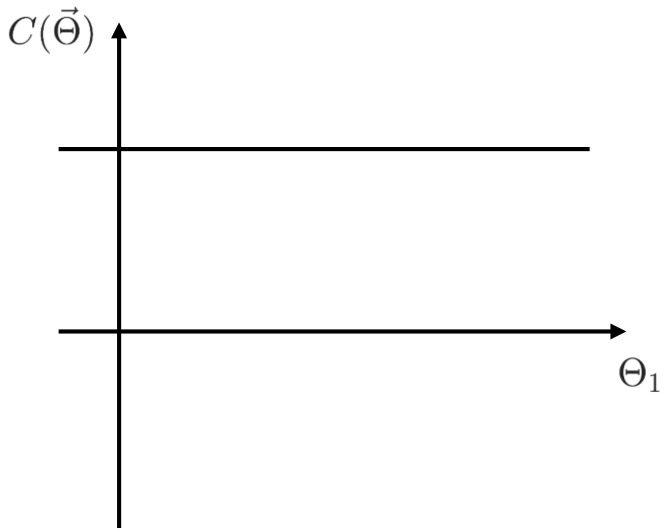
- The likelihood of suboptimal solutions increases with the frequency of each feature

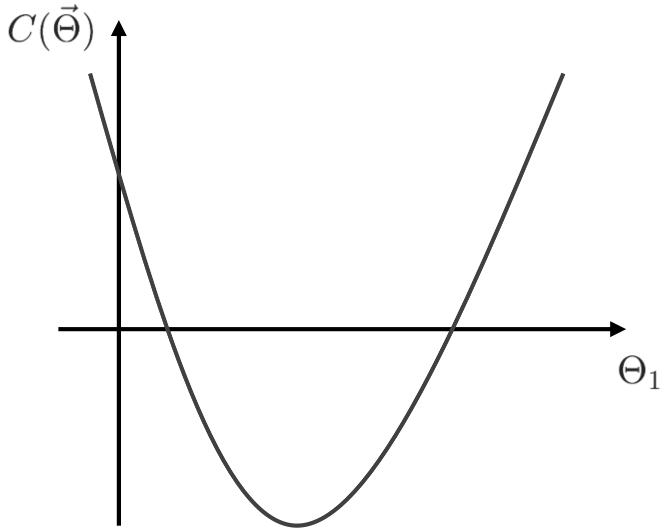


# Background – Gradient Behavior

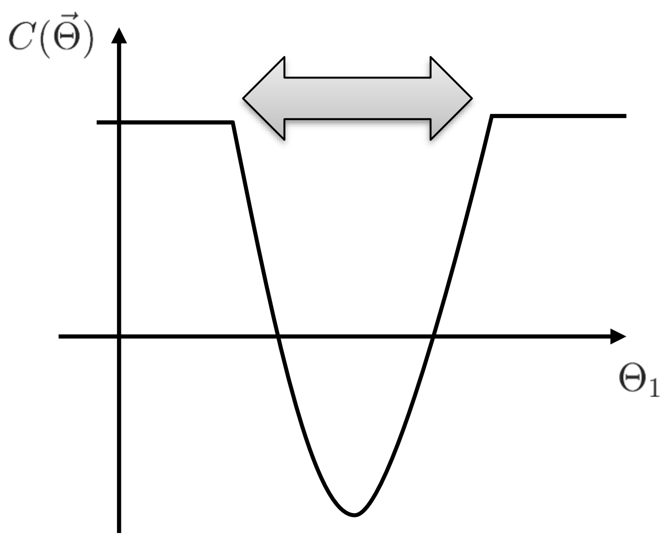
- Optimizer is not able to navigate in an improving direction

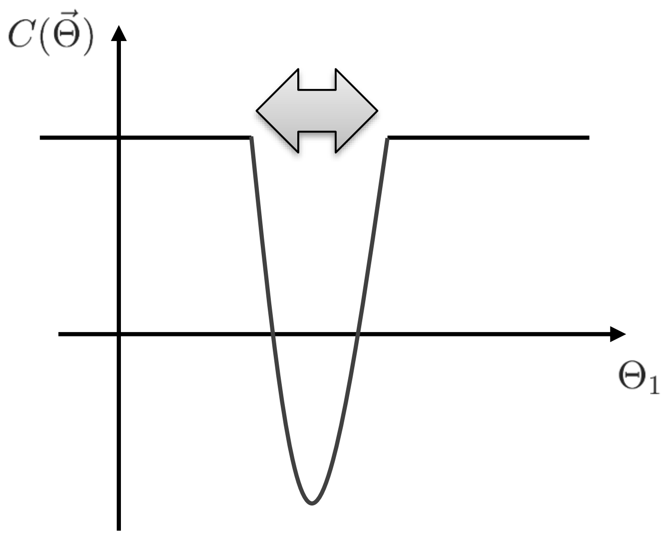


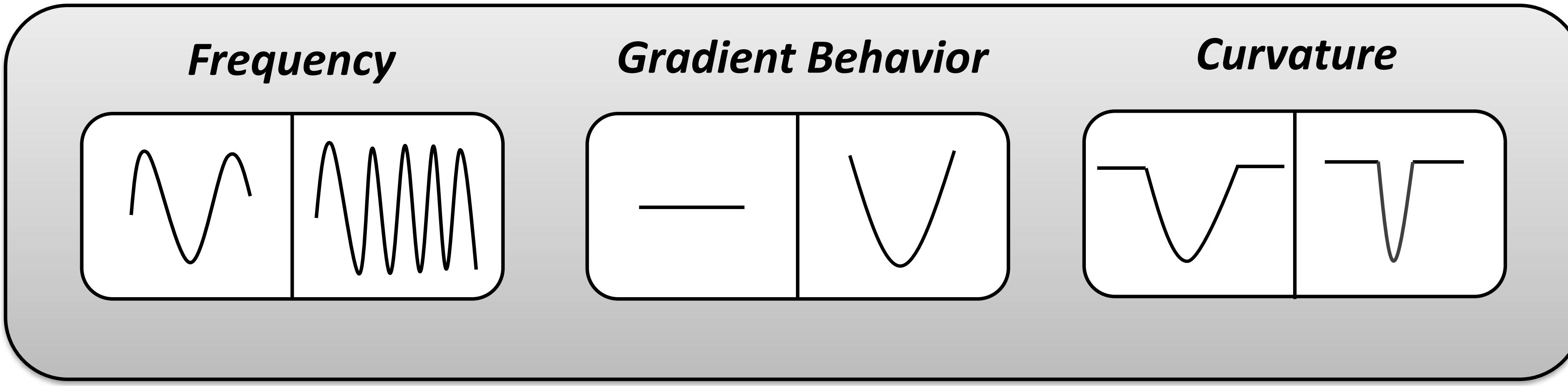
# Background – Curvature

- Optimizer has to adapt step size based on the size of the feature

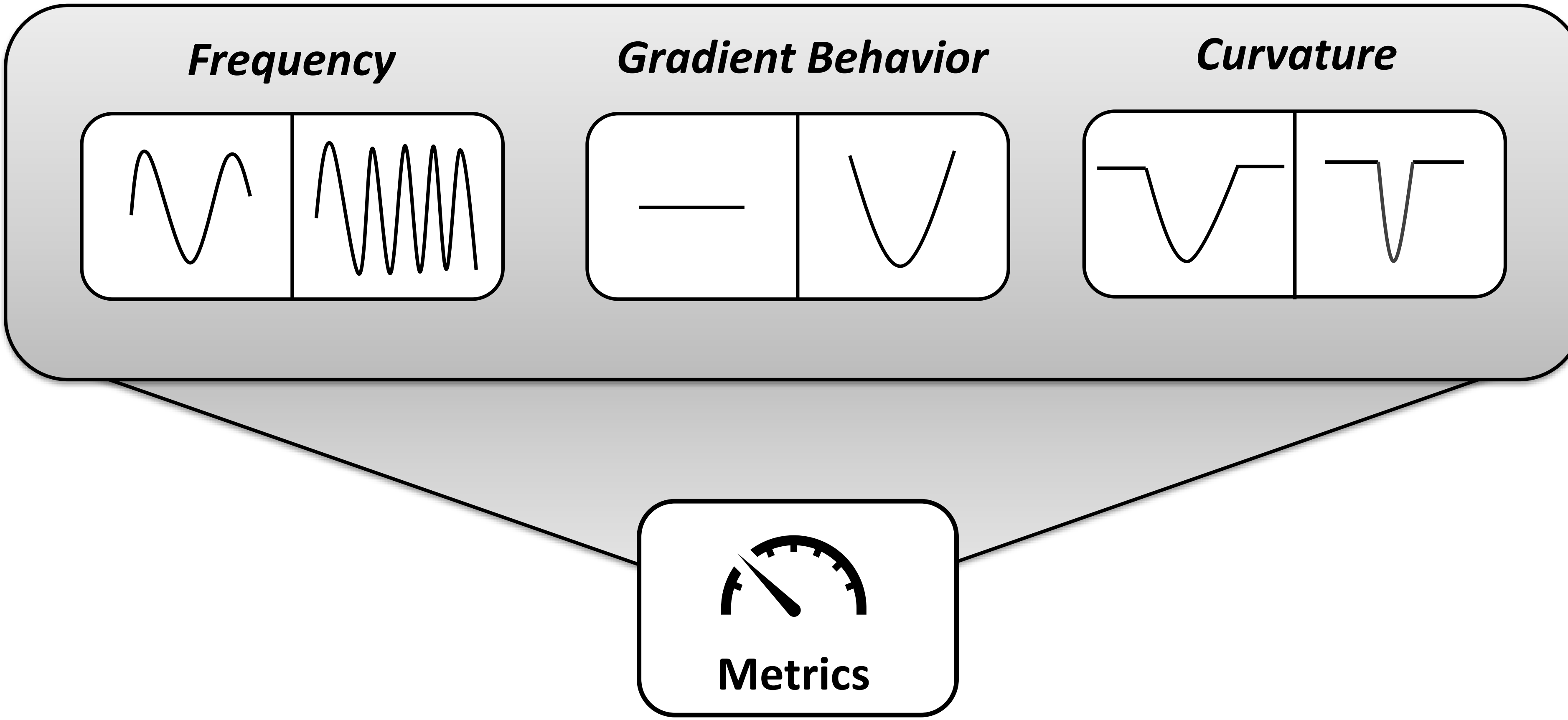


# Background



# Background



# Metrics – Overview

- Total Variation
- Fourier Density
- Inverse Gradient Standard Deviation
- Scalar Curvature

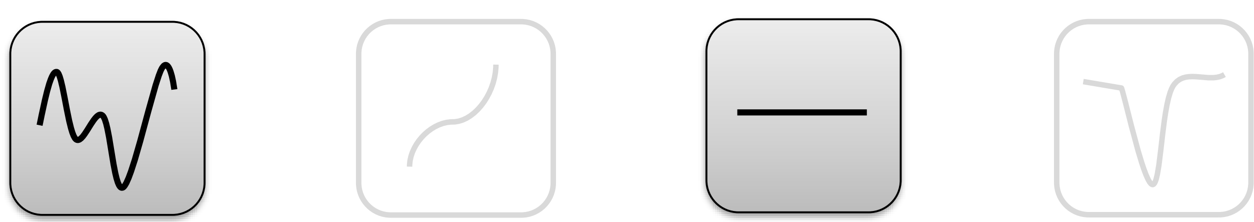
# Metrics – Total Variation

- Total Variation ($TV$) captures how much a function varies across a finite domain

$$TV(C) = \int_{\vec{\theta} \in A} \left| \nabla C(\vec{\theta}) \right| d\vec{\theta}$$

where $\nabla C$ is the gradient of the cost function

- High $TV$ values → rough landscape
- Low $TV$ values → flat areas



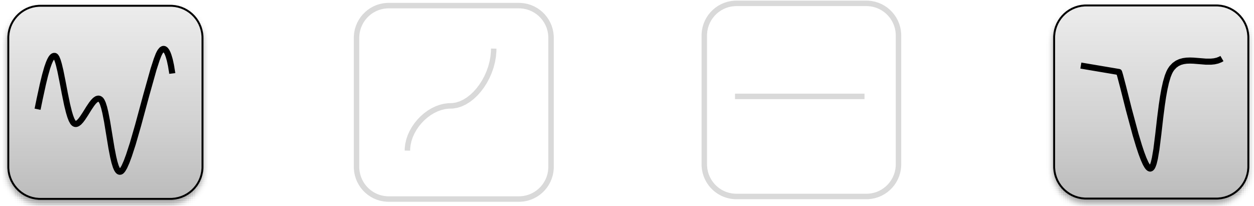
# Metrics – Fourier Density

- Fourier Density ($FD$) is defined as the number of nonzero Fourier coefficients

$$FD(C) = \frac{\|\vec{c}_\omega\|_1^2}{\|\vec{c}_\omega\|_2^2}$$

where $\vec{c}_\omega$ is a vector of Fourier coefficients, $\|\cdot\|_p$ is the $p$-Norm

- High $FD$ values → sharp transitions
- Low $FD$ values → less rough landscape



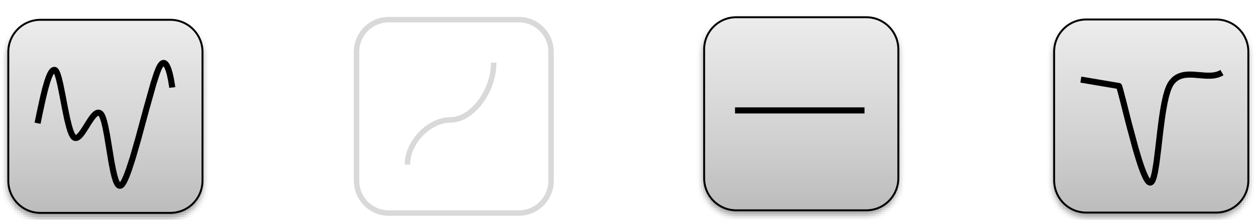
# Metrics – Fourier Density

- Fourier Density ($FD$) is defined as the number of nonzero Fourier coefficients

$$FD(C) = \frac{\|\vec{c}_\omega\|_1^2}{\|\vec{c}_\omega\|_2^2}$$

where $\vec{c}_\omega$ is a vector of Fourier coefficients, $\|\cdot\|_p$ is the $p$-Norm

- High $FD$ values → sharp transitions
- Low $FD$ values → less rough landscape
- Low $TV$ values & high $FD$ values → indicator for barren plateaus, narrow gorges



Research

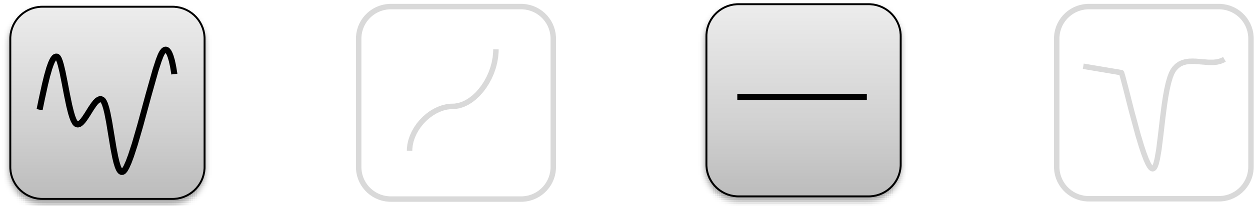
# Metrics – Inverse Gradient Standard Deviation

- Inverse Gradient Standard Deviation ($IGSD_i$) describes the variation of the gradient values from the gradient mean in the $i$-th direction given by

$$IGSD_i(C) = \frac{1}{SD\left(\partial_i C(\vec{\theta})\right)}$$

with $SD$ is the standard deviation and $\partial_i$ is the derivative of the cost function

- High $IGSD_i$ values → flat areas
- Low $IGSD_i$ values → high dispersion of values



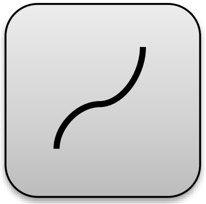
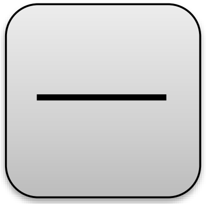
# Metrics – Scalar Curvature

- Scalar Curvature ($SC$) is defined as

$$SC(P,C) = \beta\left(Tr\left(H_C(P)\right)^2 - Tr(H_C^2(P))\right) + 2\beta^2\left(\nabla C^T(P)\left(H_C^2(P) - Tr\left(H_C(P)\right)H_C(P)\nabla C(P)\right)\right)$$

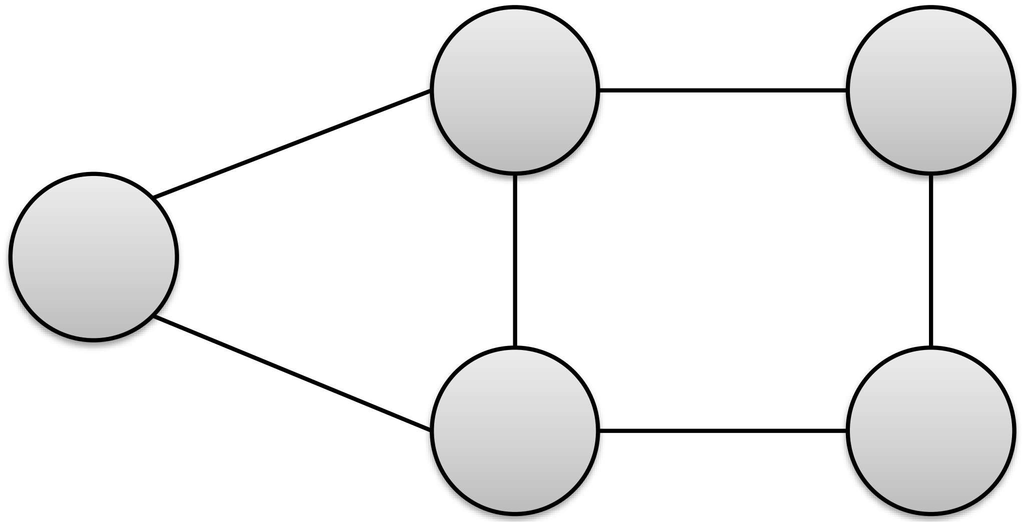
with $\beta = (1 + \|\nabla C(P)\|^2)^{-1}$, $P$ point on the landscape, $\nabla C$ Gradient of $C$ and $H_C$ is the Hessian of $C$

- $SC(P,C) > 0$ → local optimum
- $SC(P,C) < 0$ → saddle point
- $SC(P,C) = 0$ → flat

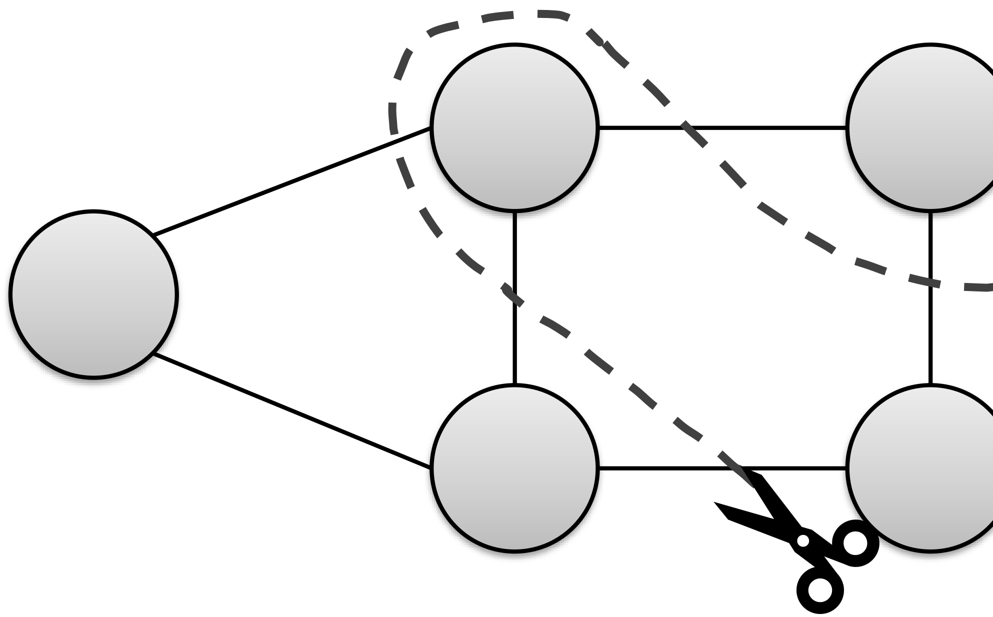

# Use Case – Solving the Maximum Cut Problem

- The maximum cut problem is to partition the node set of a graph into two sets such that the number of edges between nodes of the different sets is maximal



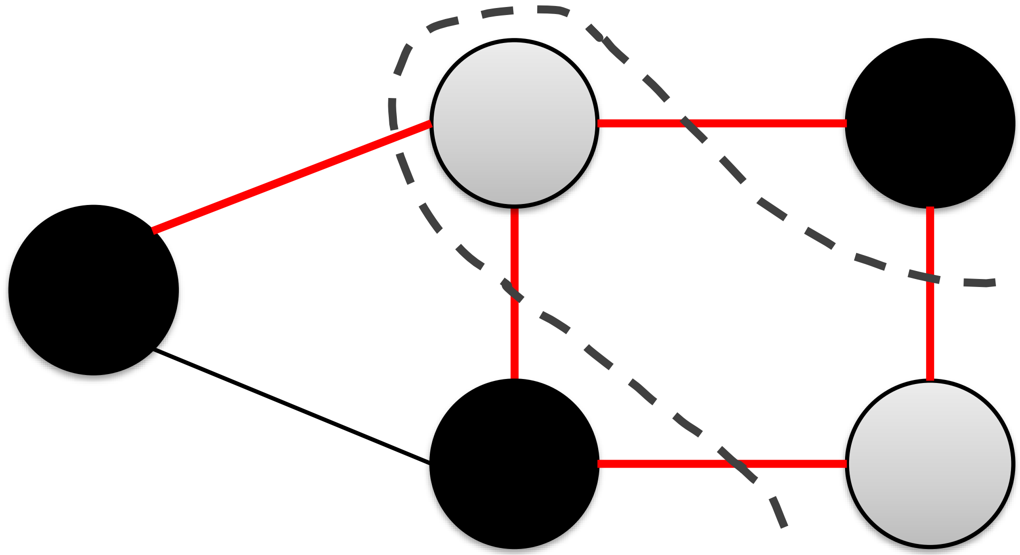
# Use Case – Solving the Maximum Cut Problem

- The maximum cut problem is to partition the node set of a graph into two sets such that the number of edges between nodes of the different sets is maximal

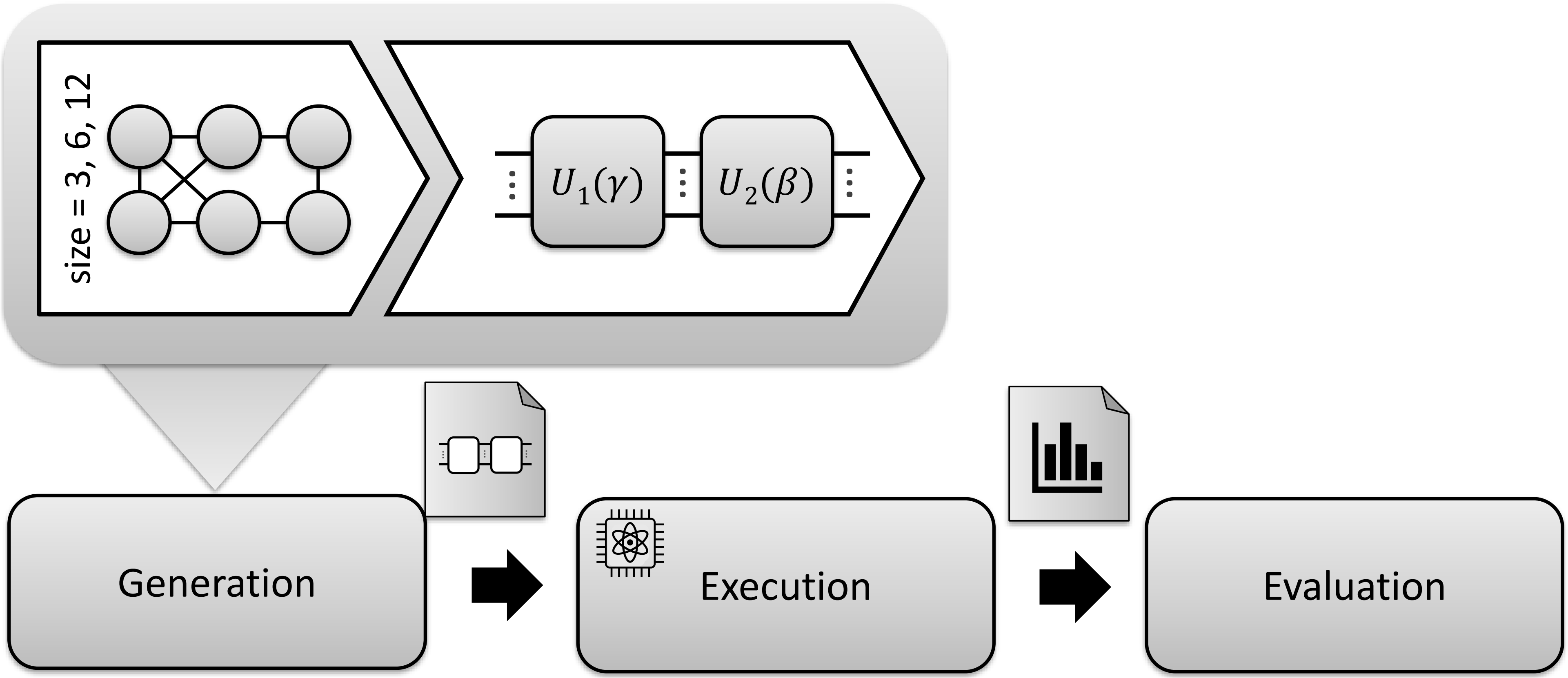


# Use Case – Solving the Maximum Cut Problem

- The maximum cut problem is to partition the node set of a graph into two sets such that the number of edges between nodes of the different sets is maximal



# Use Case – Results

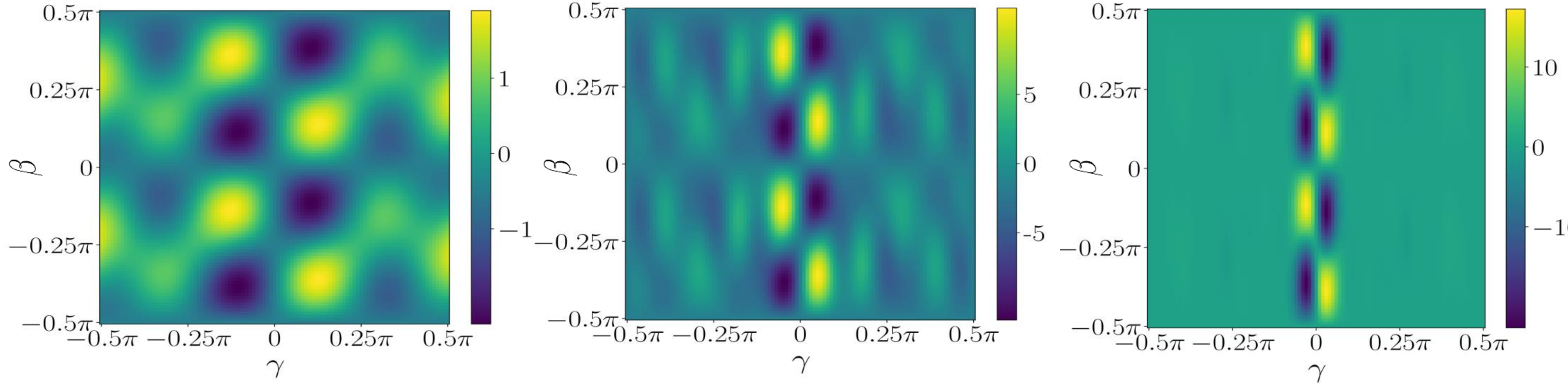


# Use Case – Results

- With increasing number of nodes, the margins get flatter and the cost landscape concentrates in the center

  → Less local optima

size = 3 | size = 6 | size = 12

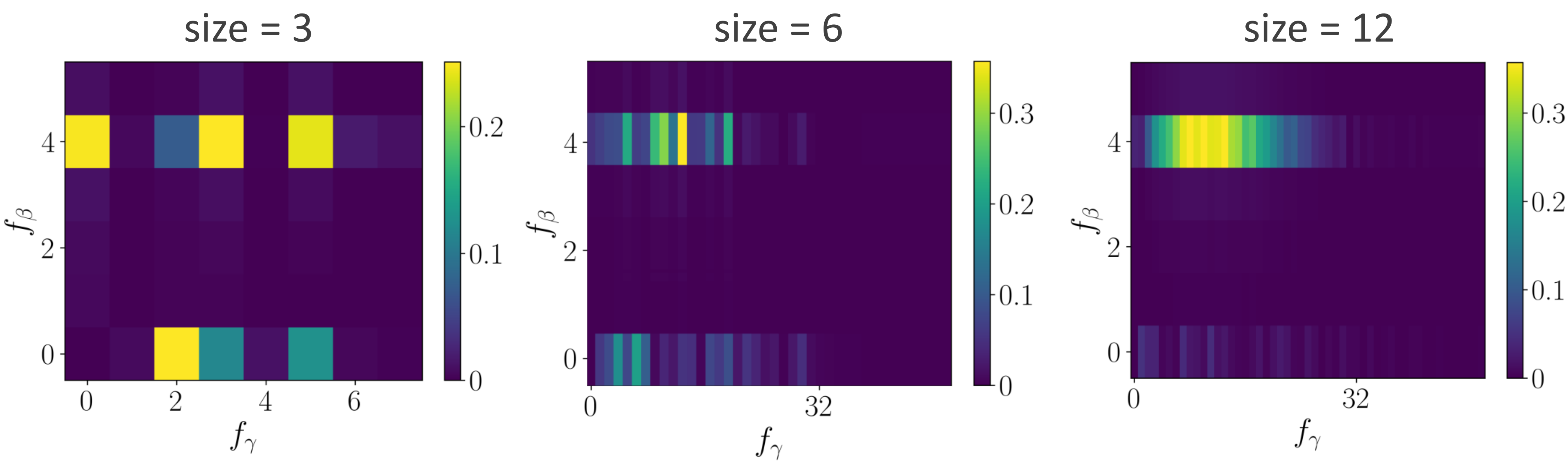



$TV$ = 4.12 | $TV$ = 3.80 | $TV$ = 1.83

# Use Case – Results

- The progression of $FD$ values from 25.07 to 210.19 indicates a steepening of the cost landscape

  → Cancellation of most of the frequencies

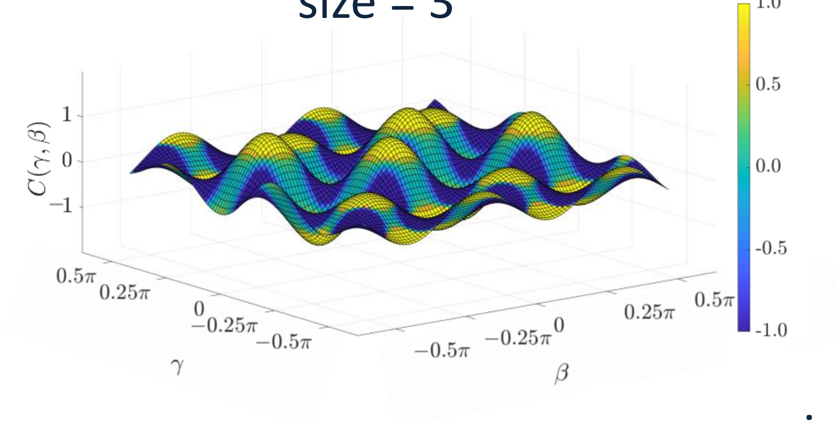



$FD$ = 25.07 $FD$ = 179.88 $FD$ = 210.19

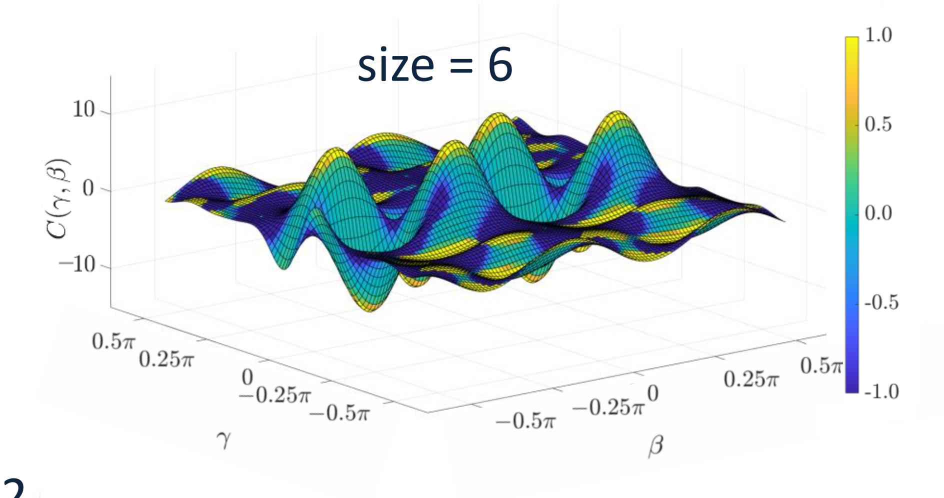
# Use Case – Results

- With increasing number of vertices the landscape gets flatter and the local optima become steeper

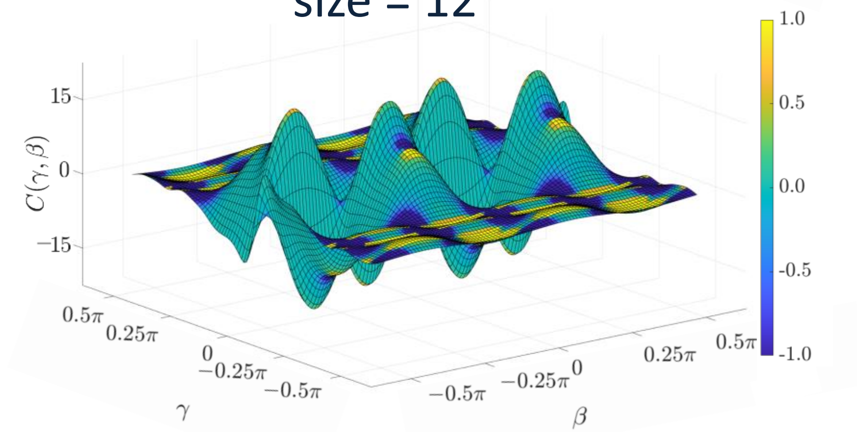
size = 3




size = 6




size = 12




$IGSD_\gamma$ = 0.94
$IGSD_\beta$ = 0.85

$IGSD_\gamma$ = 1.93
$IGSD_\beta$ = 1.77

$IGSD_\gamma$ = 4.02
$IGSD_\beta$ = 3.38

Research

# Conclusion & Future Work

- **Conclusion:**
  - Identification of obstacles in the optimization process
  - Overview of different metrics and their interplay

- **Future Work:**
  - Development of new metrics
  - Include these metrics in the classical optimization process

Thank you for your attention ☺