AI for Laser Micromachining

S. Walker[1], D. Brinkmeier[2], T. Menold[2], A. Michalowski[2] and B. Neuenschwander[1]

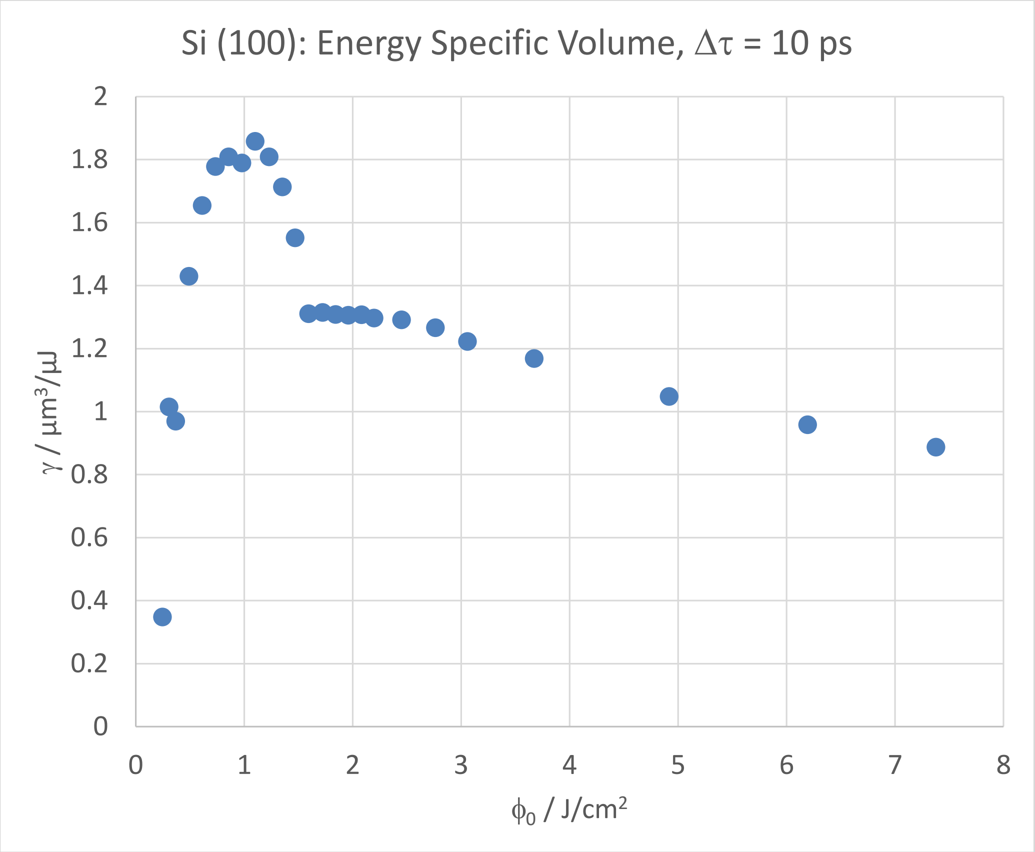
- Institute for Applied Laser, Photonics and Surface Technologies ALPS

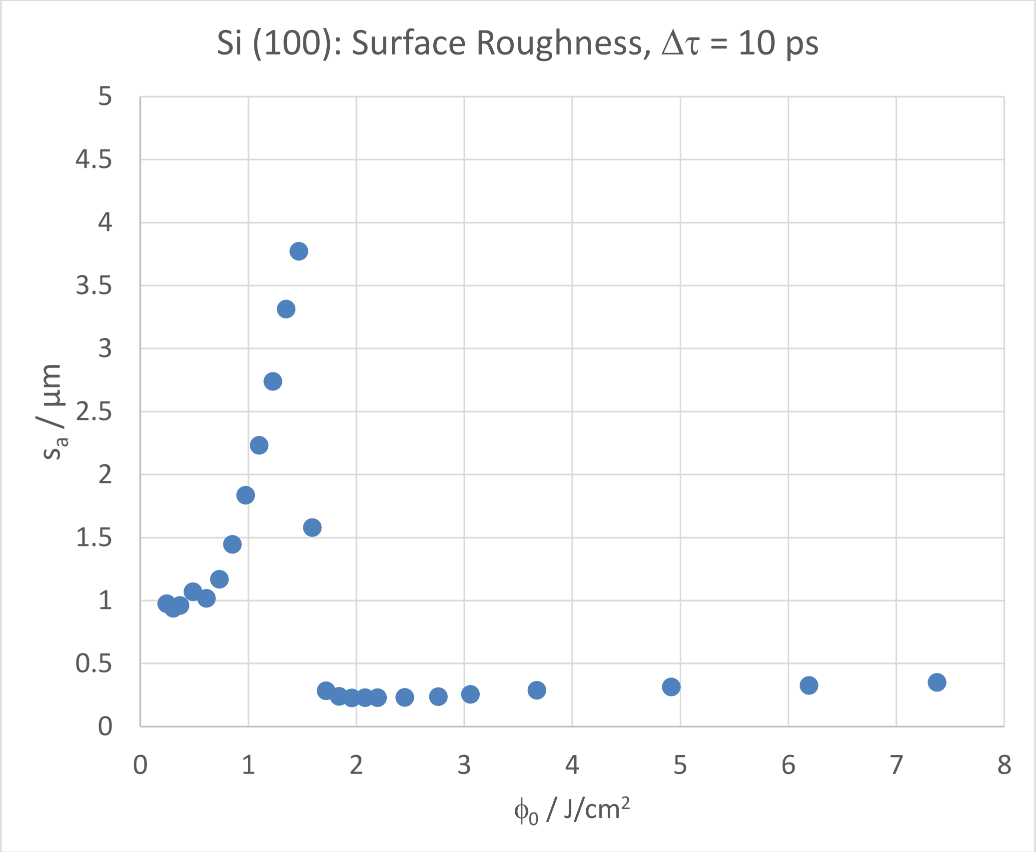
Outline

- Motivation
- Bayesian Optimization
- Examples:
  - High Throughput Maching with a 2 Step Process
  - Steel surfaces
- Summary and Conclusion

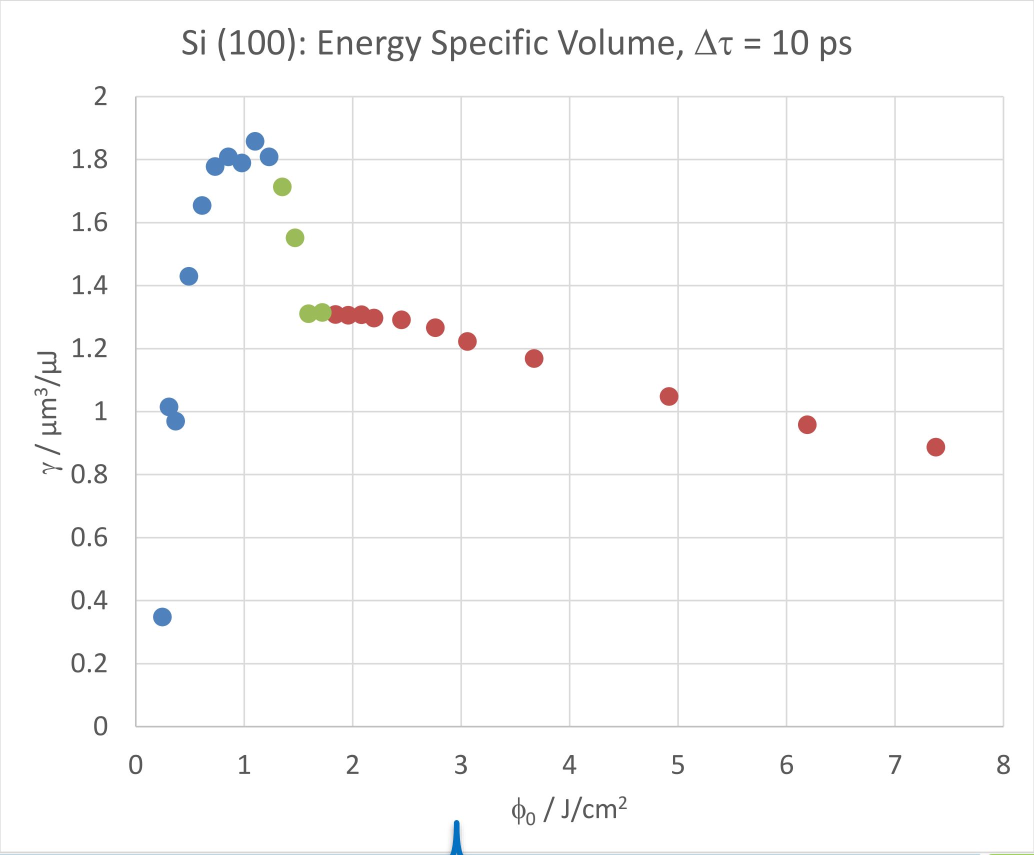
- Motivation

Si (100), 10 ps: Energy Specific Volume and Roughness

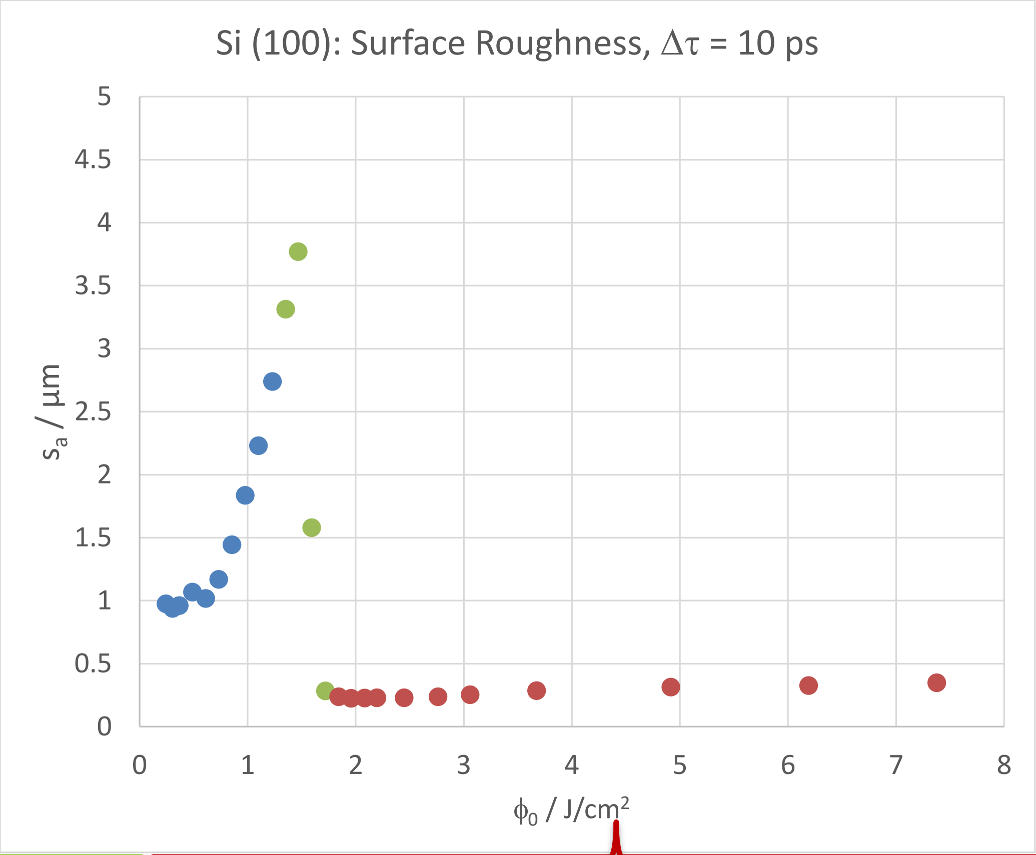


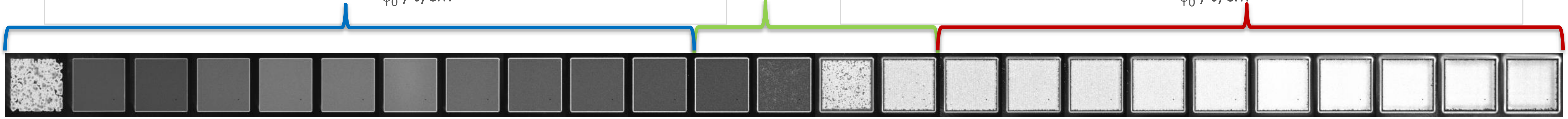


Si (100), 10 ps: Energy Specific Volume and Roughness

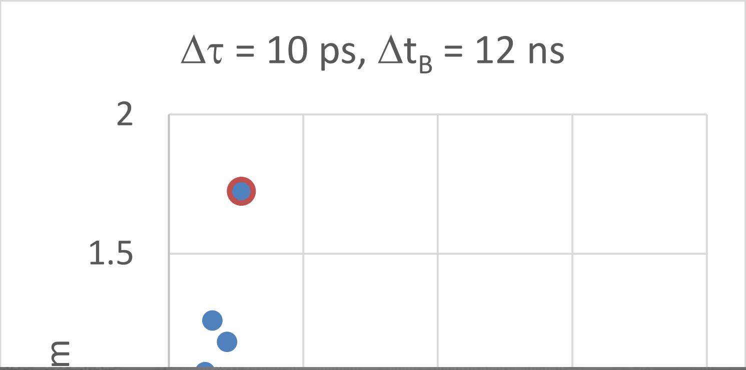


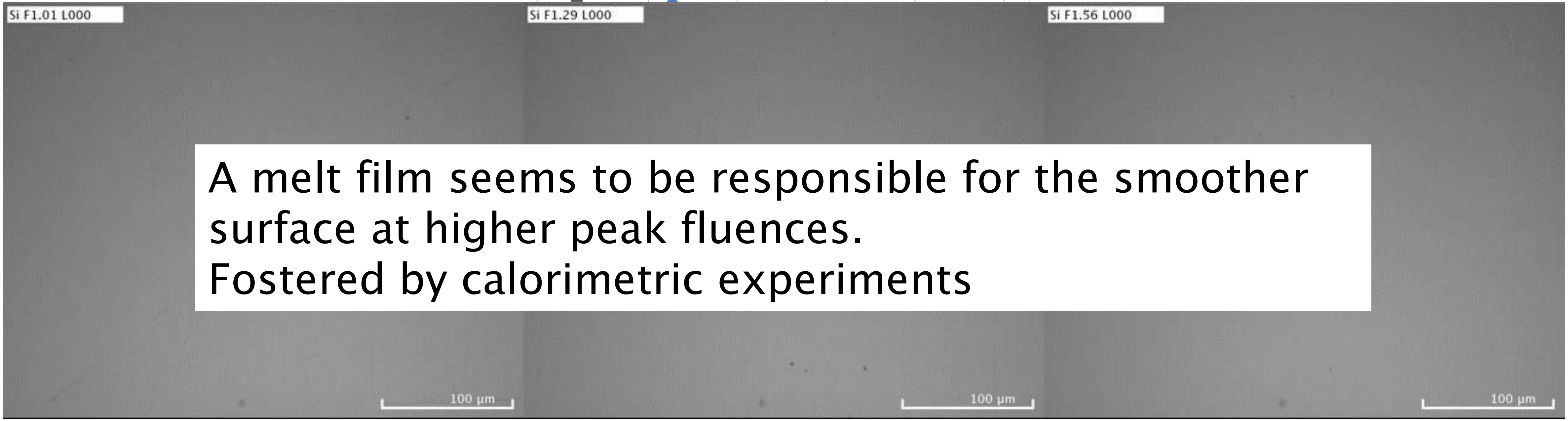
Clear transition between two regimes





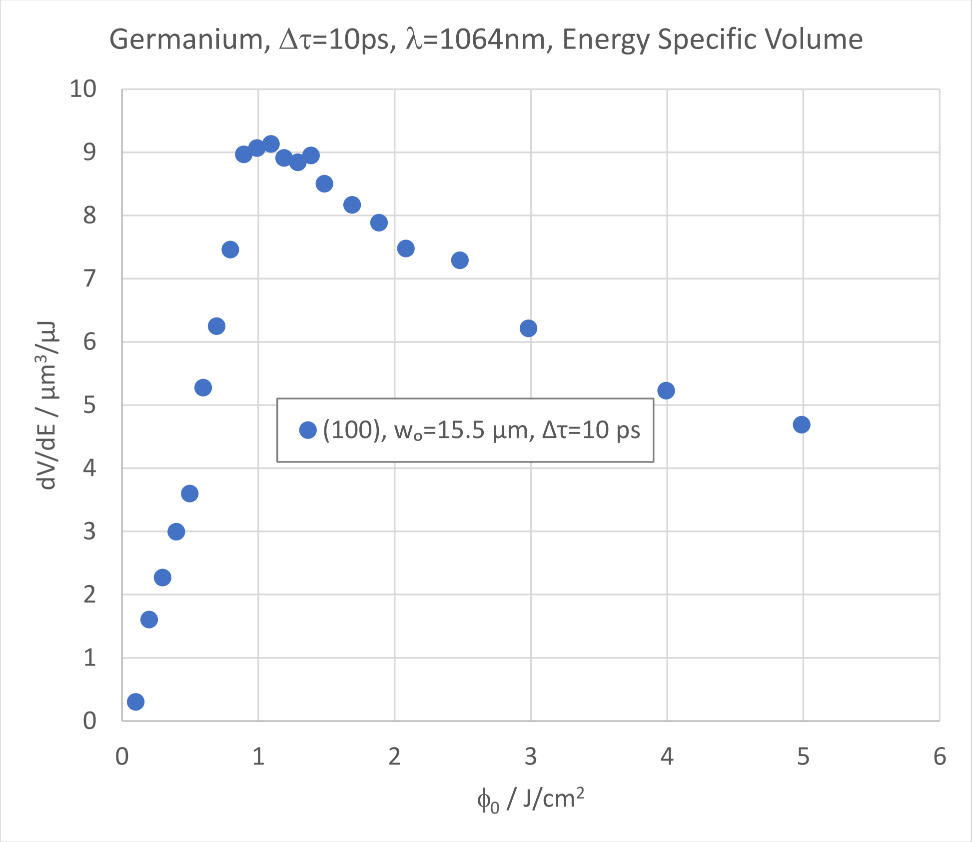
Si (100), 10 ps: Transition Region (Former Experiments)

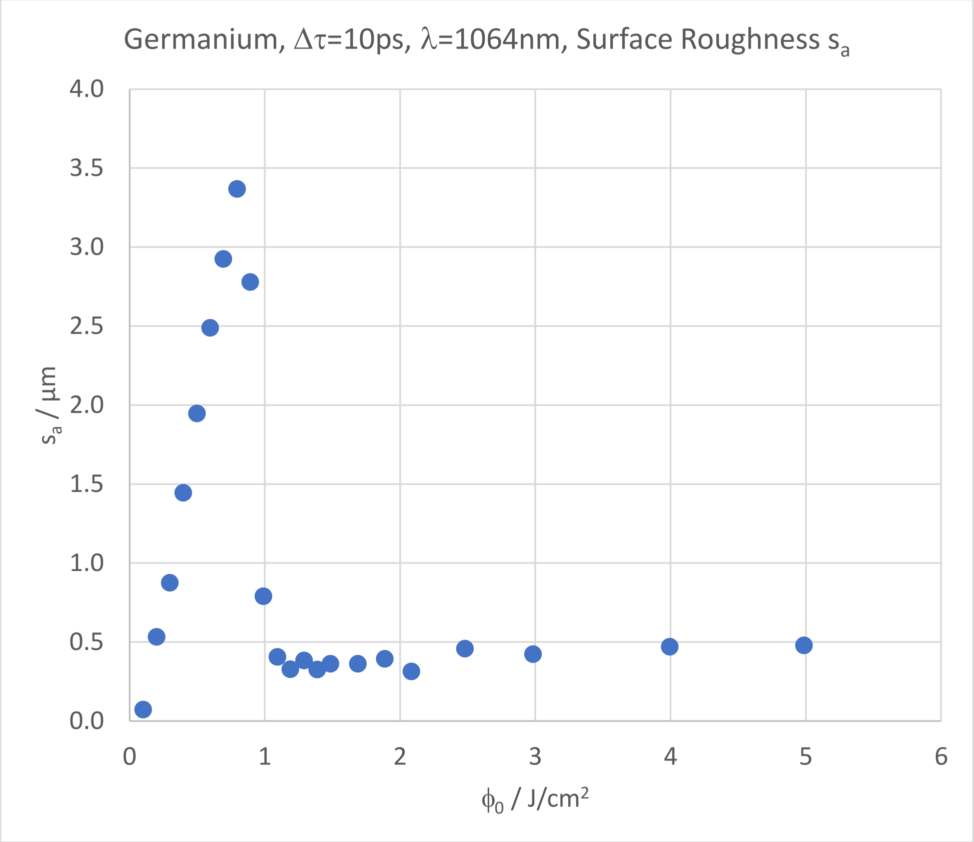




A melt film seems to be responsible for the smoother surface at higher peak fluences.
Fostered by calorimetric experiments

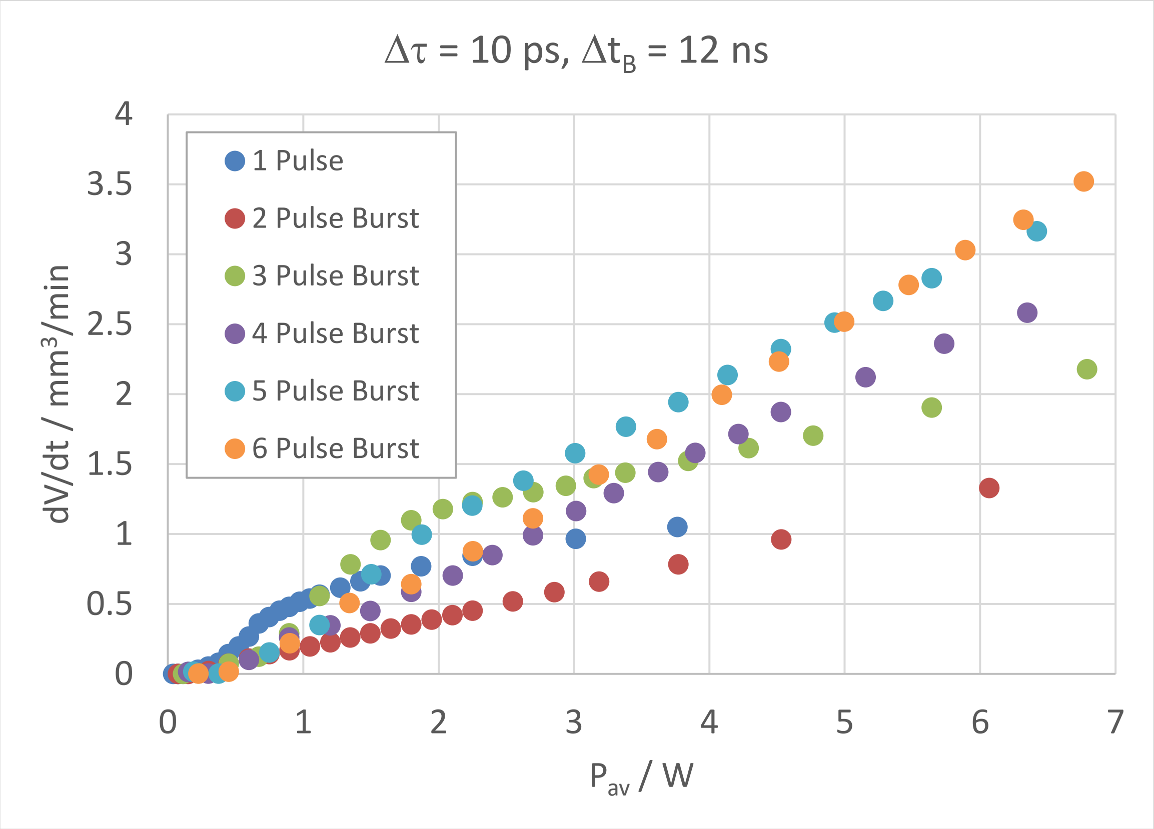
Ge (100), 10 ps: Energy Specific Volume and Roughness

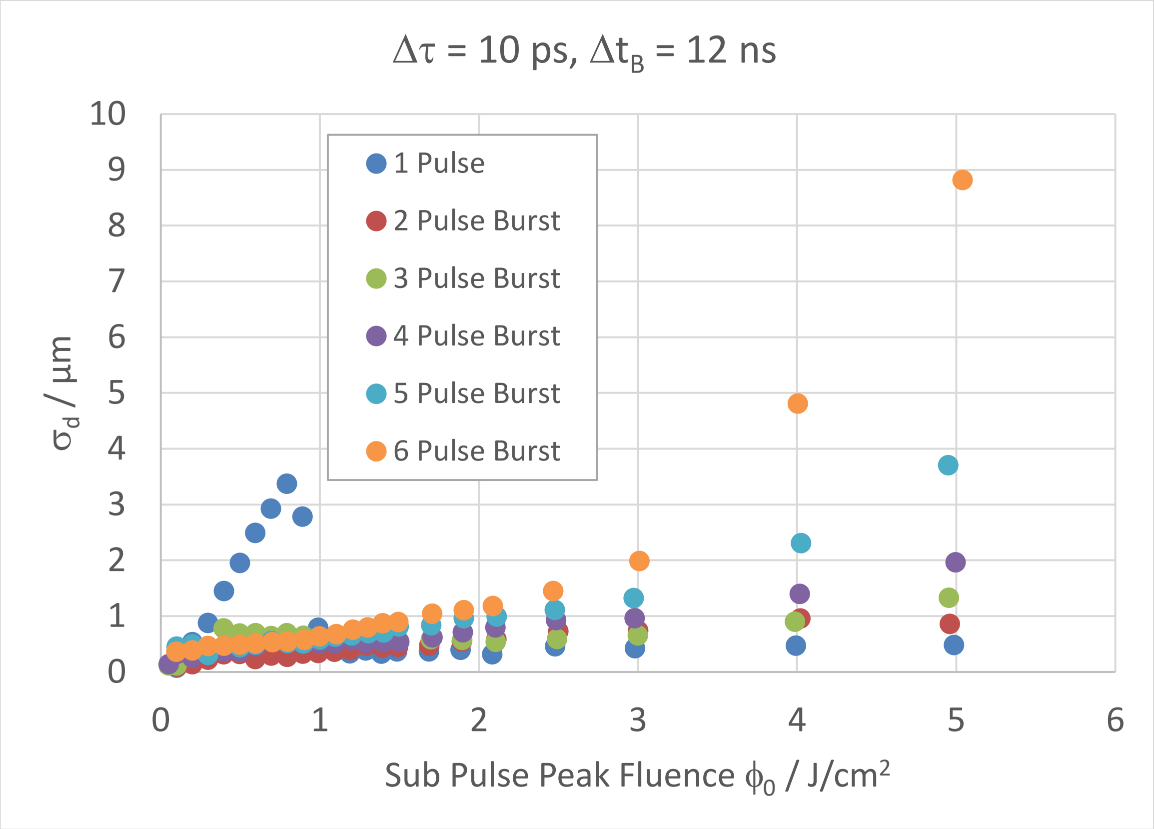




- Identical behavior for the surface roughness for Germanium.

Ge (100), 10 ps: Bursts

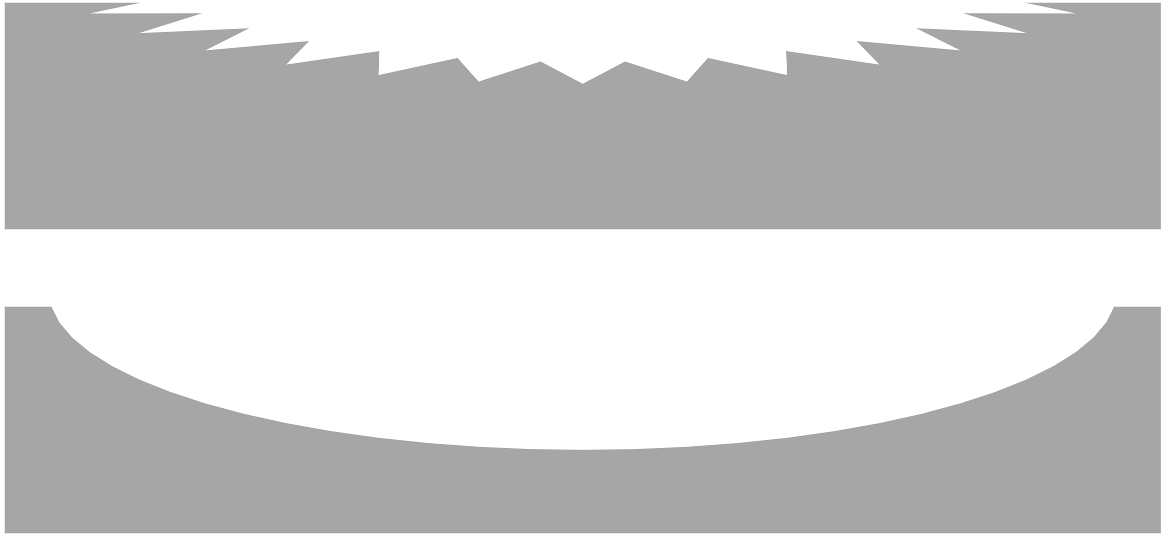




- Higher removal rates with bursts

- Identical behavior of the surface roughness
- But increasing surface roughness with bursts.

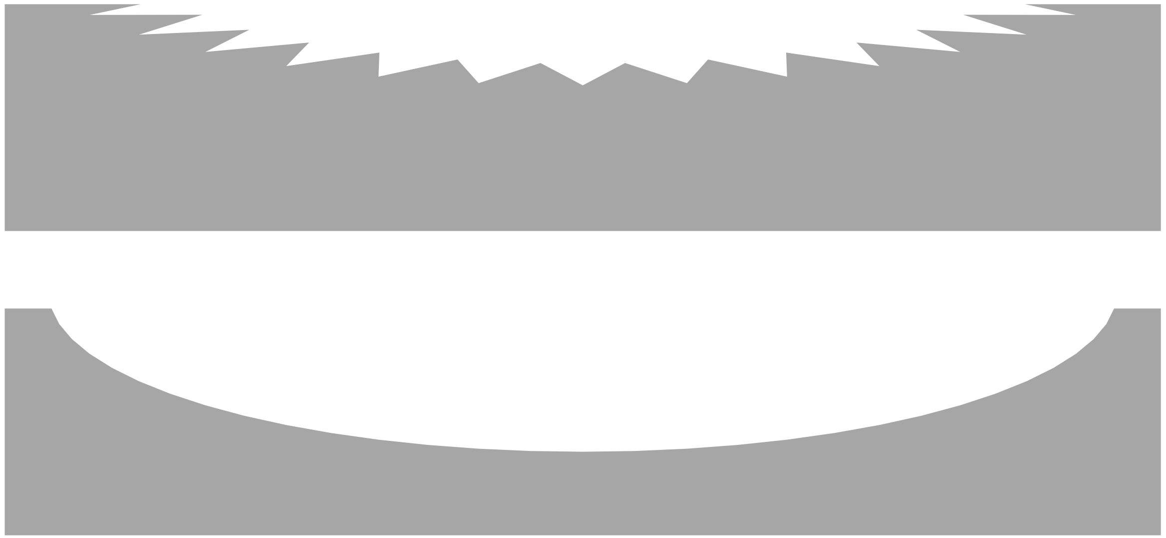
Idea for a 2-Step Process: Roughening - Smoothening



- 1. Step: Roughing
  - High removal rate with bursts
  - Rough surface
- 2. Step: Smoothening
  - Smoothening with single pulses with fluence above the transition

- To optimize: s_a (min), dV/dt (max)
- Parameters which can be varied:
  - Roughing: $n_{burst}, \phi_0, w_0, p_x, p_y, f_{rep}, n_{layer}$
  - Smoothening: $\phi_0, w_0, p_x, p_y, f_{rep}, n_{layer}$
  - For only 5 values per parameter $\approx 1.2 \cdot 10^9$ Experiments

Idea for a 2-Step Process: Roughening - Smoothening



- 1. Step: Roughing
  - High removal rate with bursts
  - Rough surface
- 2. Step: Smoothening
  - Smoothening with single pulses with fluence above the transition

- To optimize: s_a (min), dV/dt (max)
- Parameters which can be varied:
  - Roughing: $n_{burst}, \phi_0, w_0, p_x, p_y, f_{rep}, n_{layer}$
  - Smoothening: $\phi_0, w_0, p_x, p_y, f_{rep}, n_{layer}$
  - For only 5 values per parameter $\approx 1.2 \cdot 10^9$ Experiments
  - Even when we reduce $\approx 3'125$ Experiments
  - Can machine learning (ML) help?

- Bayesian Optimization

Process Optimization by Bayesian Optimization

Parameter Set $\vec{P}$

- n_b: #Pulses per Burst
- $\phi_{0,r}$: Peak Fluence of a Single Pulse
- $n_{l,r}$: #Layer roughening
- $\phi_{0,s}$: Peak Fluence of a Single Pulse
- $n_{l,s}$: #Layer smoothening
- $p_{x,s}$: Pulse-pulse distance smoothening
- $p_{y,s}$: Line-line distance smoothening
- etc.

Optimization Target

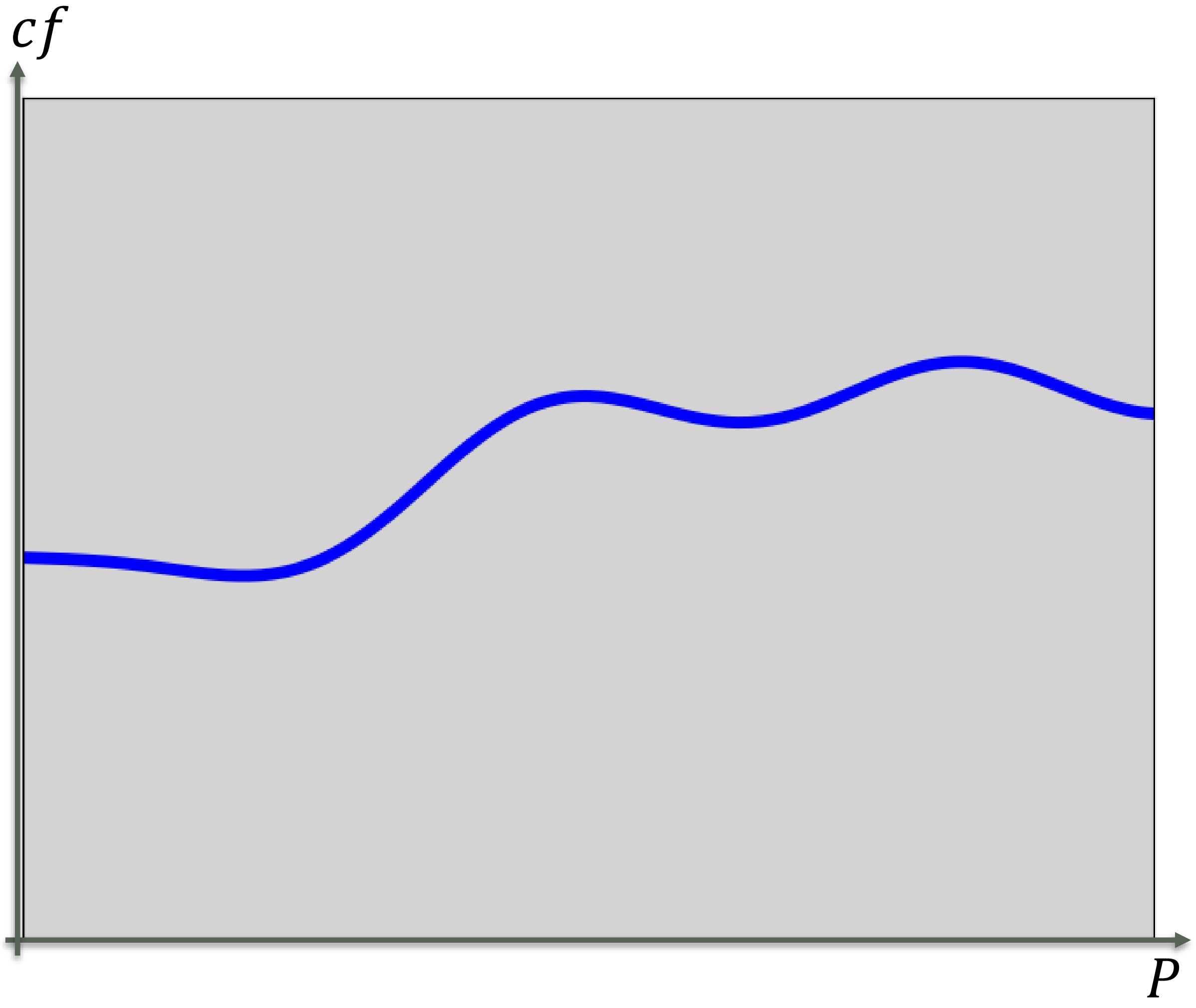
- s_a: Surface roughness
- dV/dt: Removal rate
  directly scales with the average depth per layer t

- Define scalar cost Function:

$$cf(\vec{P}) = f(s_a(\vec{P}), t(\vec{P}))$$

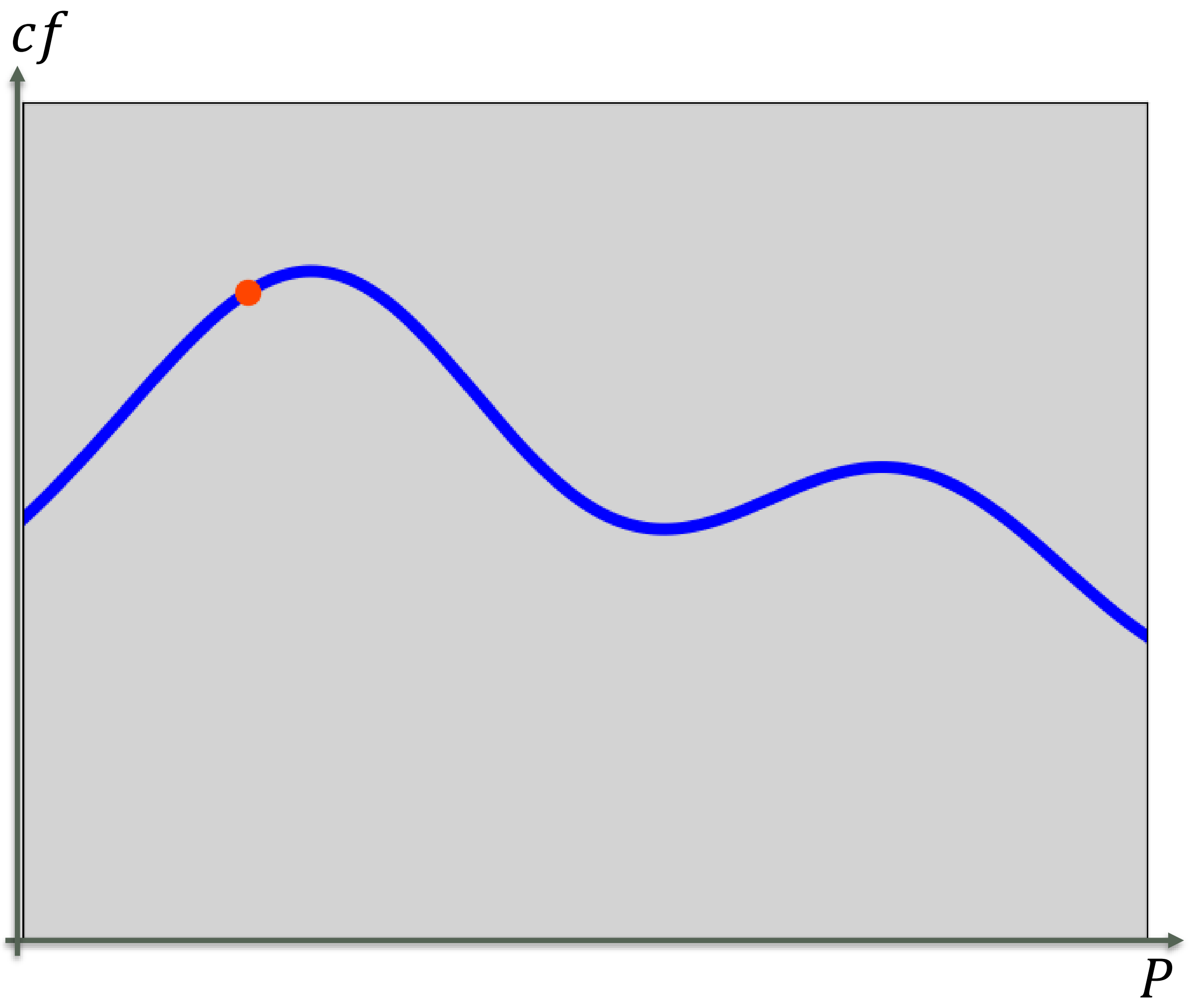
- **Goal: Approximate $cf(\vec{P})$ and find it's minimum value.**

Gaussian Processes



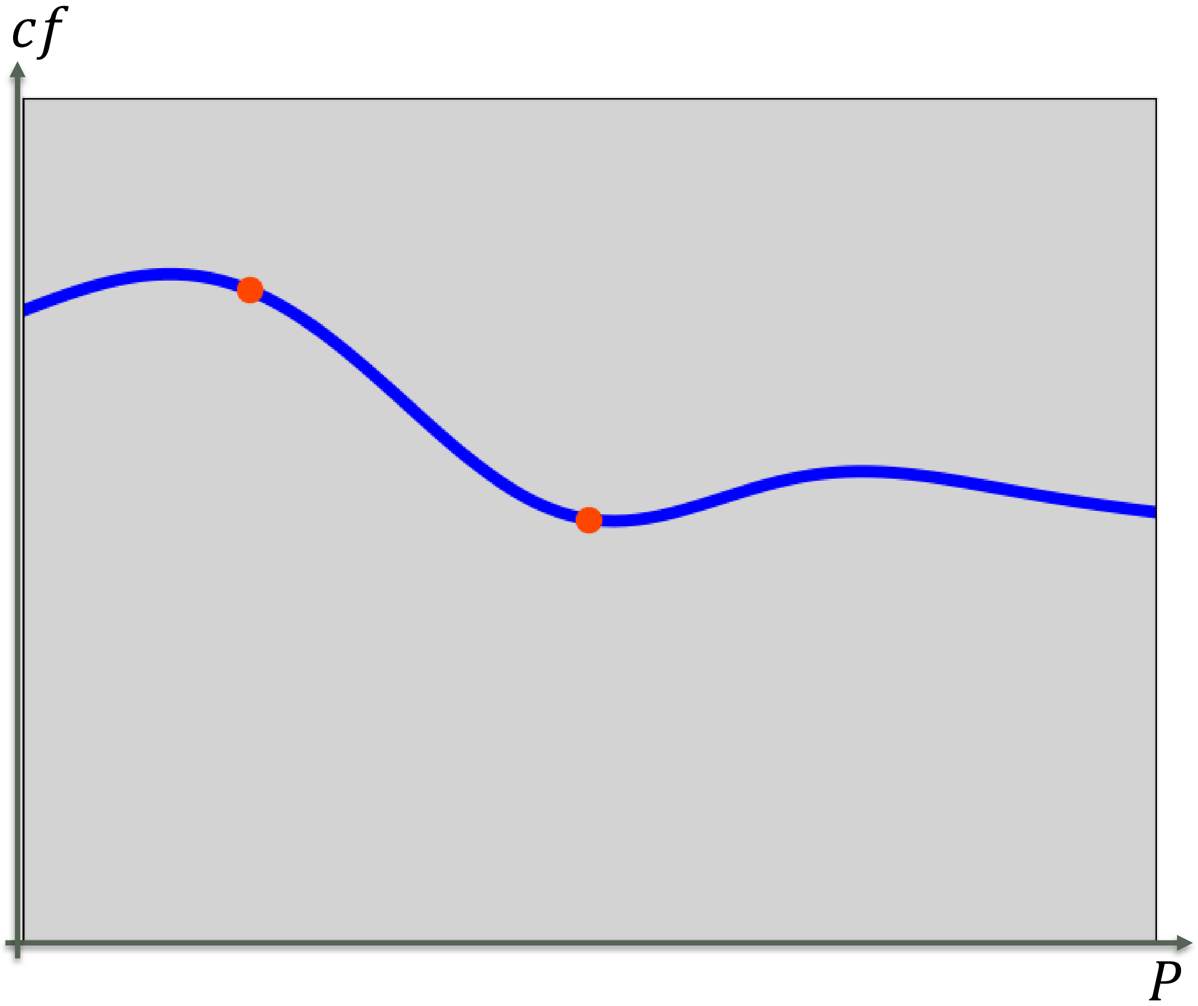
- No data available about $cf(P)$.
- Take a set out from the infinite space of functions (Gaussian process).

Gaussian Processes



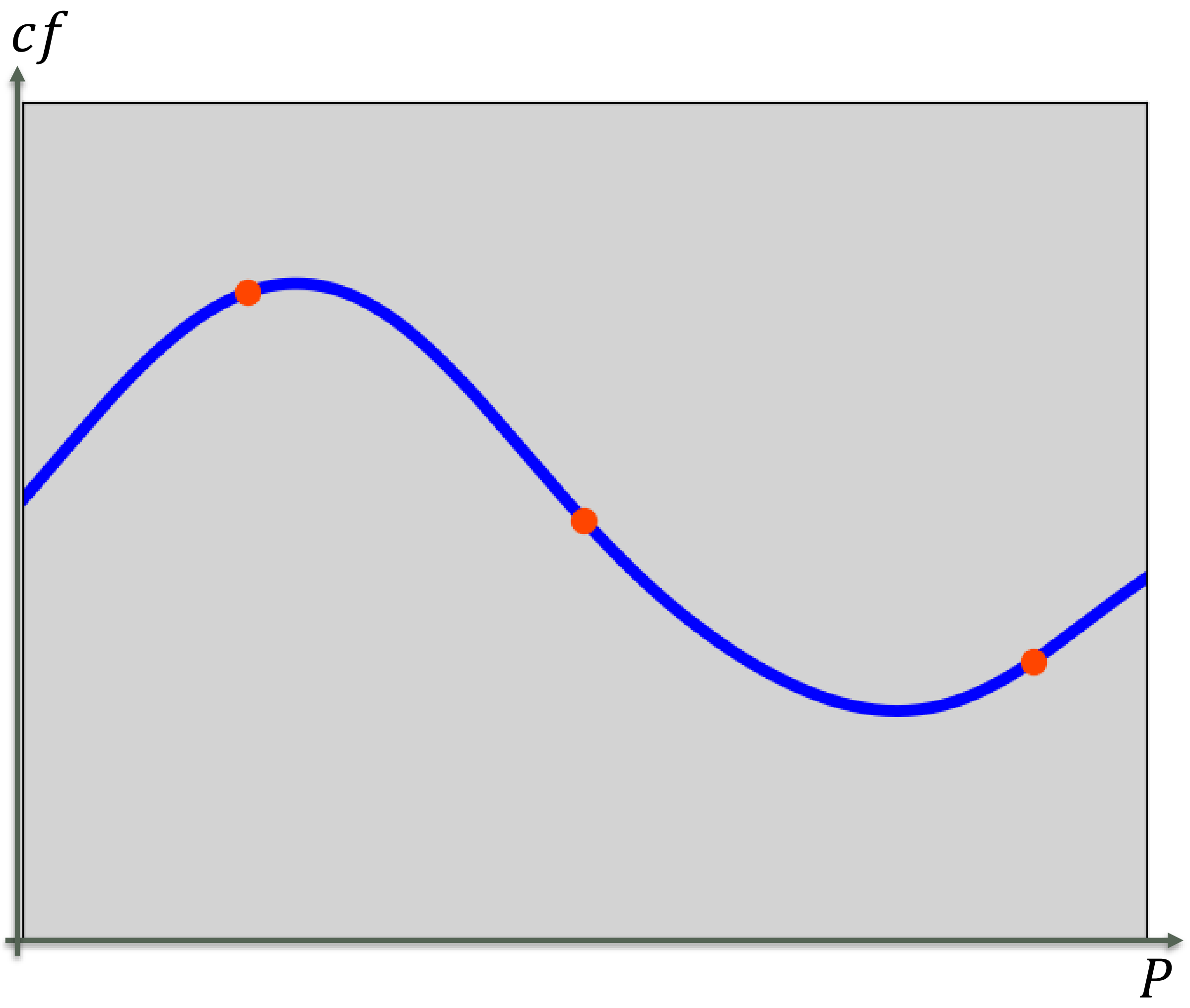
- No data available about $cf(P)$.
- Take a set out from the infinite space of functions (Gaussian process).
- A first measured value $cf(P_1)$ restricts the number of possible functions.

Gaussian Processes



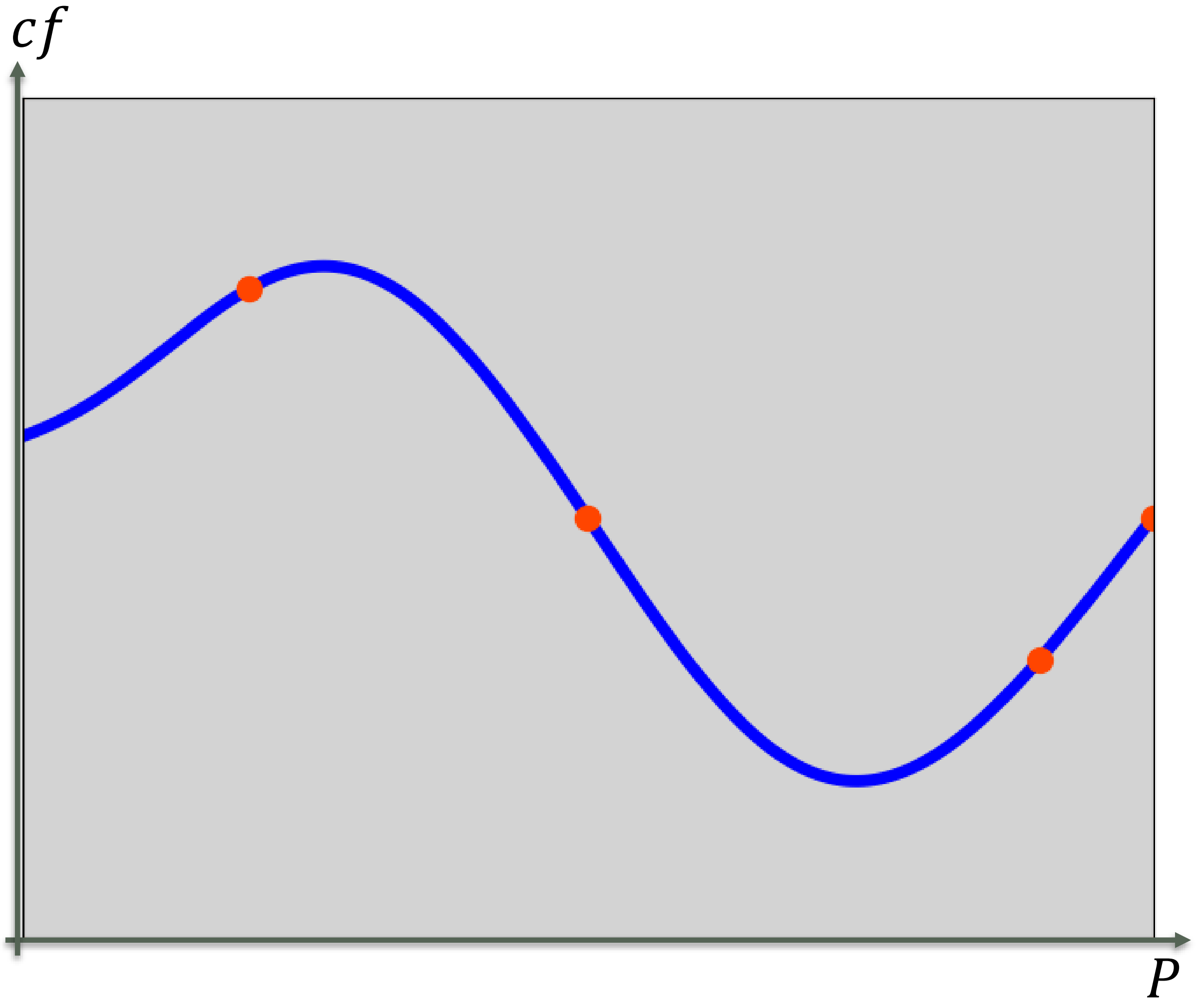
- No data available about $cf(P)$.
- Take a set out from the infinite space of functions (Gaussian process).
- A first measured value $cf(P_1)$ restricts the number of possible functions.
- Each additional point further limits these functions.

Gaussian Processes



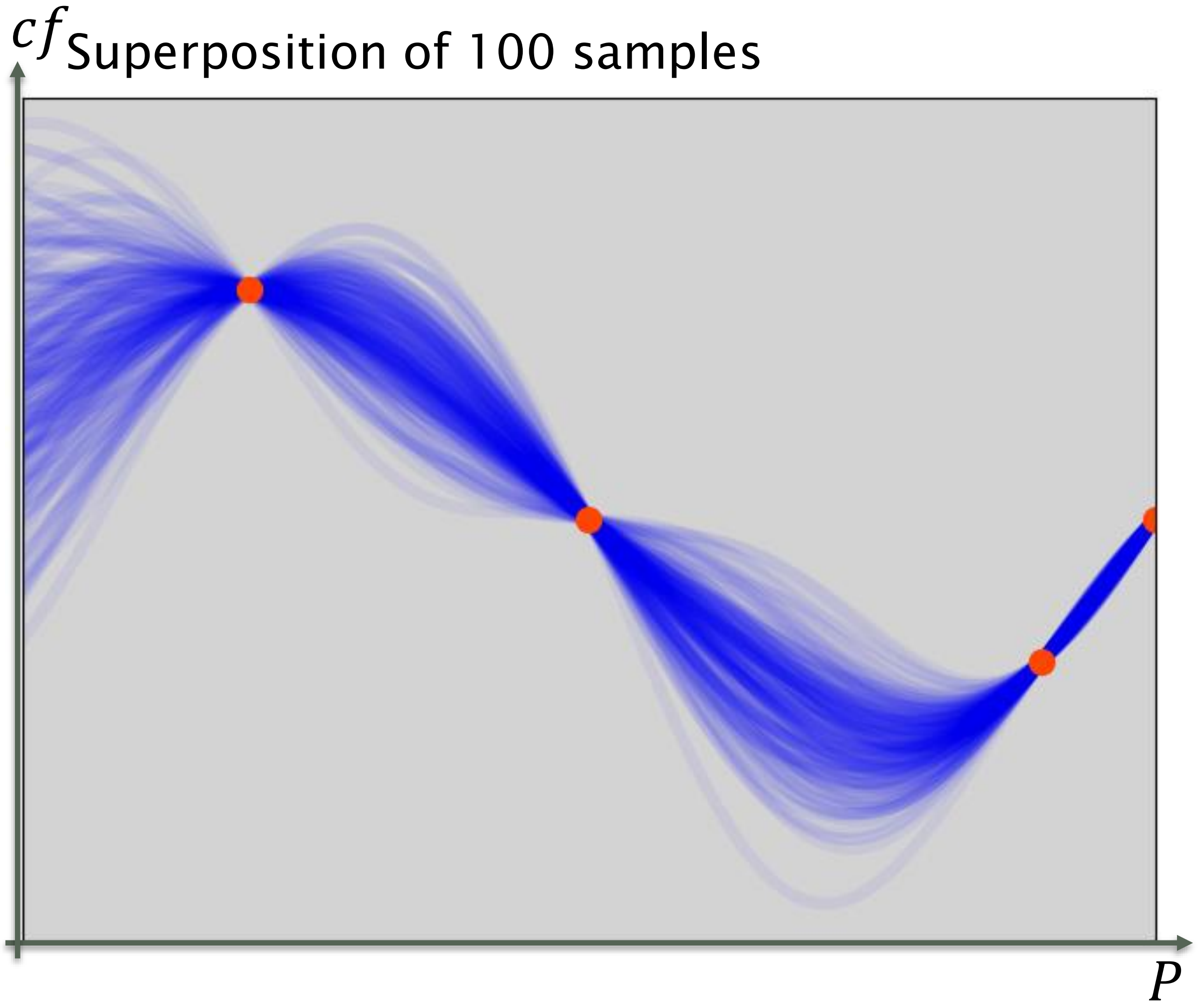
- No data available about $cf(P)$.
- Take a set out from the infinite space of functions (Gaussian process).
- A first measured value $cf(P_1)$ restricts the number of possible functions.
- Each additional point further limits these functions.

Gaussian Processes



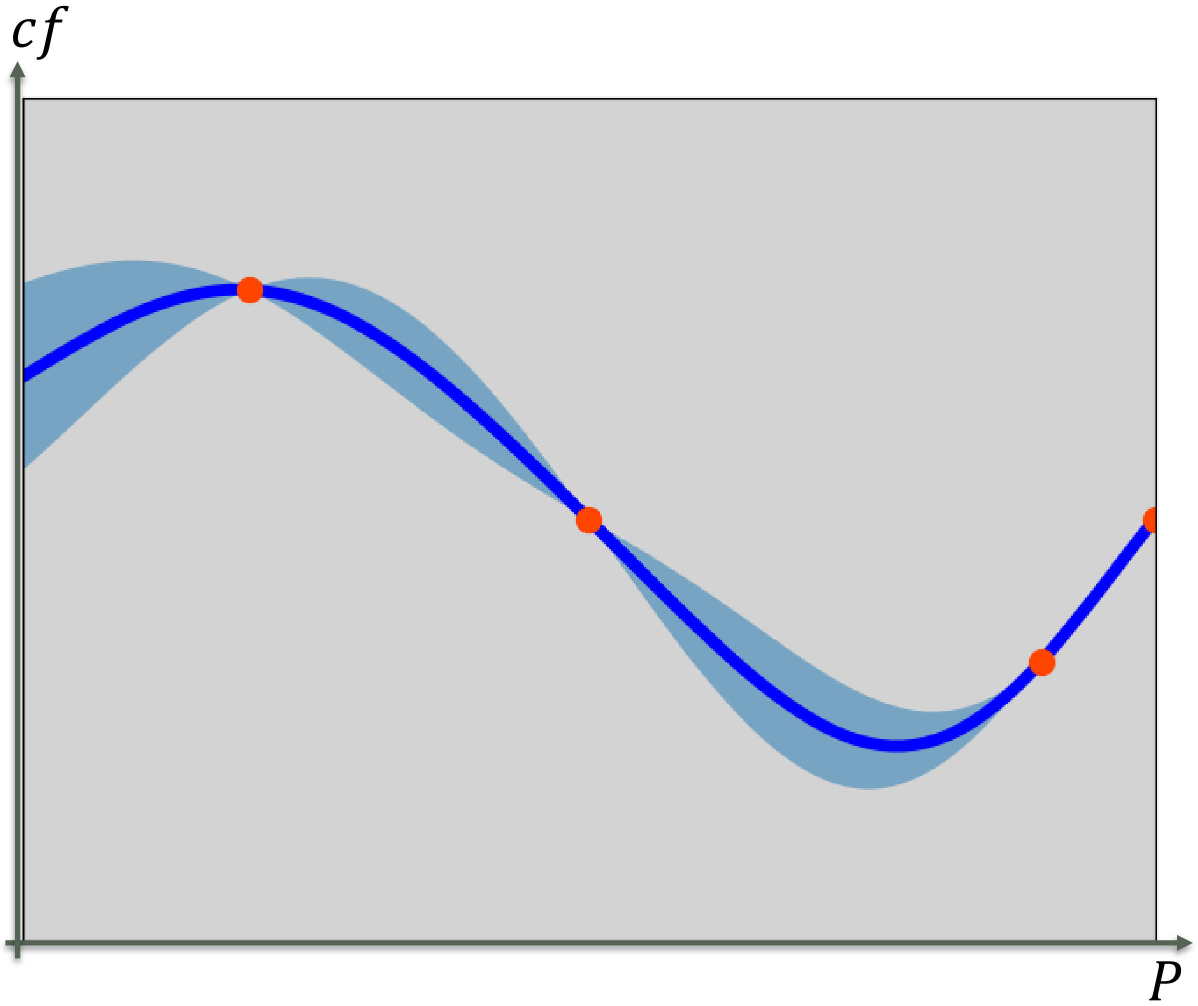
- No data available about $cf(P)$.
- Take a set out from the infinite space of functions (Gaussian process).
- A first measured value $cf(P_1)$ restricts the number of possible functions.
- Each additional point further limits these functions.

Gaussian Processes



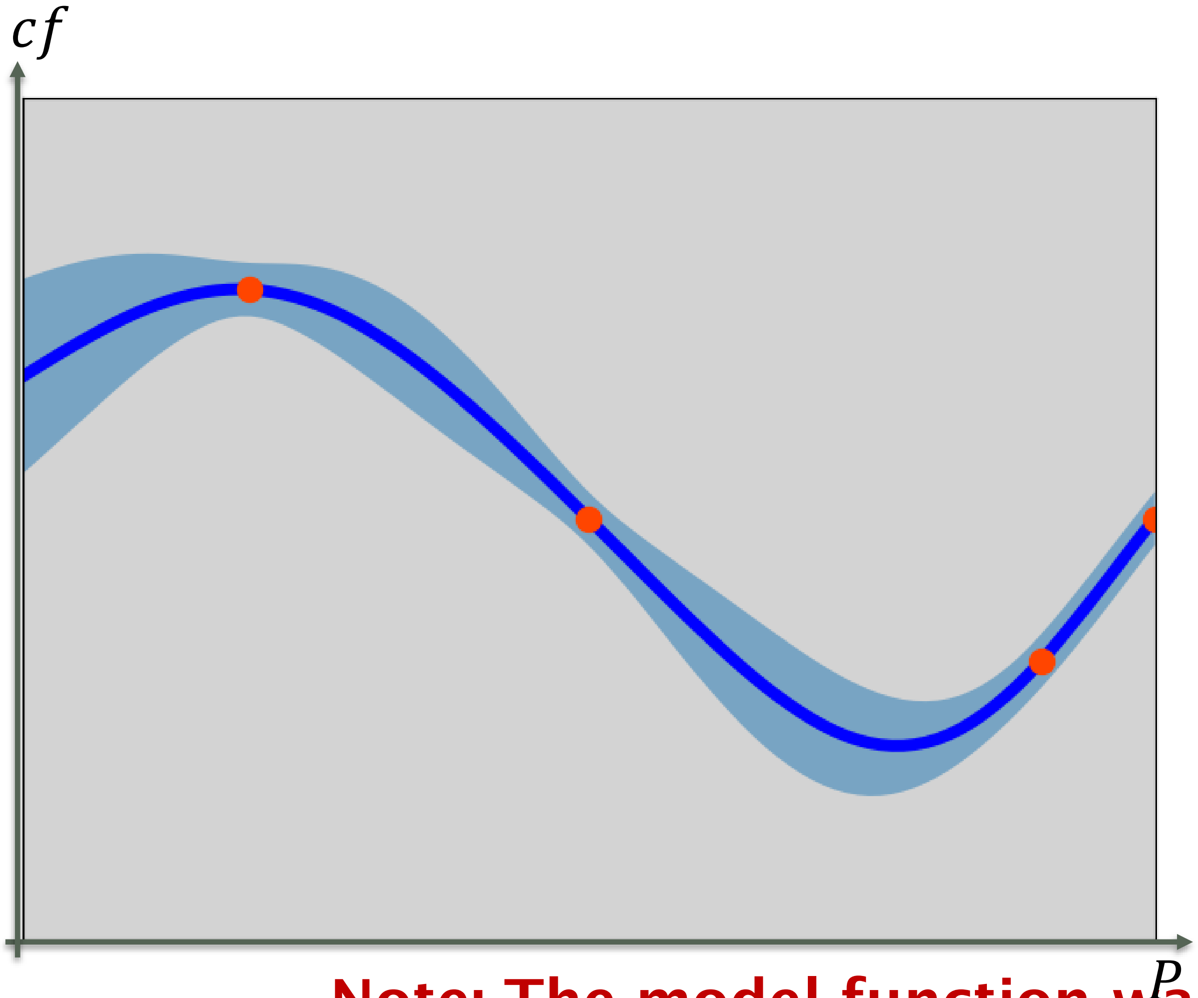
- No data available about $cf(P)$.
- Take a set out from the infinite space of functions (Gaussian process).
- A first measured value $cf(P_1)$ restricts the number of possible functions.
- Each additional point further limits these functions.
- Uncertainty for $cf(P)$ between the measured points.

Gaussian Processes



- No data available about $cf(P)$.
- Take a set out from the infinite space of functions (Gaussian process).
- A first measured value $cf(P_1)$ restricts the number of possible functions.
- Each additional point further limits these functions.
- Uncertainty for $cf(P)$ between the measured points.
- Describe with predicted mean and confidence region.

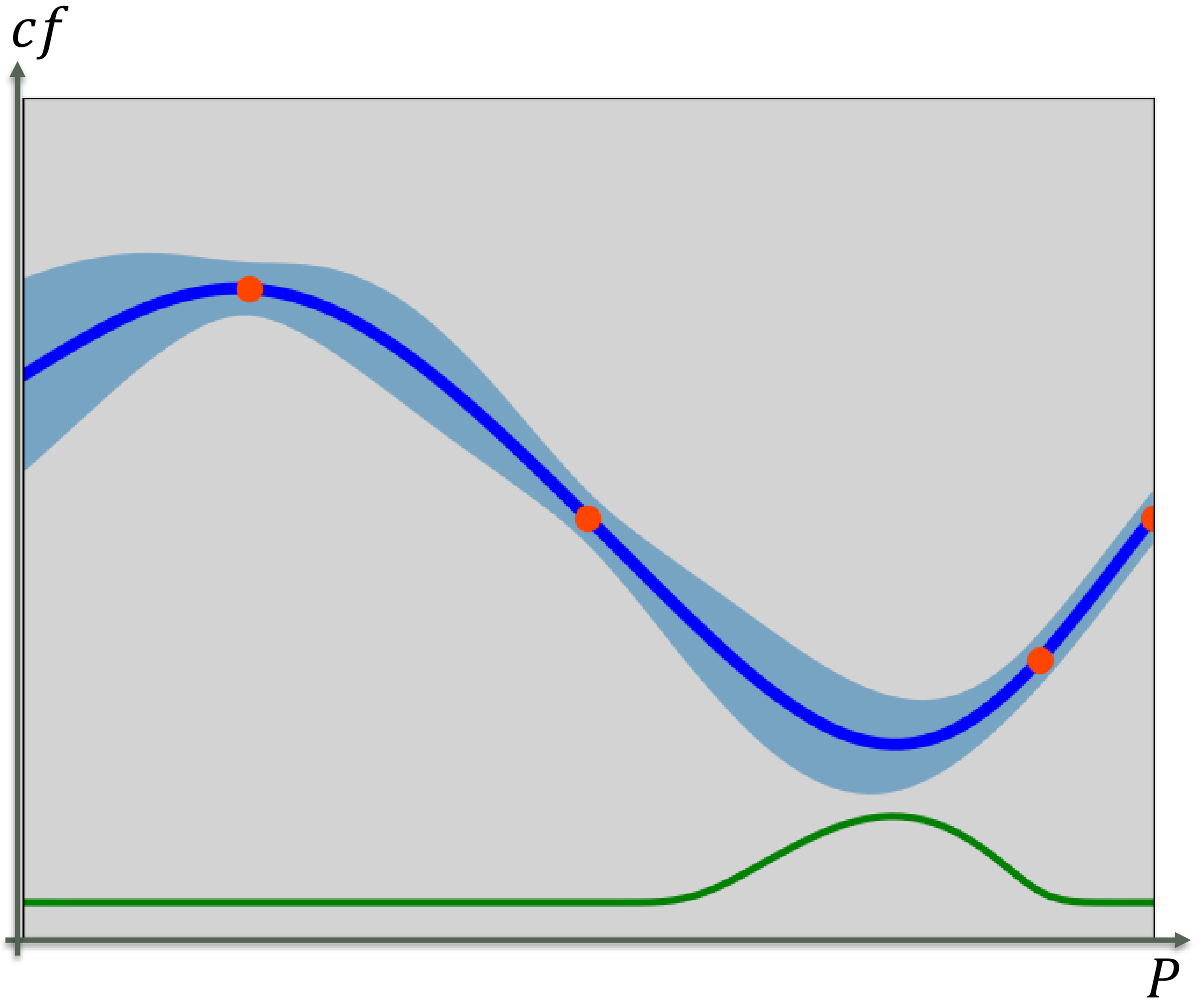
Gaussian Processes



- No data available about $cf(P)$.
- Take a set out from the infinite space of functions (Gaussian process).
- A first measured value $cf(P_1)$ restricts the number of possible functions.
- Each additional point further limits these functions.
- Uncertainty for $cf(P)$ between the measured points.
- Describe with predicted mean and confidence region.
- Uncertainty of measured values.

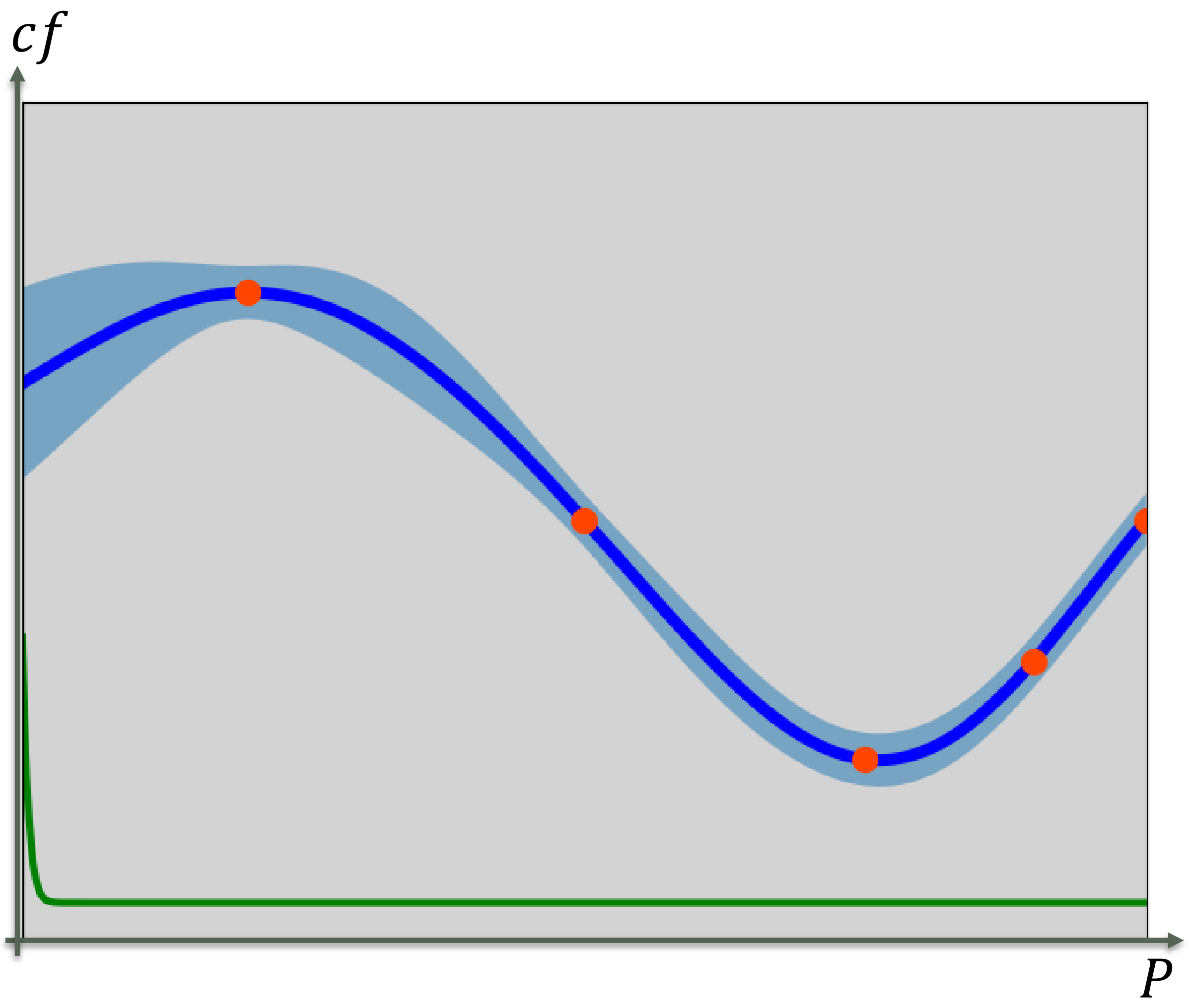
Note: The model function was self-learned, it is not a polynomial or spline or something like that!

Bayesian Optimization



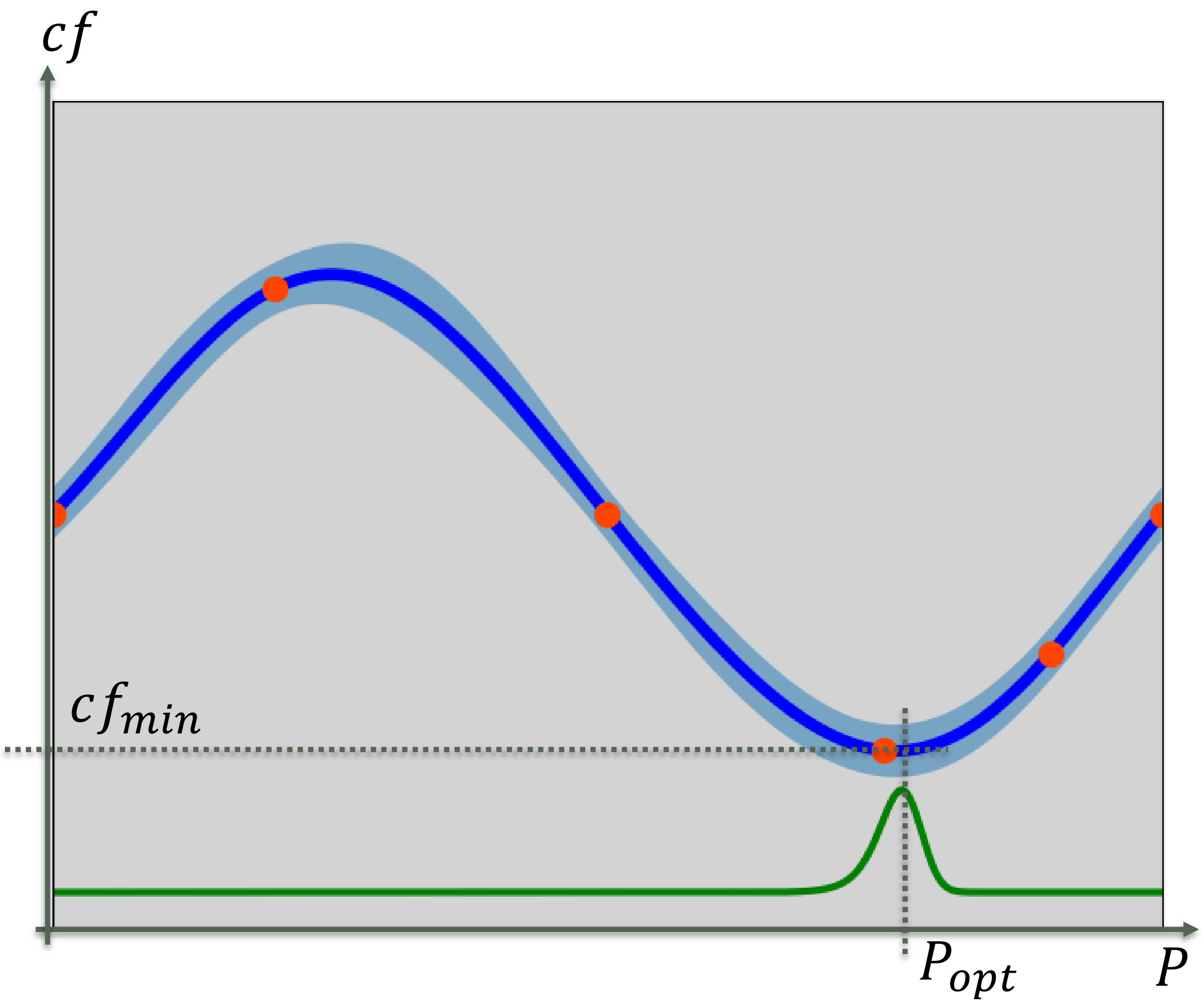
- The acquisition function determines the P with most expected information.

Bayesian Optimization



- The acquisition function determines the P with most expected information.

Bayesian Optimization



- The acquisition function determines the P with most expected information.
- After e few steps $cf(P)$ can be approximated in demanded precision to estimate
  - $cf_{min}(P)$
  - P_{opt}
- This method can be extended to multi-dimensional parameter space $\vec{P}$

- Example Ge

Experimental Procedure

Fixed Parameters:

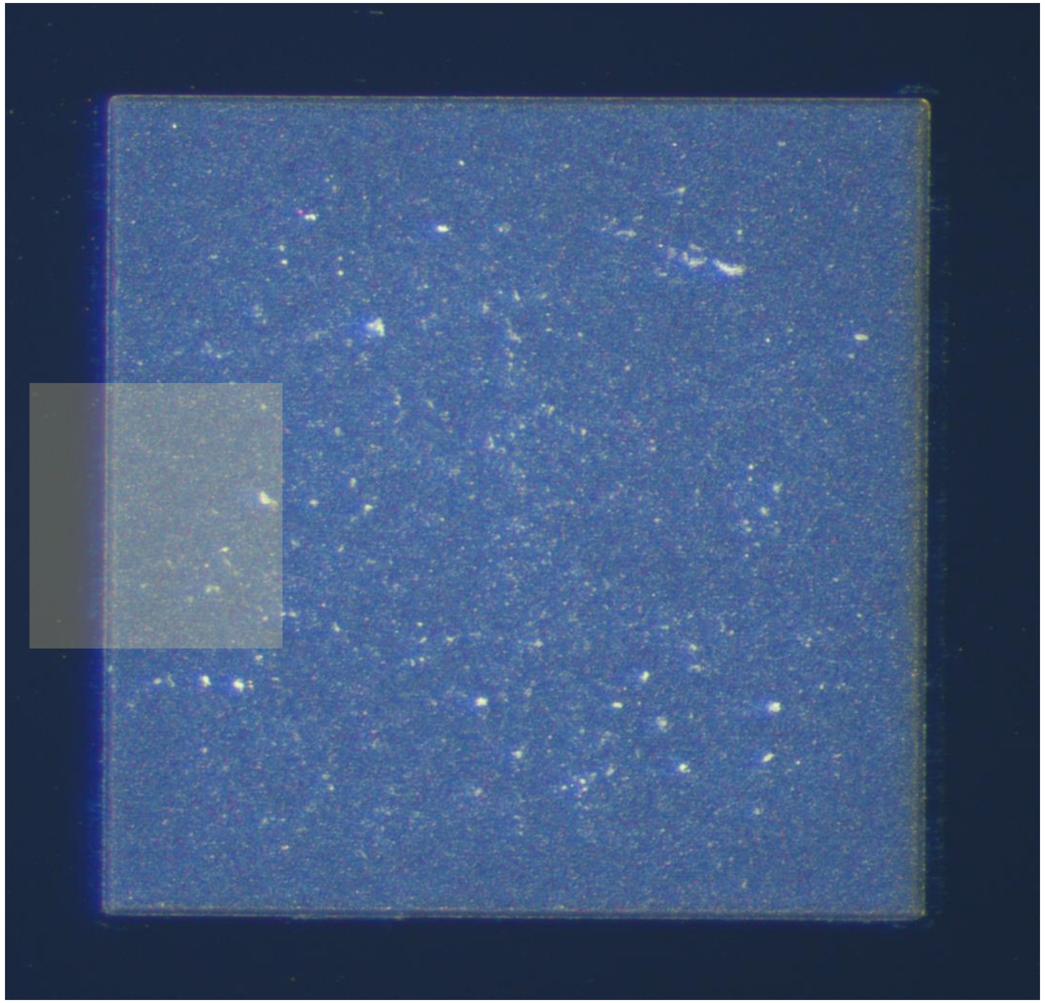
- $f_{rep} = 200\ kHz$
- $w_0 = 14\ \mu m$
- $p_x = p_y = 7\ \mu m$
- $\Delta\tau = 10\ ps$
- $\lambda = 1064\ nm$
- $N_{rough,smooth} = 1$
- No autotracking of the focal position

Varied Parameters:

- $n_{burst,rough} = 1, 2, \dots 8$
- $0.17 \frac{J}{cm^2} \leq \phi_{0,rough} \leq 6 \frac{J}{cm^2}$
- $0.17 \frac{J}{cm^2} \leq \phi_{0,smooth} \leq 6 \frac{J}{cm^2}$
- $n_{layer,rough} = 1, 2, \dots 10$
- $n_{layer,smooth} = 5, 6, \dots 25$

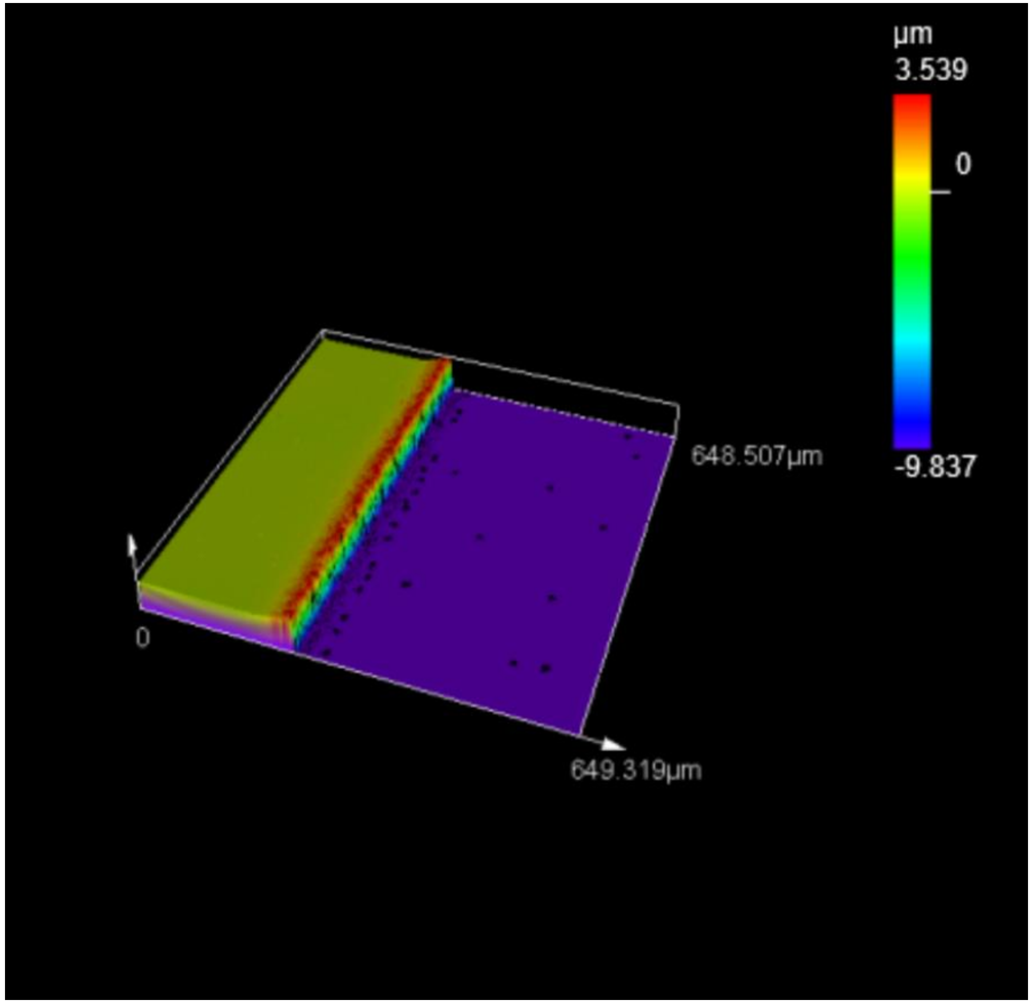
- Start with arbitrary set of parameters.
- Calculate cf and next set of parameters by Bayesian optimization
- Stop after 40 experiments.

Experimental Procedure



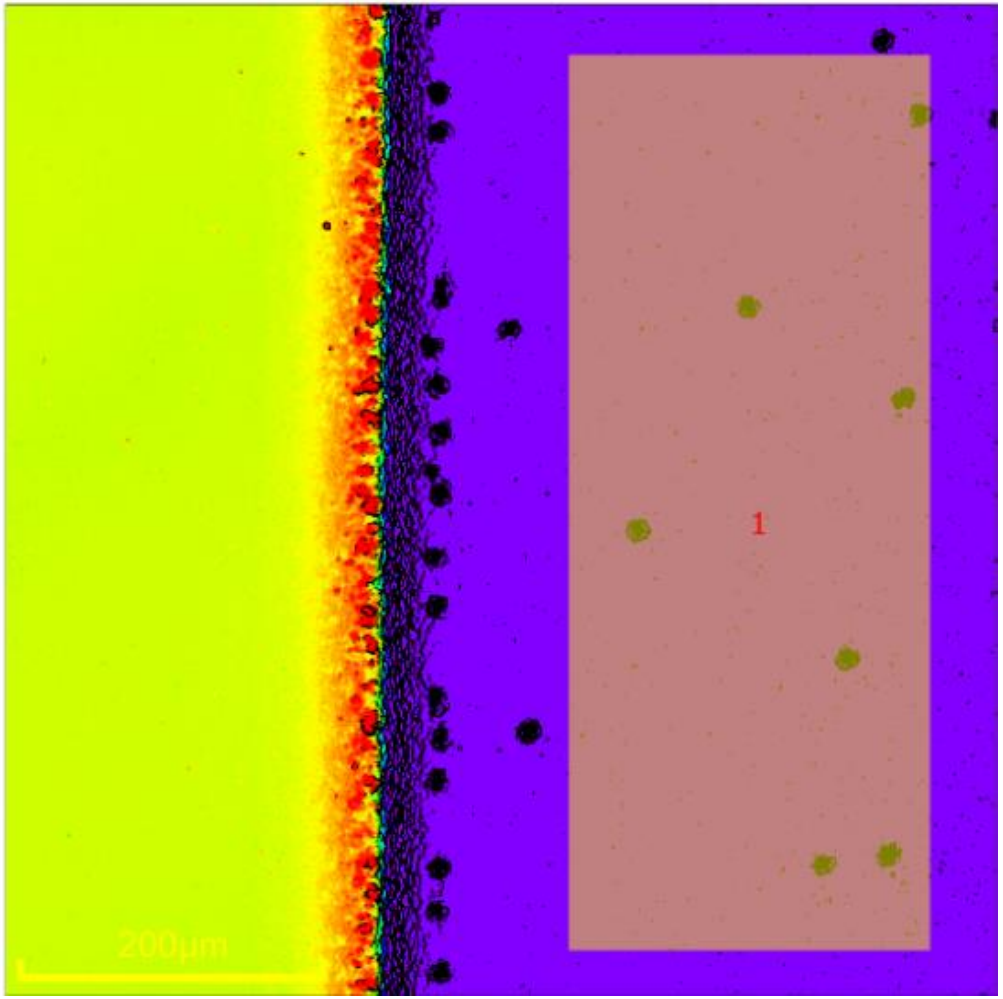
- Machine a square with a side length of $s = 1.0\ mm$ with a set of parameter for roughing and smoothening.
- Measure surface topography with a laser scanning microscope.

Experimental Procedure



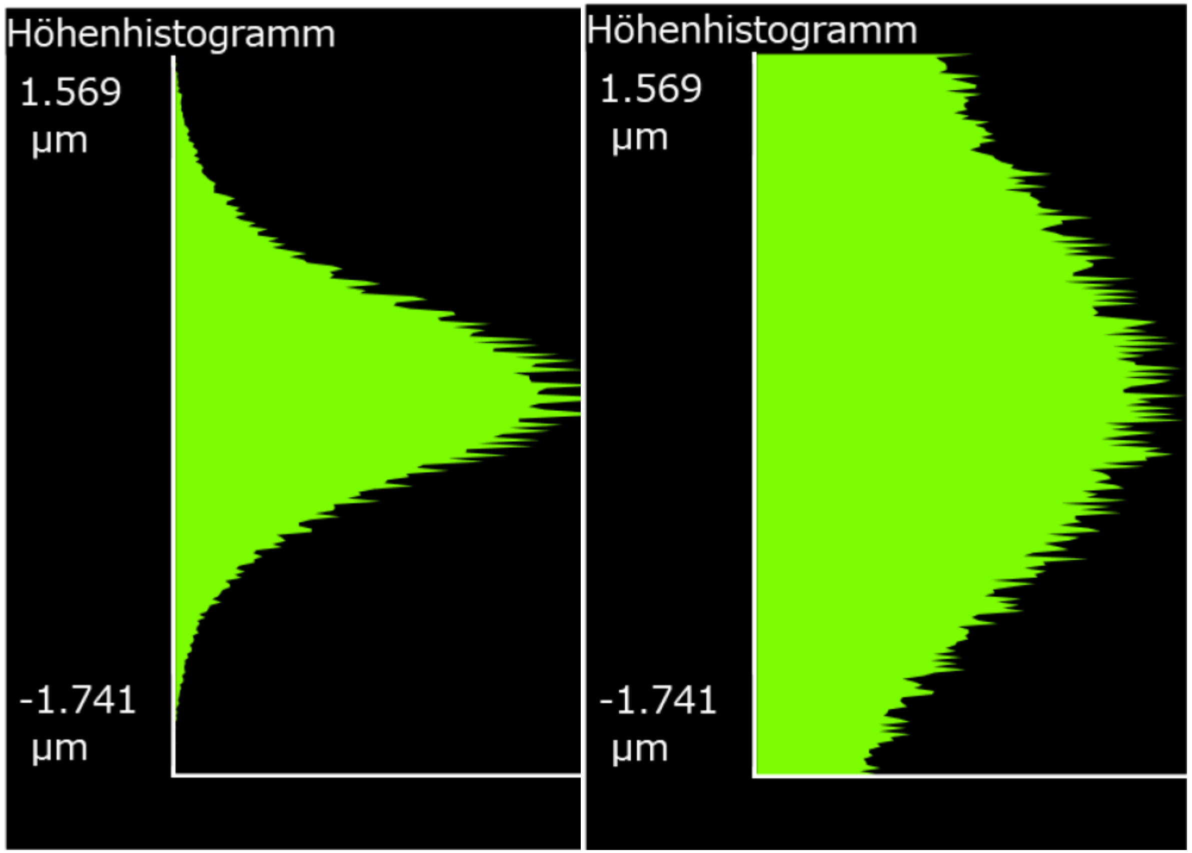
- Machine a square with a side length of $s = 1.0\ mm$ with a set of parameter for roughing and smoothening.
- Measure surface topography with a laser scanning microscope.
- Deduce the step height and with this the average removal depth t per layer.

Experimental Procedure



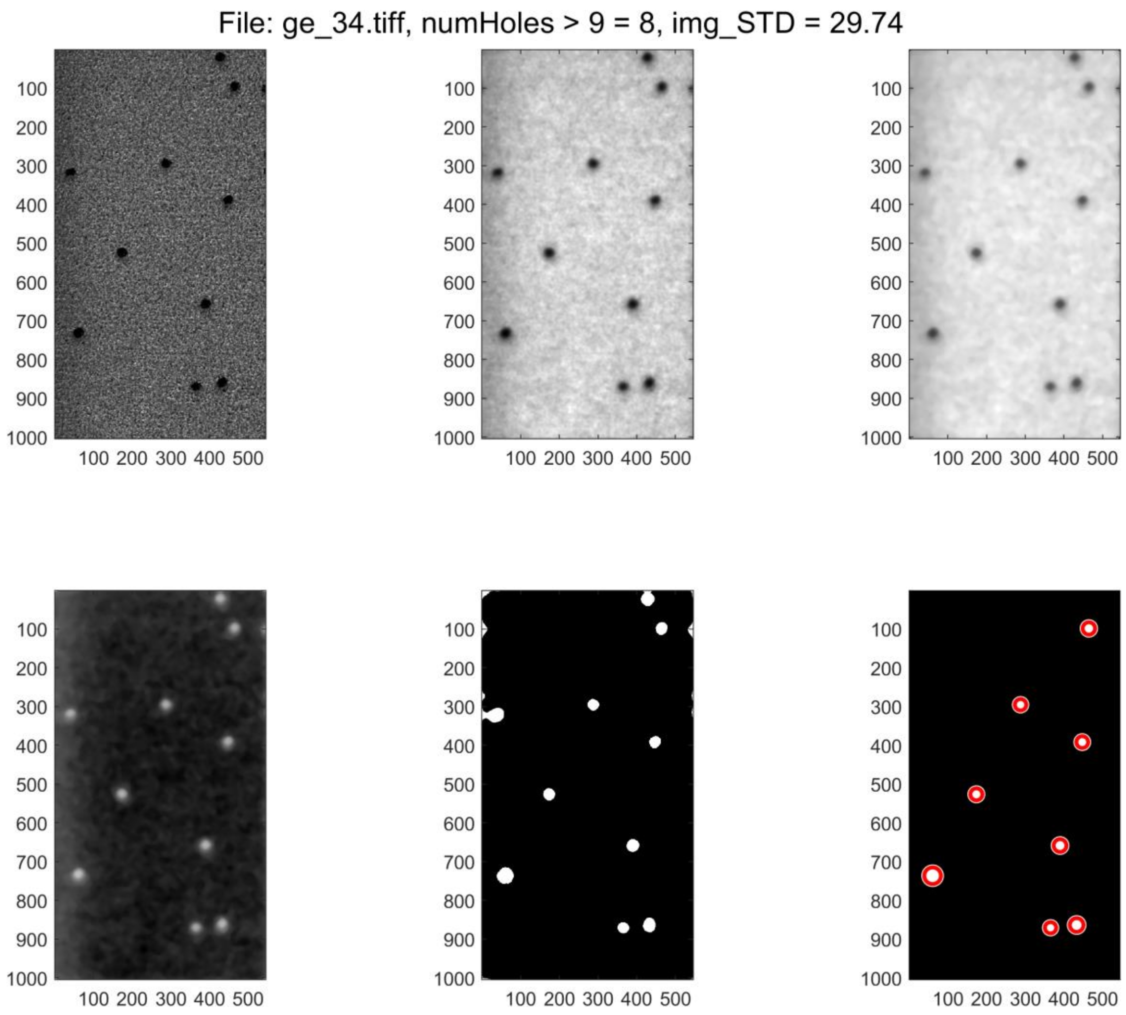
- Machine a square with a side length of $s = 1.0\ mm$ with a set of parameter for roughing and smoothening.
- Measure surface topography with a laser scanning microscope.
- Deduce the step height and with this the average removal depth t per layer.
- Deduce the value of the surface roughness s_a in a selected machined area.

Experimental Procedure



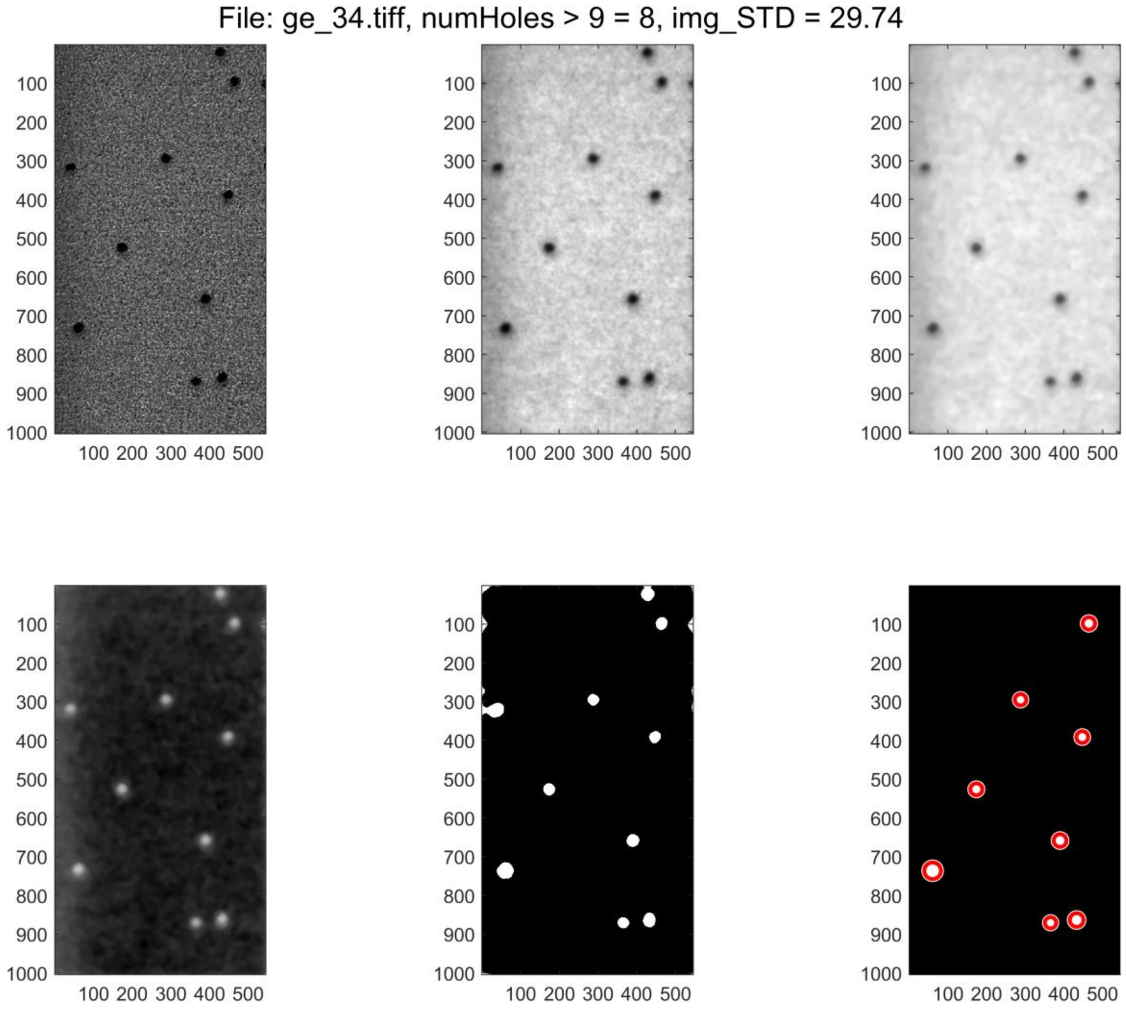
- Machine a square with a side length of $\ell = 1.0\,mm$ with a set of parameter for roughing and smoothening.
- Measure surface topography with a laser scanning microscope.
- Deduce the step height and with this the average removal depth t per layer.
- Deduce the value of the surface roughness s_a in a selected machined area.
- Calculate in this selected area the standard deviation σ of the measured heights.

Experimental Procedure



- Machine a square with a side length of $\ell = 1.0\ mm$ with a set of parameter for roughing and smoothening.
- Measure surface topography with a laser scanning microscope.
- Deduce the step height and with this the average removal depth t per layer.
- Deduce the value of the surface roughness s_a in a selected machined area.
- Calculate in this selected area the standard deviation σ of the measured heights.
- Deduce the number of holes N in the selected area with image processing.

Experimental Procedure

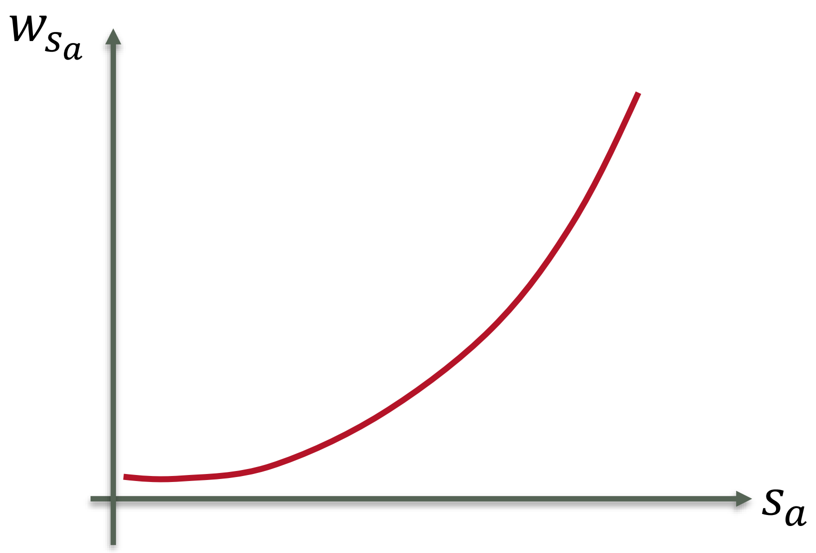


- Machine a square with a side length of $\ell = 1.0\,mm$ with a set of parameter for roughing and smoothening.
- Measure surface topography with a laser scanning microscope.
- Deduce the step height and with this the average removal depth $\boldsymbol{t}$ per layer.
- Deduce the value of the surface roughness $\boldsymbol{s_a}$ in a selected machined area.
- Calculate in this selected area the standard deviation $\boldsymbol{\sigma}$ of the measured heights.
- Deduce the number of holes $\boldsymbol{N}$ in the selected area with image processing.

Experimental Procedure

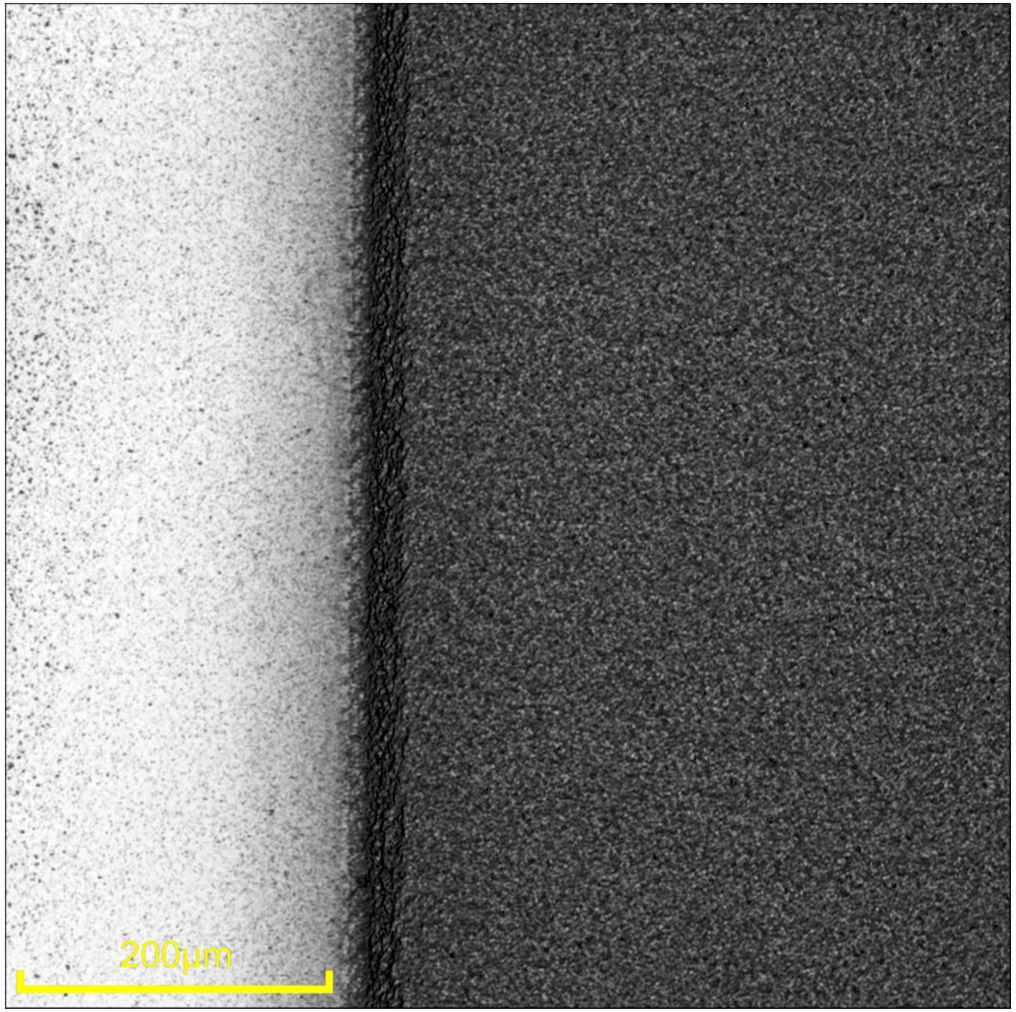
- Cost function:

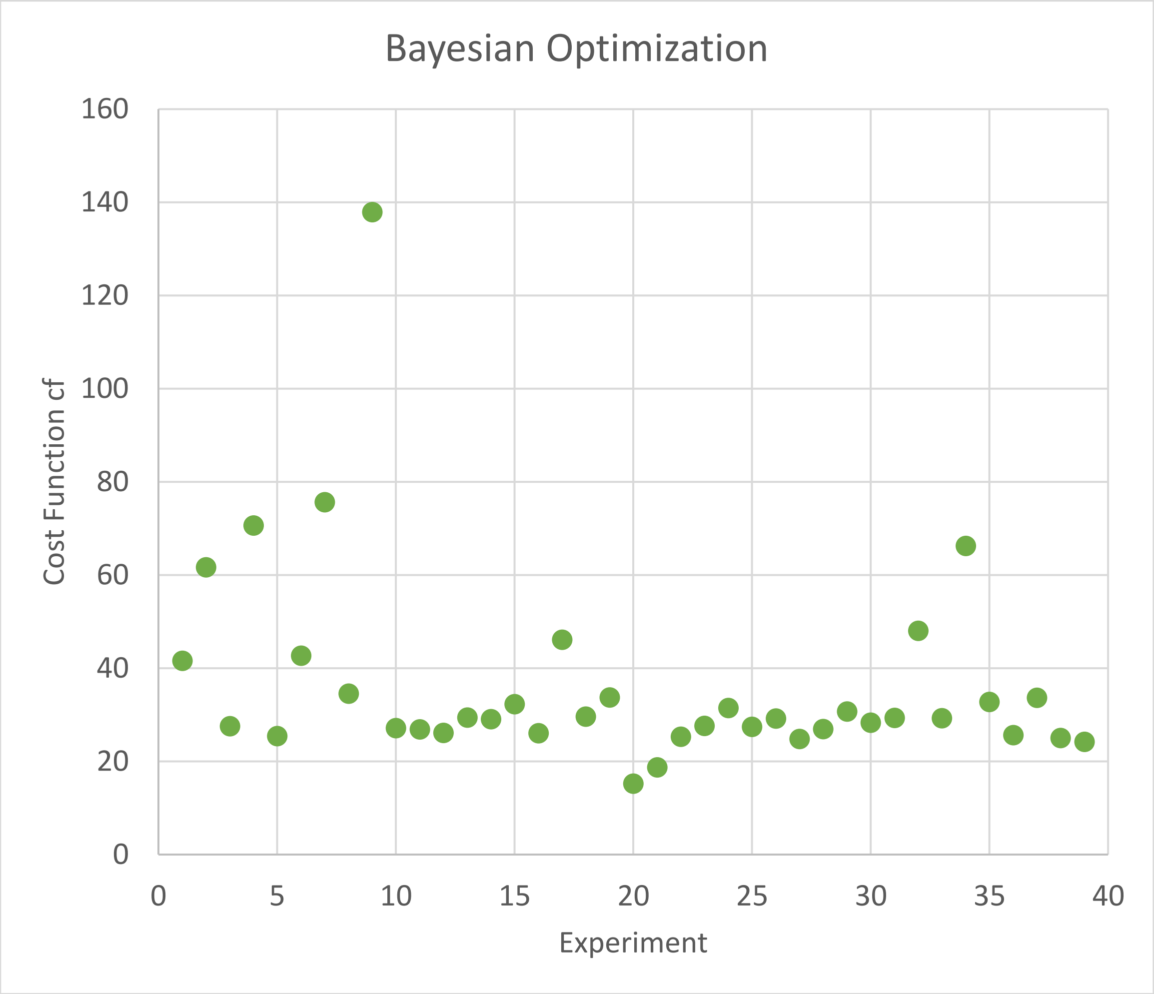
$$cf = w_{s_a}(s_a) \cdot s_a - w_t \cdot t + w_\sigma \cdot \sigma + w_N \cdot N$$



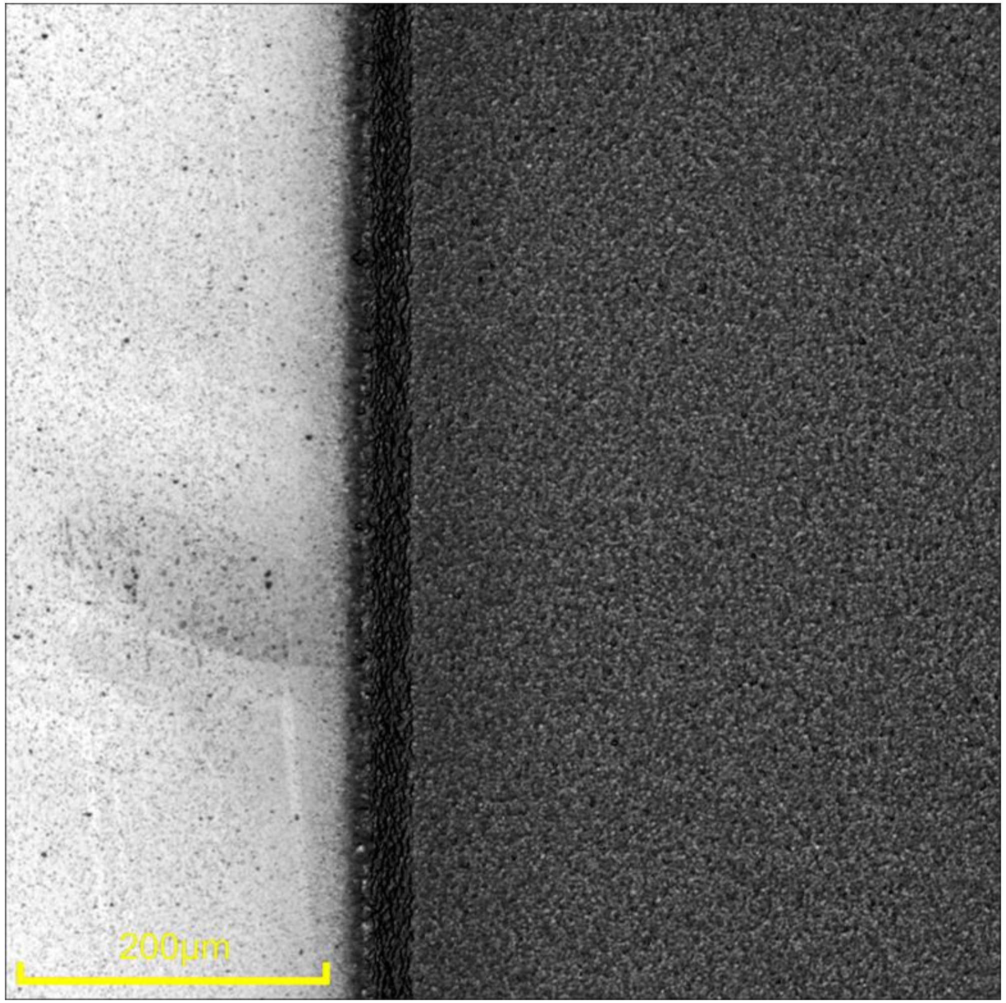
- Machine a square with a side length of $\ell = 1.0\,mm$ with a set of parameter for roughing and smoothening.
- Measure surface topography with a laser scanning microscope.
- Deduce the step height and with this the average removal depth $\boldsymbol{t}$ per layer.
- Deduce the value of the surface roughness $\boldsymbol{s_a}$ in a selected machined area.
- Calculate in this selected area the standard deviation $\boldsymbol{\sigma}$ of the measured heights.
- Deduce the number of holes $\boldsymbol{N}$ in the selected area with image processing.
- Calculate the value of the cost function

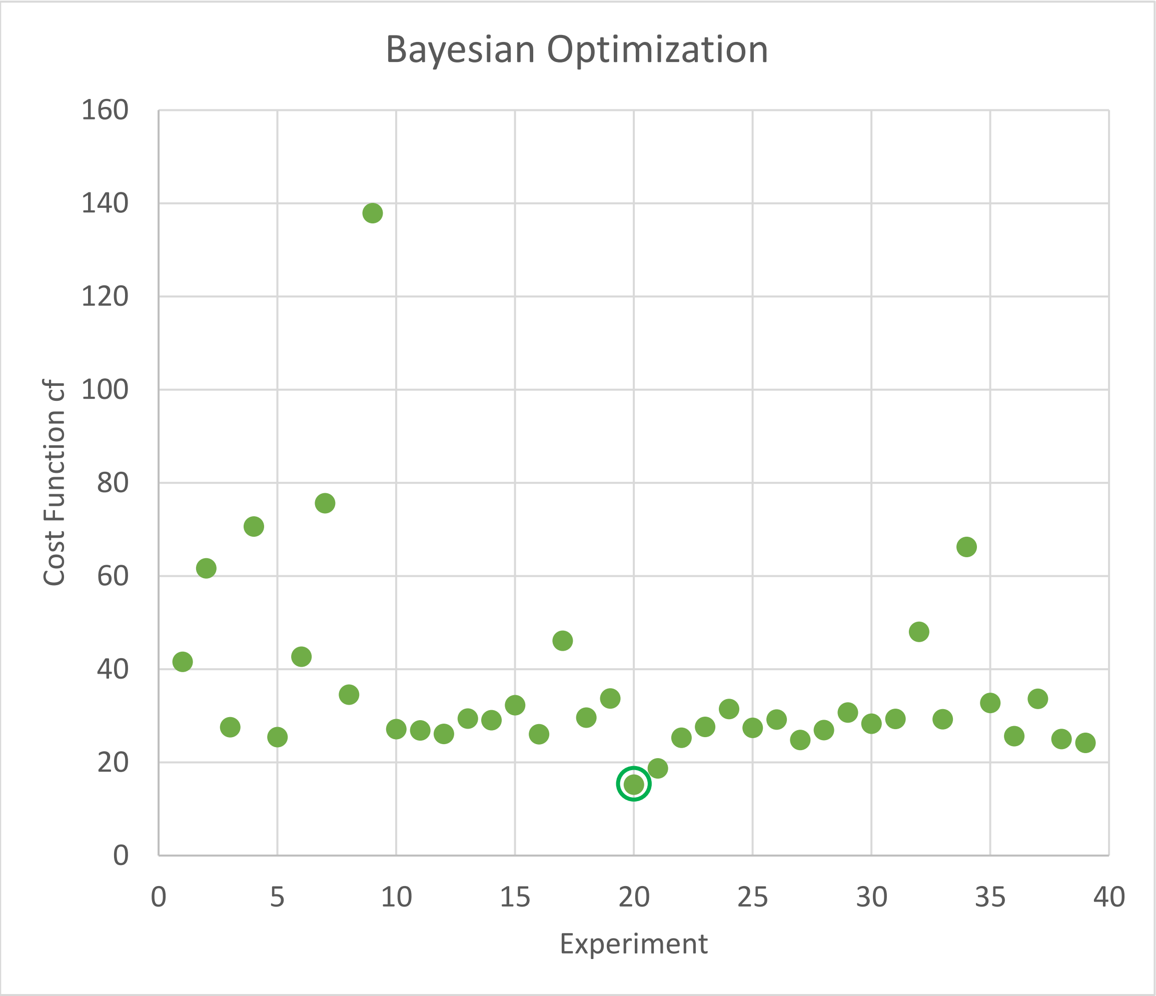
Ge: Bayesian Optimization



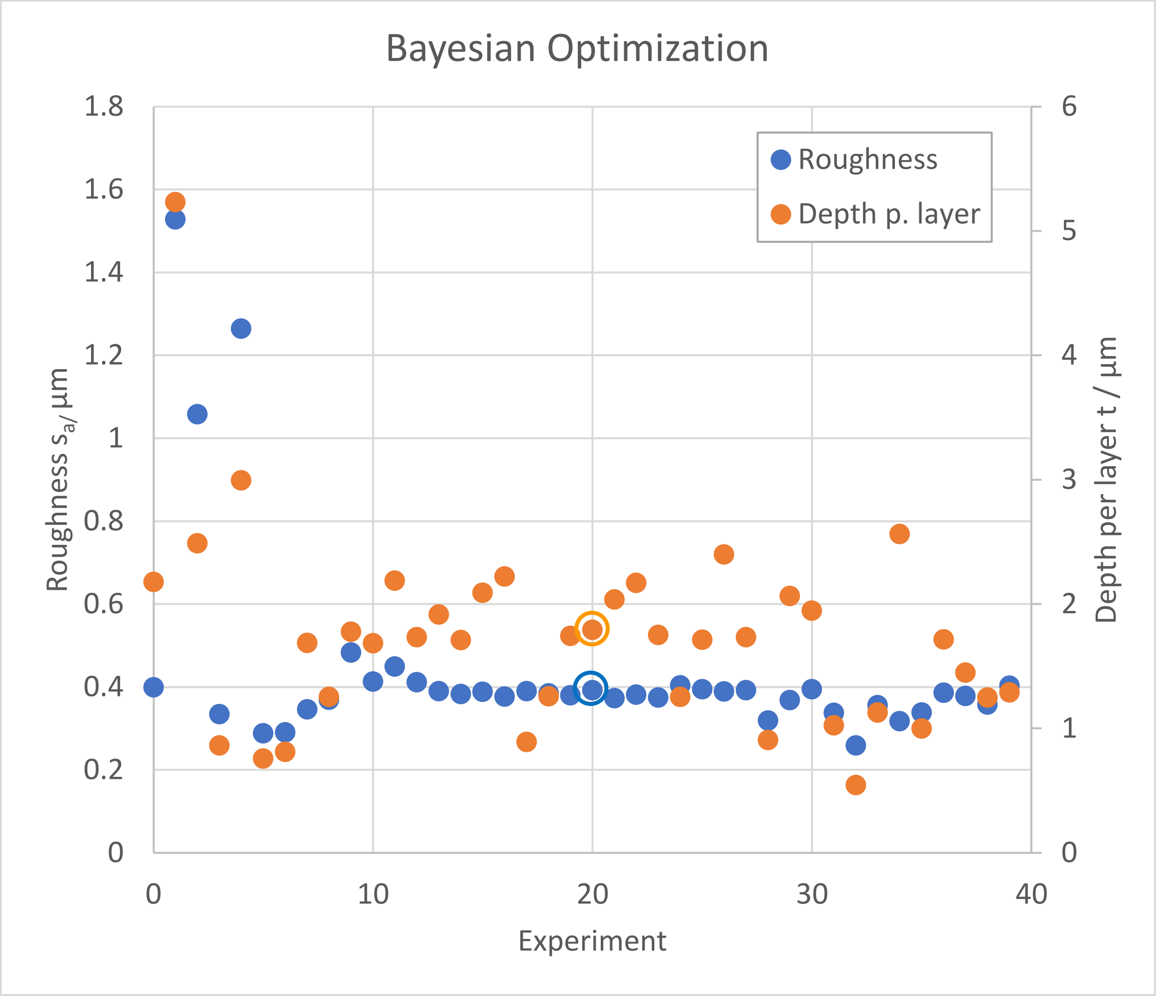


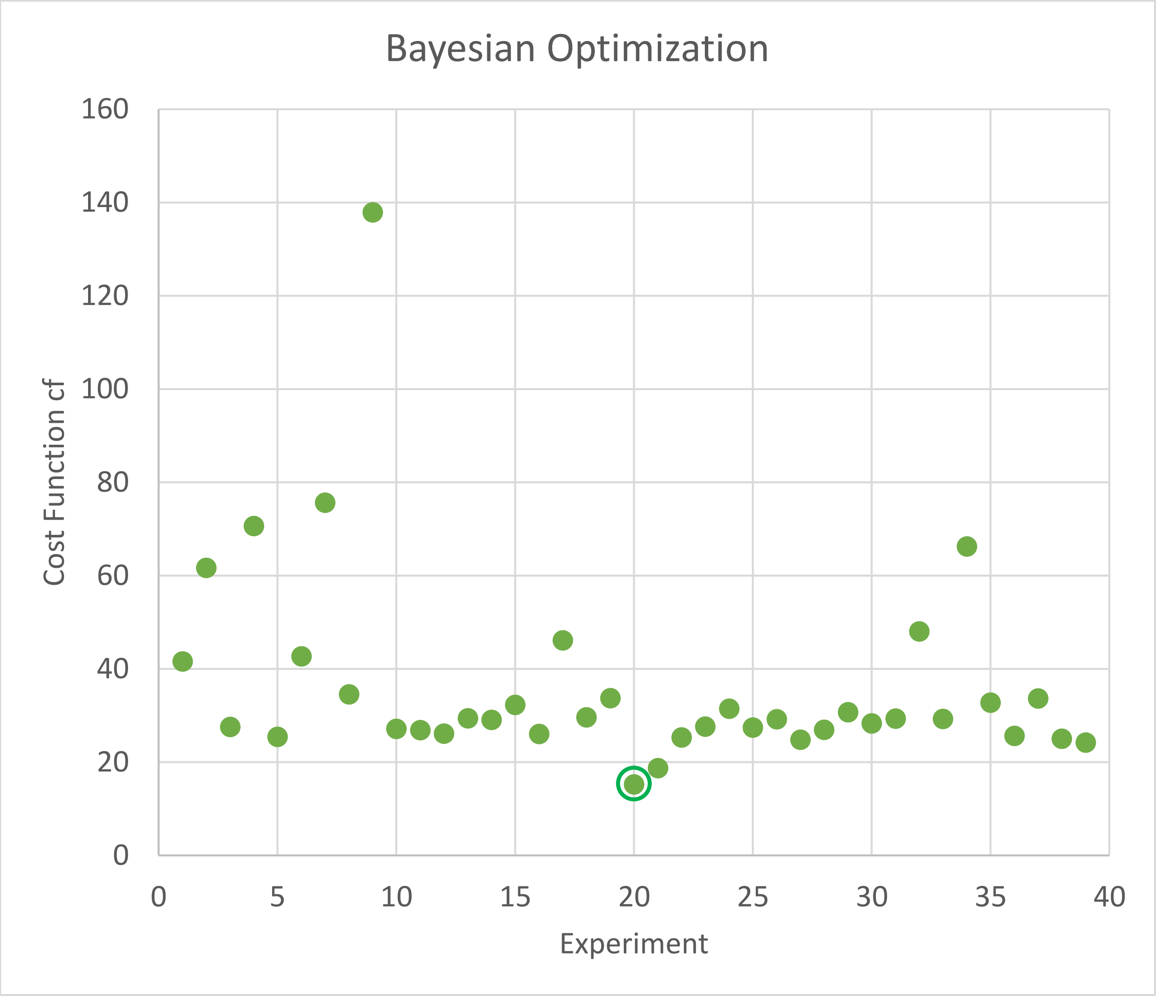
Ge: Bayesian Optimization



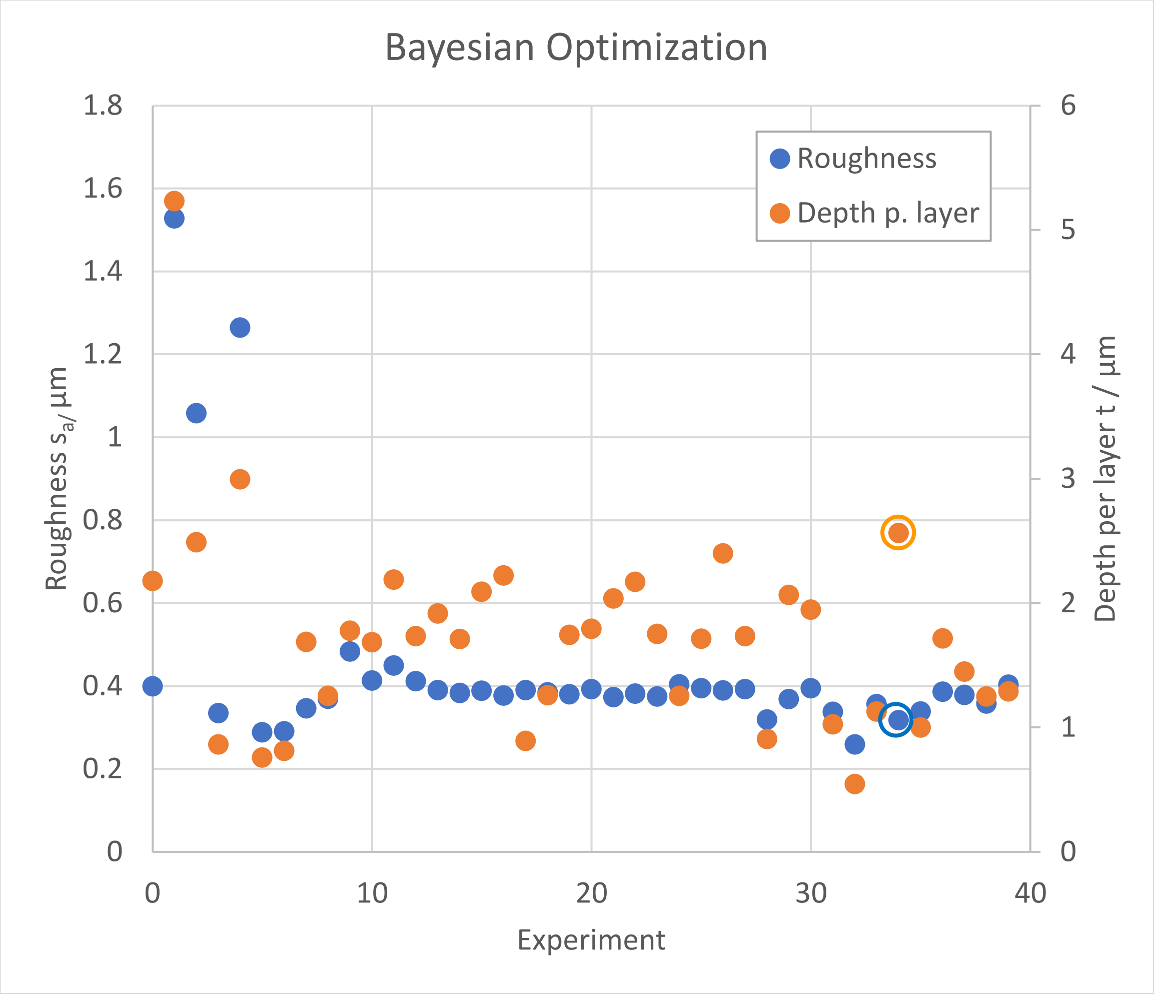


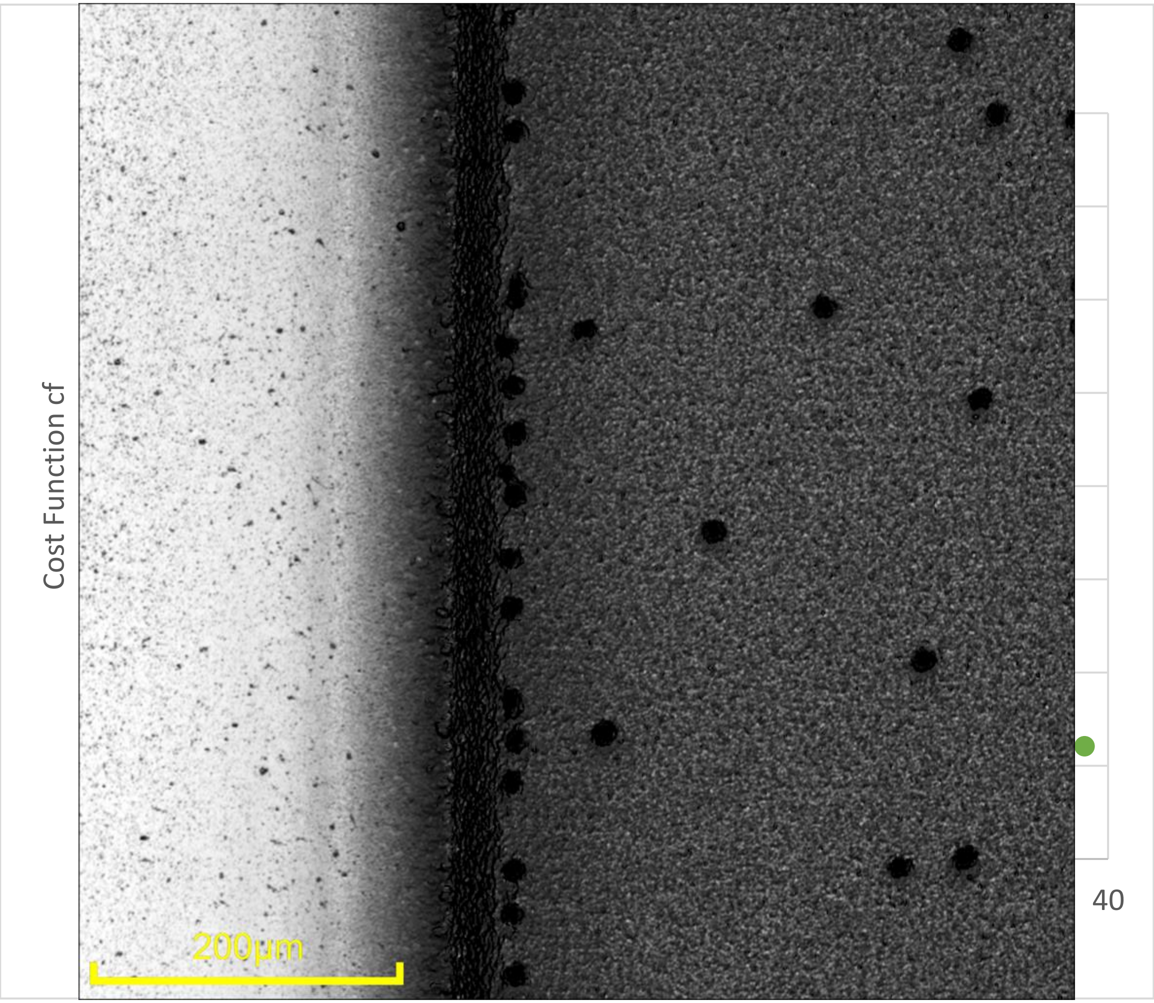
Ge: Bayesian Optimization





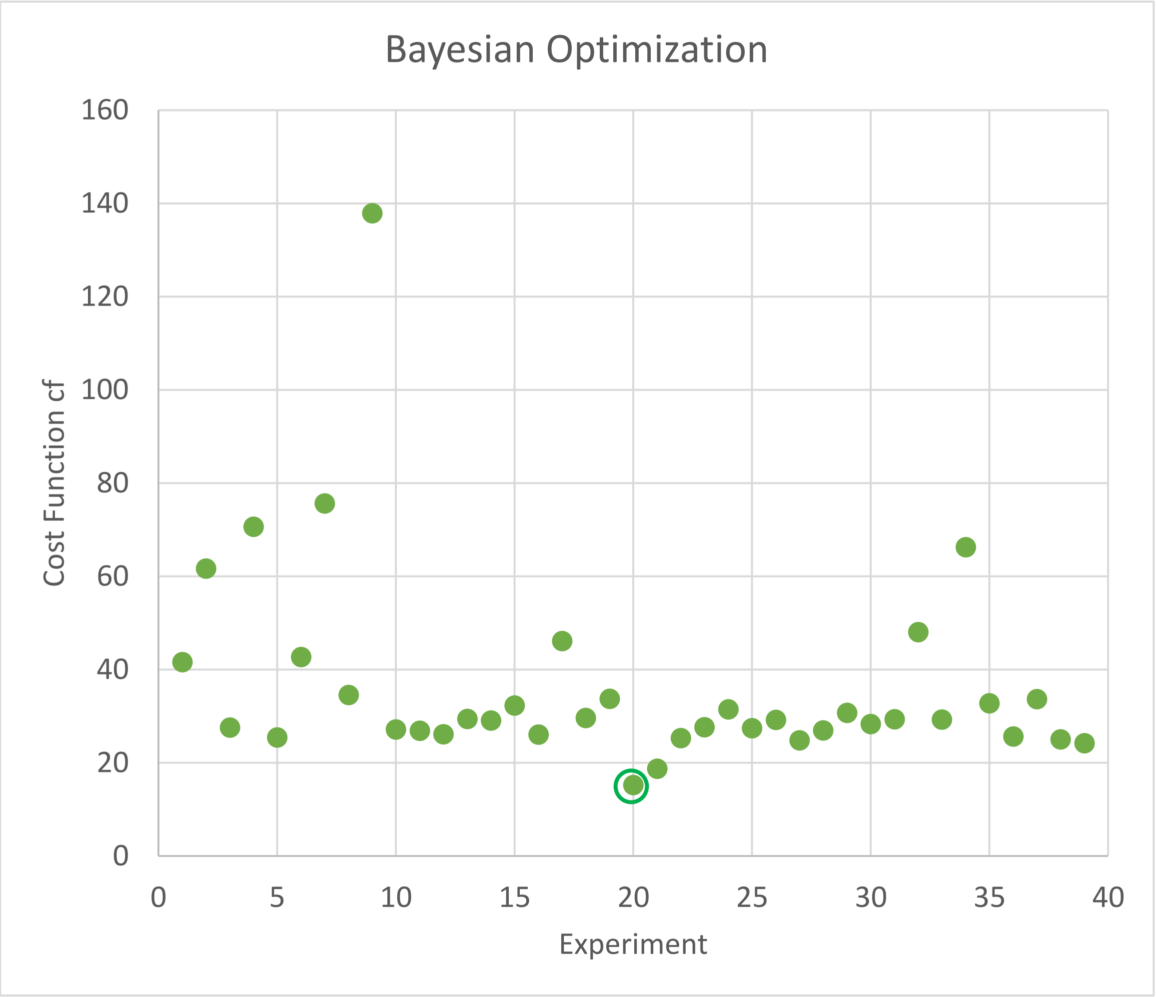
Ge: Bayesian Optimization





Ge: Bayesian Optimization

- Lowest value after 20 experiments.
- Additional 20 experiments no further improvement.
- Best parameters:
  - $n_{burst,rough} = 3$
  - $\phi_{0,rough} = 2.33 \frac{J}{cm^2}$
  - $\phi_{0,smooth} = 6 \frac{J}{cm^2}$
  - $n_{layer,rough} = 5$
  - $n_{layer,smooth} = 13$
- Results
  - $s_a = 390\ nm,\ t = 1.79\ \mu m \rightarrow \frac{\mathrm{dV}}{\mathrm{dt}} = 1.05 \frac{\mathrm{mm}^3}{\mathrm{min}}$



- Example: Steel Surfaces

Smart Laser Micromachining Platform - Scheme

IDSIA Istituto Dalle Molle di studi sull'intelligenza artificiale USI/SUPSI

Parameter Set $\vec{P}$

- $\phi_{0,SP}$: Peak Fluence of a Single Pulse
- n_b: #Pulses per Burst
- o: Overlap
- r: Random
- f_r: Repetition rate
- $E_{P,b}$: "burst dynamics"
- $\Delta\tau$: Pulse duration
- w_0: Spot size

Optimization Target

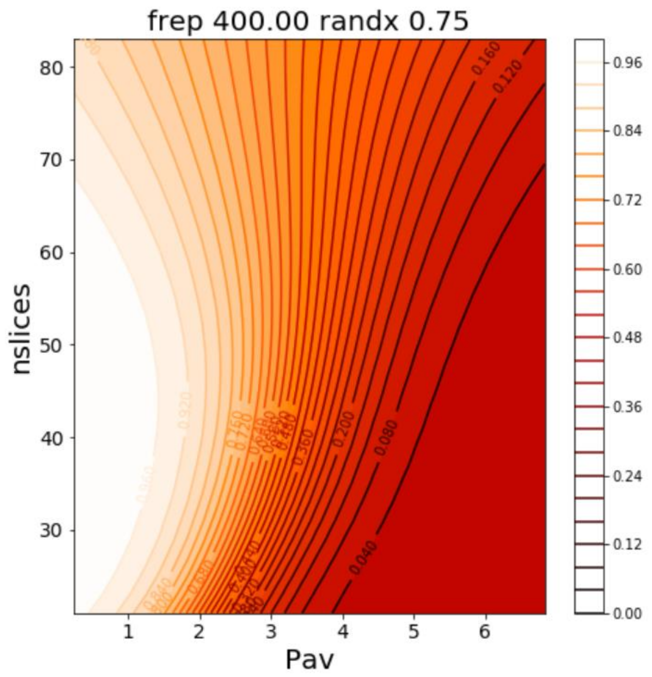
- s_a: Surface roughness
- Define cost Function:

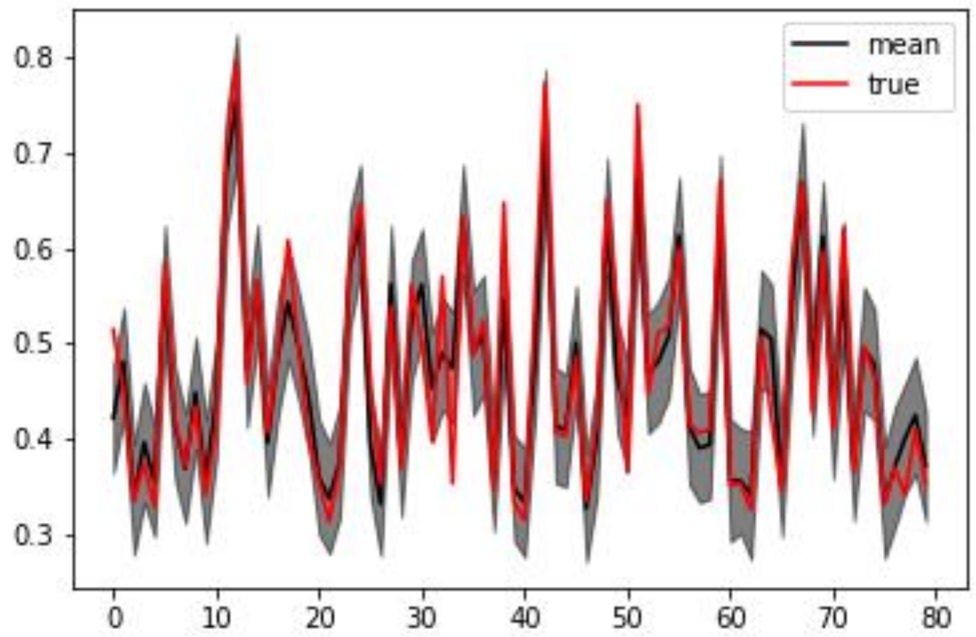
$$cf = s_a + Penalty(d < 2\ \mu m) + Penalty(Bad\ regions)$$

- **Goal: Autonomously find minimum of $\boldsymbol{cf}$**

Model Learning from Data

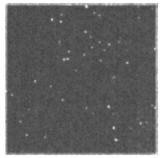
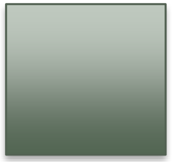
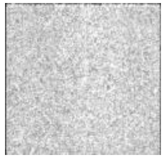
- Learning models from data with AISI 304
  Gaussian processes are employed to learn models from data.
  - Learned models for the roughness in laser-micromachining:





ML Based Optimization on AISI 420 with Agile fs System

Material: Steel AISI 420 (1.2083)
Laser source: NKT/CSEM
Number of samples: 110
Optimization method: AI based

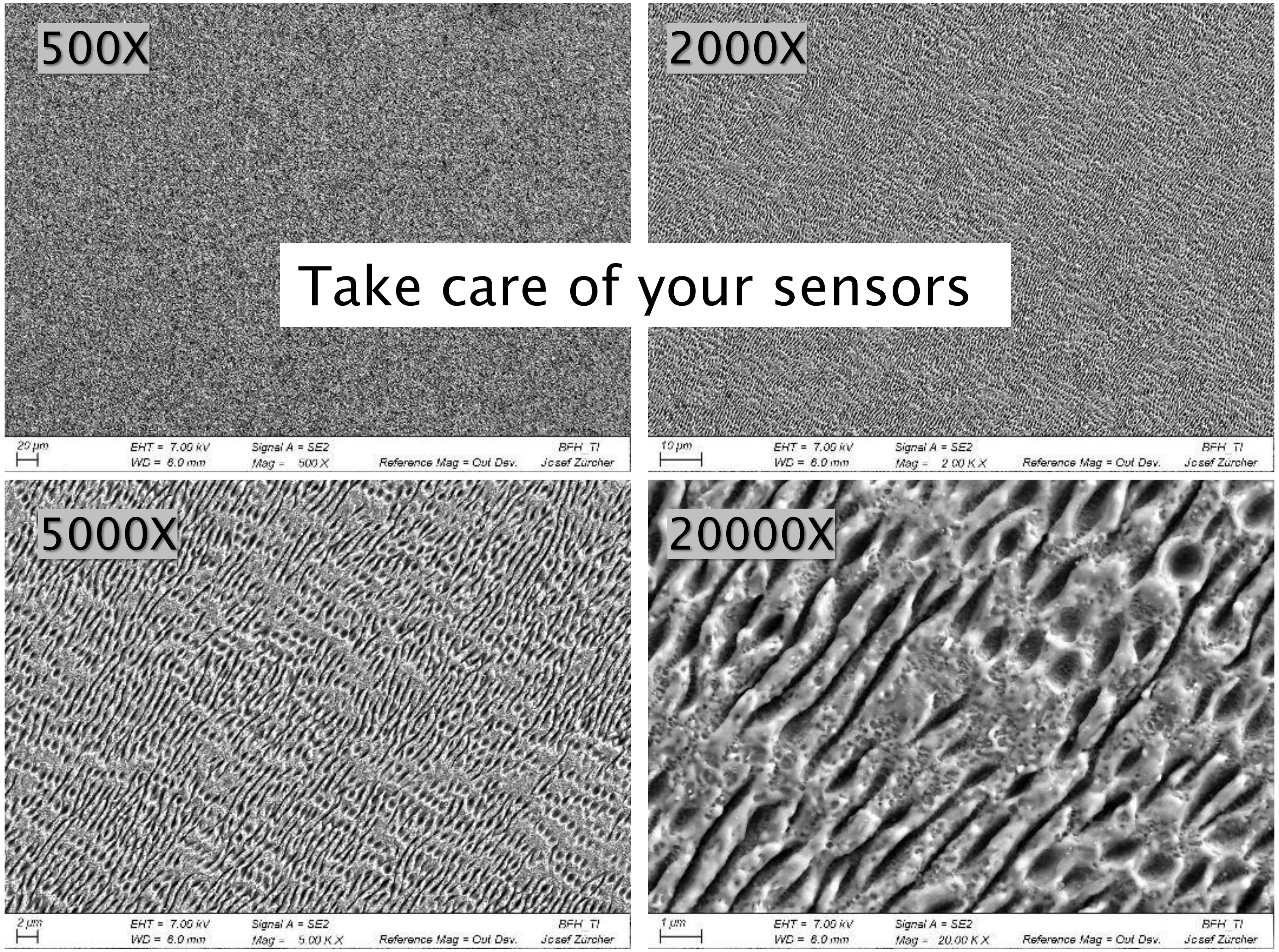




ABLATION	Sample 48	Sample 50	Sample 52			
Best roughness value s_a / μm	0,294	0,292	0,139			
Pitch $p_x = p_y$ / µm	5,6	8	8			
Random / μm	1,4	2	2			
n_{slices}	200	200	200			
f_r / kHz	2000	2000	2000			
$\phi_{0,SP}$ / J/cm^2	0,510	1,706	0,153			
n_b	2	2	2			
w_0 / μm	10,7	12,4	12,4			
$\Delta\tau$ / ps	2ps	350fs	2ps			

ML Based Optimization on AISI 420 with Agile fs System

Material: Steel AISI 420 (1.2083)
Laser source: NKT/CSEM
Number of samples: 110
Optimization method: AI based

ABLATION	Sample 52
Best roughness value s_a / μm	0,139
Pitch $p_x = p_y$ / µm	8
Random / μm	2
n_{slices}	200
f_r / kHz	2000
$\phi_{0,SP}$ / J/cm^2	0,153
n_b	2
w_0 / μm	12,4
$\Delta\tau$ / ps	2ps



Summary / Outlook

- We demonstrated the efficient optimization of a two-step process (roughening – smoothening) with Bayesian optimization.
- This method represents a powerful tool for a tremendous reduction of the demanded number of experiments and can be automized.

- For Ge a good set of parameter was found after 40 experiments instead of ≈ 1000 for a systematic study.

- But take care
  - The definition of the cost function uses detailed knowledge about the goals.
  - Use adequate sensors.