Inverse problems: Take an investigation the wrong way round, and you'll find it easier to solve!

Nathanaël Munier, Emmanuel Soubies and Pierre Weiss

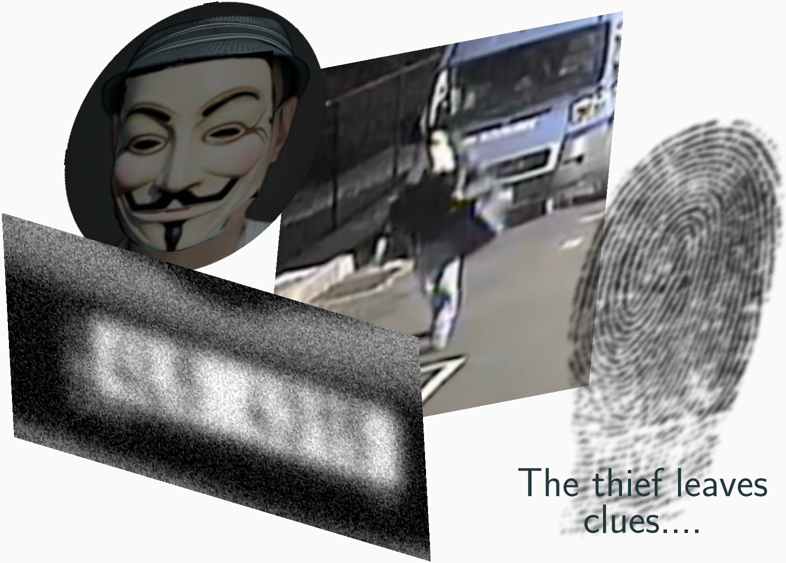
February 8, 2024

SOMEONE HAS STOLEN HAÏ'S LAPTOP !!!!!

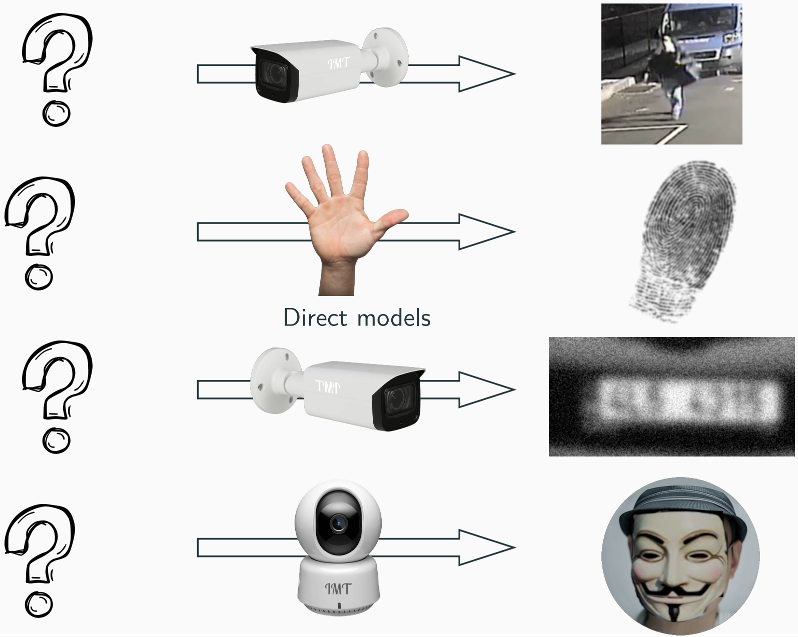
Crime scene and clues presentation

Crime scene and clues presentation

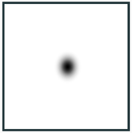
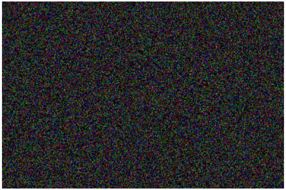
Crime scene and clues presentation



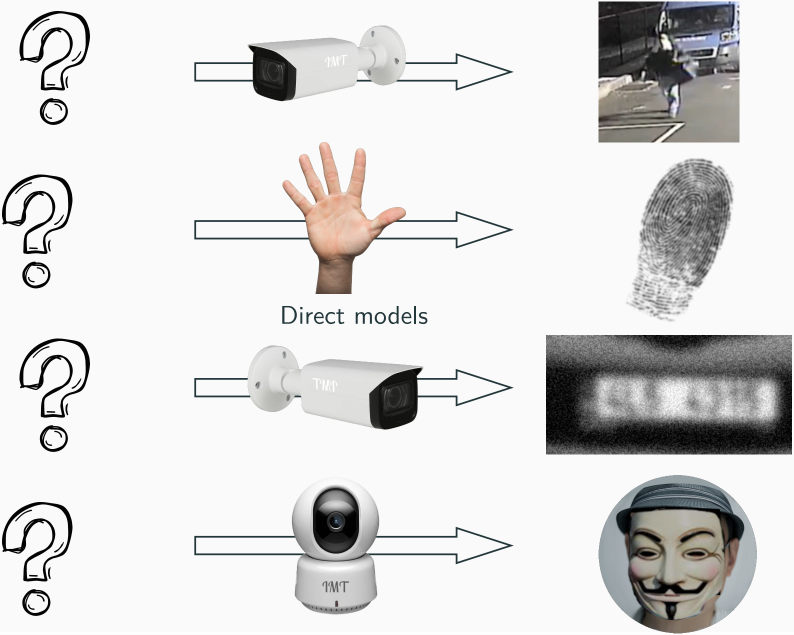
The thief leaves clues....


Direct models

$=$

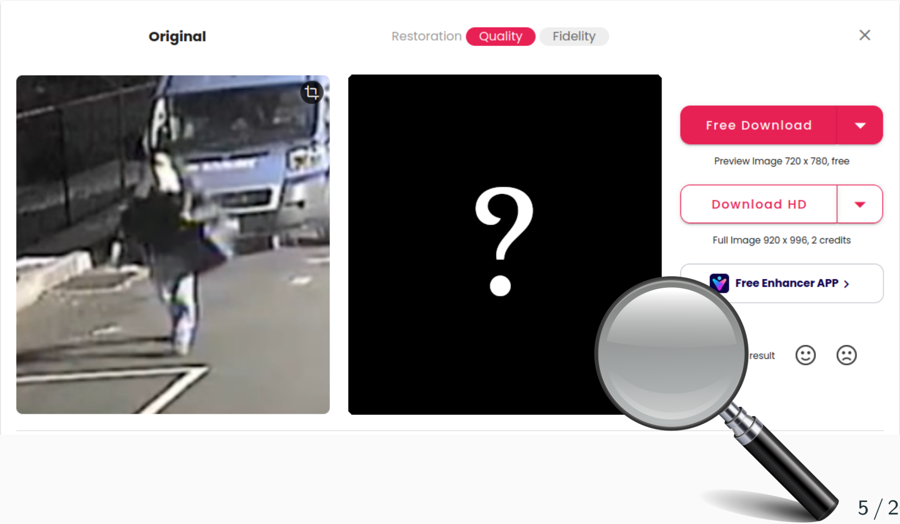
 $*$  $+$ 

$$y = x * h + \varepsilon, \qquad x * h(s) = \int x(t)\, h(s-t)\, dt$$


IMT
Direct models
IMT

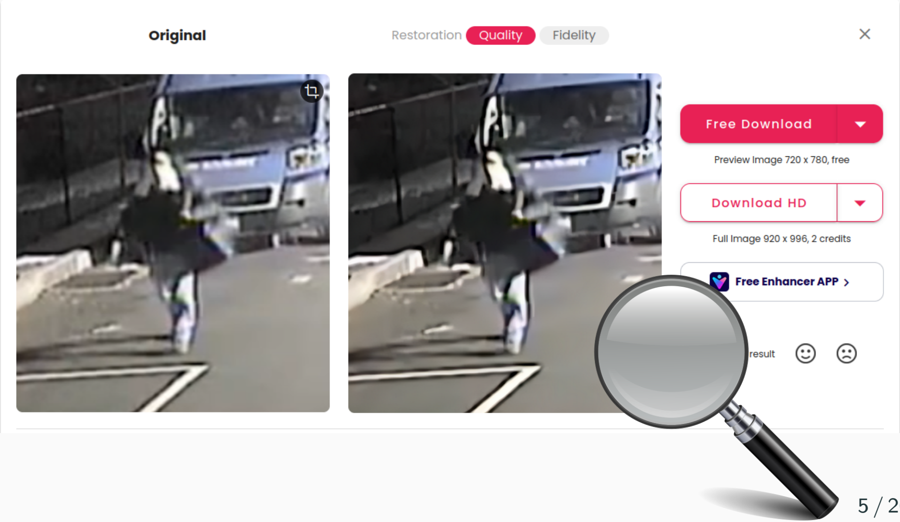
Clue 1: Enhance it!!!!

Let's do as in our best TV detective series



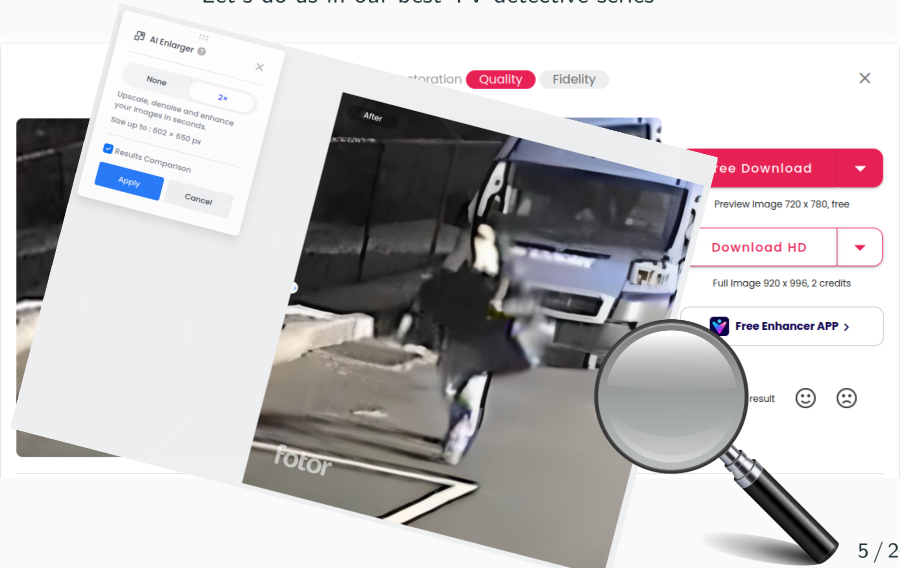
Clue 1: Enhance it!!!!

Let's do as in our best TV detective series



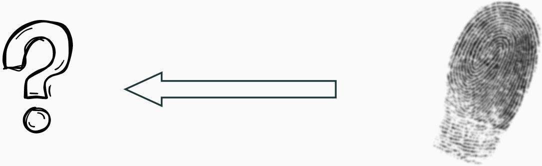
Clue 1: Enhance it!!!!

Let's do as in our best TV detective series

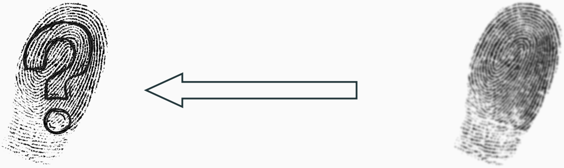


Clue 1: Enhance it!!!!

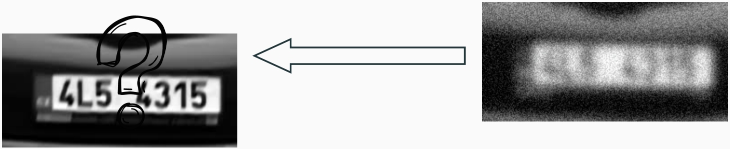
Let's do as in our best TV detective series



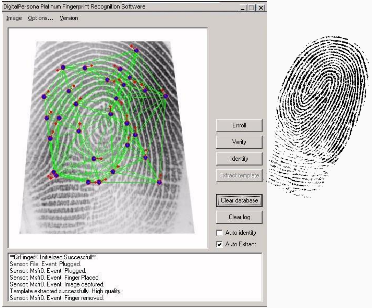
Inverse problems



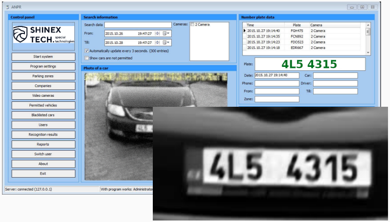
Inverse problems

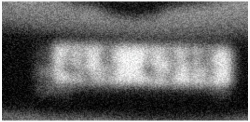


Clues 2 & 3



Inverse problems



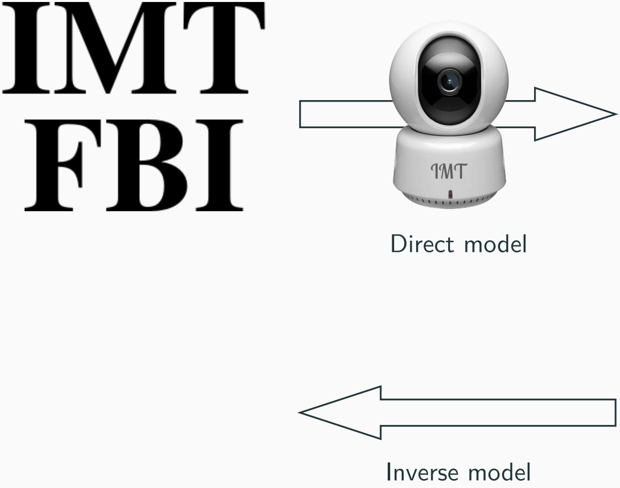


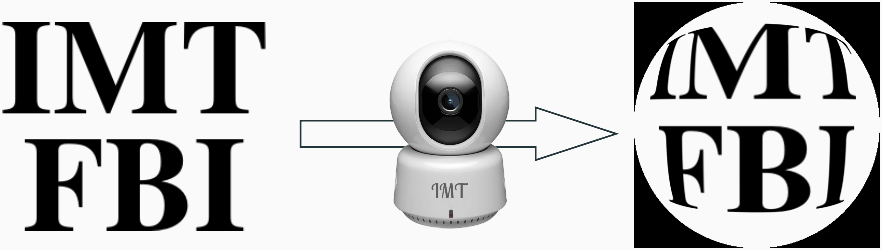
Clue 4: last but not least

Direct model

Inverse model

Clue 4: last but not least

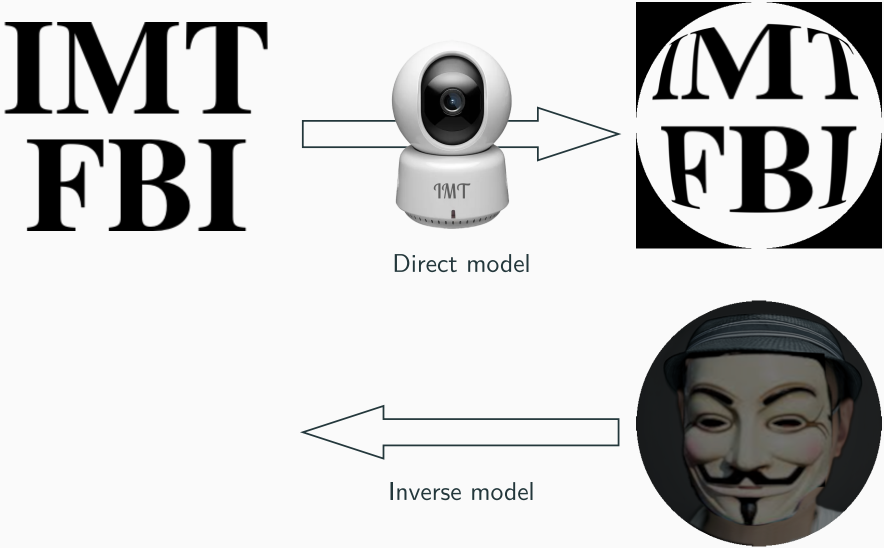


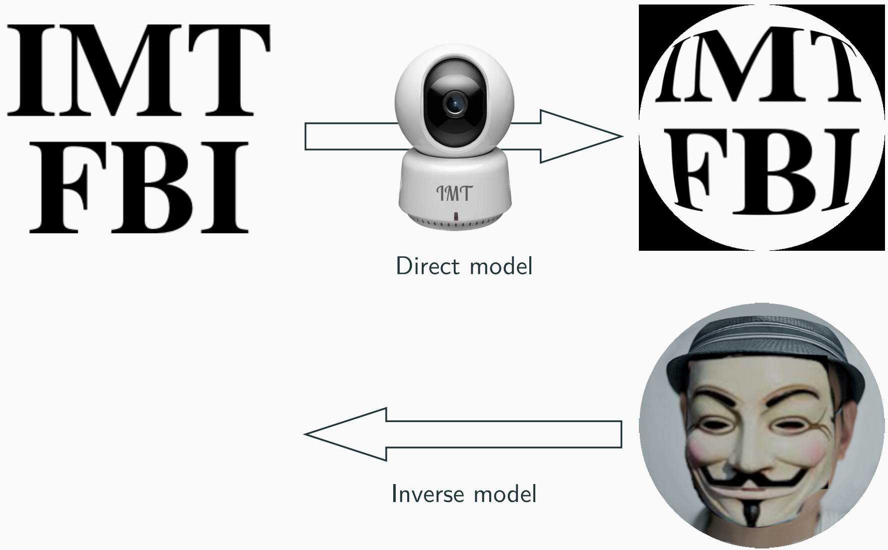


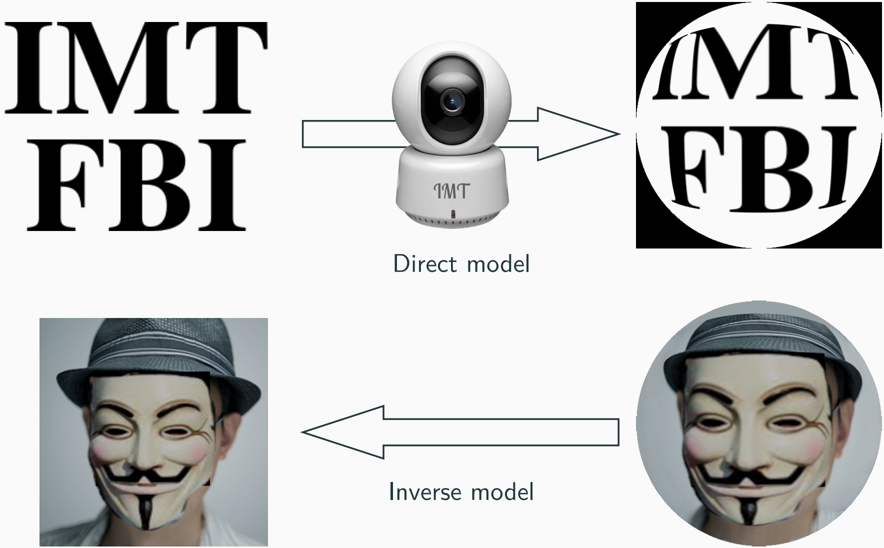
Direct model

Inverse model

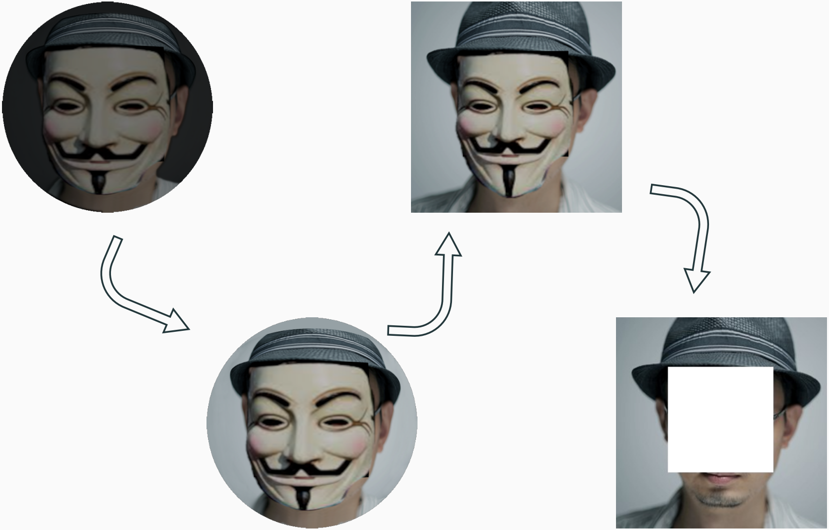
Clue 4: last but not least




IMT
FBI
IMT
IMT
FBI
Direct model
Inverse model


IMT
FBI
IMT
IMT
FBI
Direct model
Inverse model

Clue 4: last but not least



Inverse problem framework

$$y = Ax + \varepsilon$$

$$y \in \mathbb{R}^M, \quad x \in \mathbb{R}^N, \quad A : \mathbb{R}^N \to \mathbb{R}^N, \quad \varepsilon \sim \mathcal{N}(0, \sigma^2 \mathrm{Id})$$

Inverse problem framework

$$y = Ax + \varepsilon$$

$$y \in \mathbb{R}^M, \quad x \in \mathbb{R}^N, \quad A : \mathbb{R}^N \to \mathbb{R}^N, \quad \varepsilon \sim \mathcal{N}(0, \sigma^2 \mathrm{Id})$$

Stochastic framework:

- X of pdf (probability density function) p_X
- E of pdf (probability density function) $p_E \propto \exp\left(-\frac{\|\varepsilon\|_2}{2\sigma^2}\right)$
- Y of pdf (probability density function) p_Y

with the relation $Y = AX + E$.

Inverse problem framework

$$y = Ax + \varepsilon$$

$$y \in \mathbb{R}^M, \quad x \in \mathbb{R}^N, \quad A : \mathbb{R}^N \to \mathbb{R}^N, \quad \varepsilon \sim \mathcal{N}(0, \sigma^2 \mathrm{Id})$$

For simplicity, let's denote $\forall x \in \mathbb{R}^N, y \in \mathbb{R}^M$,

$$p(x) \stackrel{\text{def}}{=} p_X(x), \quad p(y) \stackrel{\text{def}}{=} p_Y(y),$$

$$p(x|y) \stackrel{\text{def}}{=} p_{X|Y}(x|y), \quad p(x) \stackrel{\text{def}}{=} p_{Y|X}(y|x).$$

$$y = Ax + \varepsilon$$

Knowing the clues y, is it possible to recover informations x on the robber?

$$y = Ax + \varepsilon$$

Knowing the clues y, is it possible to recover informations x on the robber?

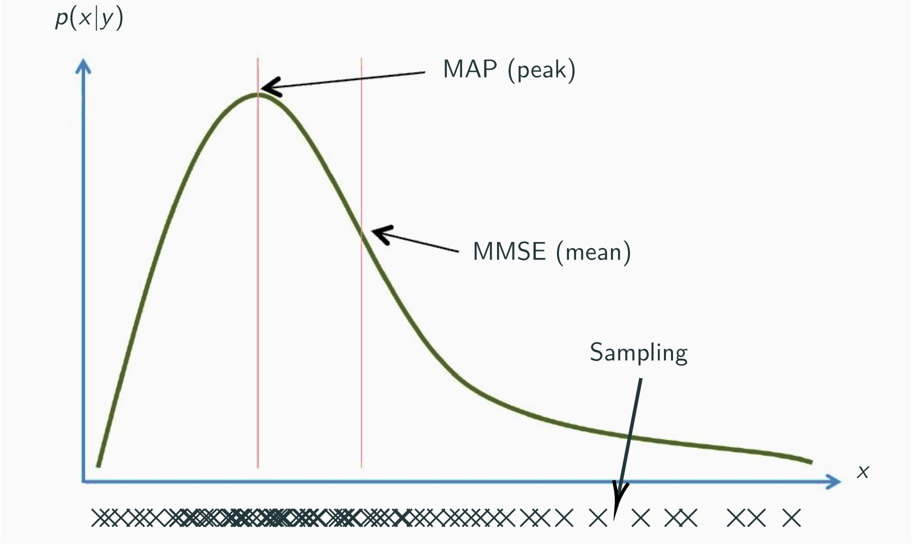
All of this is linked with the probability $p(x|y)$.

$$\hat{x}_{\mathsf{MAP}} = \operatorname*{argmax}_x \; p(x|y) \qquad \text{(maximum a posteriori)}$$

$$\hat{x}_{\mathsf{MMSE}} = \mathbb{E}\,(X|Y=y) = \int x \; p(x|y) \, \mathrm{d}x \qquad \text{(minimum mean square error)}$$

Sampling $\hat{x}_k \sim p(X|y)$ i.i.d

All of this is linked with the probability $p(x|y)$.



Maximum a posteriori (MAP)

Bayesian formulation

$$\hat{x}_{\mathsf{MAP}} = \operatorname*{argmax}_{x} \; p(x|y)$$

Maximum a posteriori (MAP)

Bayesian formulation

$$\begin{aligned}\hat{x}_{\mathsf{MAP}} &= \operatorname*{argmax}_x \; p(x|y) \\ &= \operatorname*{argmax}_x \; \frac{p(y|x) \cdot p(x)}{p(y)}\end{aligned}$$

Maximum a posteriori (MAP)

Bayesian formulation

$$\hat{x}_{\text{MAP}} = \operatorname*{argmax}_x \; p(x|y)$$
$$= \operatorname*{argmax}_x \; \frac{p(y|x) \cdot p(x)}{p(y)}$$
$$= \operatorname*{argmax}_x \; p(y|x) \cdot p(x)$$

Maximum a posteriori (MAP)

Bayesian formulation

$$
\begin{aligned}
\hat{x}_{\mathsf{MAP}} &= \underset{x}{\operatorname{argmax}} \; p(x|y) \\
&= \underset{x}{\operatorname{argmax}} \; \frac{p(y|x) \cdot p(x)}{p(y)} \\
&= \underset{x}{\operatorname{argmax}} \; p(y|x) \cdot p(x) \\
&= \underset{x}{\operatorname{argmin}} \; -\log p(y|x) - \log p(x)
\end{aligned}
$$

Maximum a posteriori (MAP)

Bayesian formulation

$$
\begin{aligned}
\hat{x}_{\mathrm{MAP}} &= \operatorname*{argmax}_x \; p(x|y) \\
&= \operatorname*{argmax}_x \; \frac{p(y|x) \cdot p(x)}{p(y)} \\
&= \operatorname*{argmax}_x \; p(y|x) \cdot p(x) \\
&= \operatorname*{argmin}_x \; -\log p(y|x) - \log p(x) \\
&= \operatorname*{argmin}_x \; \frac{1}{2\sigma^2} \|Ax - y\|_2^2 - \log p(x)
\end{aligned}
$$

Maximum a posteriori (MAP)

Bayesian formulation

$$
\begin{aligned}
\hat{x}_{\mathsf{MAP}} &= \underset{x}{\operatorname{argmax}}\ \ p(x|y) \\
&= \underset{x}{\operatorname{argmax}}\ \ \frac{p(y|x) \cdot p(x)}{p(y)} \\
&= \underset{x}{\operatorname{argmax}}\ \ p(y|x) \cdot p(x) \\
&= \underset{x}{\operatorname{argmin}}\ \ -\log p(y|x) - \log p(x) \\
&= \underset{x}{\operatorname{argmin}}\ \ \frac{1}{2\sigma^2} \|Ax - y\|_2^2 - \log p(x)
\end{aligned}
$$

The remaining task is to estimate $p(x)$ that is the pdf of X.

Idea 1: Solve it by hand

What do we have at our disposal?

Idea 1: Ask random people to do the work for you.

Idea 1: Solve it by hand

Idea 1: Solve it by hand

What do we have at our disposal?

Idea 2: Every image are equally probable.

i.e. $X \sim \mathcal{U}([0,1]^N)$ and $\operatorname*{argmin}_{x \in [0,1]^N} \|Ax - y\|_2^2$

What do we have at our disposal?

Idea 2: Every image are equally probable.

i.e. $X \sim \mathcal{U}([0,1]^N)$ and $\underset{x \in [0,1]^N}{\operatorname{argmin}} \|Ax - y\|_2^2$

$\hat{x} \stackrel{\text{def}}{=} A^+ y$ is the solution that has the minimal ℓ^2-norm.

With the Moore-Penrose inverse matrix A^+ verifying

$$AA^+A = A, \quad (AA^+)^* = AA^+,$$
$$A^+AA^+ = A^+, \quad (A^+A)^* = A^+A.$$

Here $A^+ = A$ and thus $\hat{x} = A^+ y = y$!

The mystery remains...

Idea 3: Let's choose a simple structure

What do we have at our disposal?

Idea 3: Give the image a structure

$-\log p(x) = \lambda \|x\|_2^2$ - Tikhonov regularization

$-\log p(x) = \lambda \|x\|_1^2$ - LASSO regularization

for some $\lambda > 0$.

Idea 3: Let's choose a simple structure

What do we have at our disposal?

Idea 3: Give the image a structure

$-\log p(x) = \lambda \|x\|_2^2$ - Tikhonov regularization

$-\log p(x) = \lambda \|x\|_1^2$ - LASSO regularization

for some $\lambda > 0$.

As before, the solution is  .

What do we have at our disposal?

Idea 4: Some face database.

There are several face database like Flickr-Faces-HQ (FFHQ).

Let's say we have a set of face pictures $(x_i)_i$. At the limit

$$\frac{1}{n}\sum_{i=1}^{n} \delta_{x_i} \underset{n\to\infty}{\rightharpoonup} p_X.$$

How to improve $\frac{1}{n}\sum_{i=1}^{n}\delta_{x_i}$ to get a better approximation of p_X?

How to improve $\frac{1}{n}\sum_{i=1}^{n}\delta_{x_i}$ to get a better approximation of p_X?

$\rightarrow$ Play with this database, noise and denoise it!?!

Noising: $x \mapsto x + \varepsilon$

Denoising: $\mathcal{D}_{\theta^*} : x \in \mathbb{R}^N \mapsto \mathcal{D}_{\theta^*}(x) \in \mathbb{R}^N, \quad \theta^* \in \mathbb{R}^P$

How to improve $\frac{1}{n}\sum_{i=1}^{n}\delta_{x_i}$ to get a better approximation of p_X?

$\rightarrow$ Play with this database, noise and denoise it!?!

Noising: $x \mapsto x + \varepsilon$

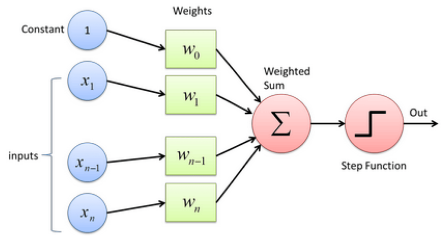
Denoising: $\mathcal{D}_{\theta^*} : x \in \mathbb{R}^N \mapsto \mathcal{D}_{\theta^*}(x) \in \mathbb{R}^N, \quad \theta^* \in \mathbb{R}^P$

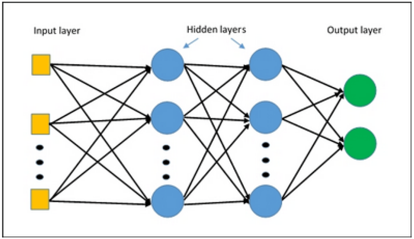
How to get such a denoising function?

$$\mathcal{D}_{\theta^*} : x \in \mathbb{R}^N \mapsto \mathcal{D}_{\theta^*}(x) \in \mathbb{R}^N, \quad \theta^* \in \mathbb{R}^P$$

How to get such a denoising function?

$$\mathcal{D}_{\theta^*} : x \in \mathbb{R}^N \mapsto \mathcal{D}_{\theta^*}(x) \in \mathbb{R}^N, \quad \theta^* \in \mathbb{R}^P$$





$\rightarrow$ From the universal approximation of neural network!

Universal approximation theorem — Let $C(X, \mathbb{R}^m)$ denote the set of continuous functions from a subset X of a Euclidean $\mathbb{R}^n$ space to a Euclidean space $\mathbb{R}^m$. Let $\sigma \in C(\mathbb{R}, \mathbb{R})$. Note that $(\sigma \circ x)_i = \sigma(x_i)$, so $\sigma \circ x$ denotes σ applied to each component of x.

Then σ is not polynomial if and only if for every $n \in \mathbb{N}$, $m \in \mathbb{N}$, compact $K \subseteq \mathbb{R}^n$, $f \in C(K, \mathbb{R}^m), \varepsilon > 0$ there exist $k \in \mathbb{N}$, $A \in \mathbb{R}^{k \times n}$, $b \in \mathbb{R}^k$, $C \in \mathbb{R}^{m \times k}$ such that

$$\sup_{x \in K} \|f(x) - g(x)\| < \varepsilon$$

where $g(x) = C \cdot (\sigma \circ (A \cdot x + b))$

$$\mathcal{D}_{\theta^*} : x \in \mathbb{R}^N \mapsto \mathcal{D}_{\theta^*}(x) \in \mathbb{R}^N, \quad \theta^* \in \mathbb{R}^P$$

Moreover, we can use optimization theory to find $\theta^* \in \mathbb{R}^P$ verifying

$$\theta^* = \underset{\theta}{\operatorname{argmin}} F(\theta).$$

It only remains to find a good function $F : \mathbb{R}^P \to \mathbb{R}$.

$$\mathcal{D}_{\theta^*} : x \in \mathbb{R}^N \mapsto \mathcal{D}_{\theta^*}(x) \in \mathbb{R}^N, \quad \theta^* \in \mathbb{R}^P$$

Moreover, we can use optimization theory to find $\theta^* \in \mathbb{R}^P$ verifying

$$\theta^* = \underset{\theta}{\operatorname{argmin}}\, F(\theta).$$

It only remains to find a good function $F : \mathbb{R}^P \to \mathbb{R}$.

$$F : \theta \to \|\mathcal{D}_{\theta^*}(x + \varepsilon) - x\|_2^2 \quad ?$$

$$\mathcal{D}_{\theta^*} : x \in \mathbb{R}^N \mapsto \mathcal{D}_{\theta^*}(x) \in \mathbb{R}^N, \quad \theta^* \in \mathbb{R}^P$$

Moreover, we can use optimization theory to find $\theta^* \in \mathbb{R}^P$ verifying

$$\theta^* = \operatorname*{argmin}_{\theta} F(\theta).$$

It only remains to find a good function $F : \mathbb{R}^P \to \mathbb{R}$.

$$F : \theta \to \|\mathcal{D}_{\theta^*}(x + \varepsilon) - x\|_2^2 \quad ?$$

Better choice:

$$F : \theta \to \frac{1}{n} \sum_{i=1}^{n} \|\mathcal{D}_{\theta^*}(x_i + \varepsilon_i) - x_i\|_2^2$$

$$\mathcal{D}_{\theta^*} : \mathbb{R}^N \to \mathbb{R}^N, \quad \text{with} \quad \theta^* = \underset{\theta}{\operatorname{argmin}} \frac{1}{n} \sum_{i=1}^{n} \|\mathcal{D}_{\theta^*}(x_i + \varepsilon_i) - x_i\|_2^2$$

$$\mathcal{D}_{\theta^*} : \mathbb{R}^N \to \mathbb{R}^N, \quad \text{with} \quad \theta^* = \operatorname*{argmin}_\theta \frac{1}{n} \sum_{i=1}^{n} \|\mathcal{D}_{\theta^*}(x_i + \varepsilon_i) - x_i\|_2^2$$

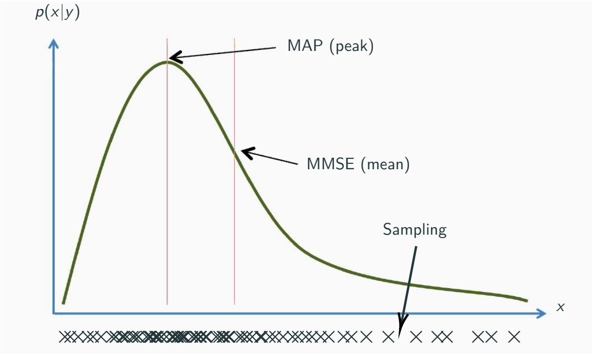
$$\underset{n\to\infty}{\longrightarrow} \operatorname*{argmin}_\theta \mathbb{E}_{E,X}(\|\mathcal{D}_\theta(Y) - X\|_2^2 \,|\, Y = X + E).$$

$$\mathcal{D}_{\theta^*} : \mathbb{R}^N \to \mathbb{R}^N, \quad \text{with} \quad \theta^* = \operatorname*{argmin}_{\theta} \frac{1}{n} \sum_{i=1}^{n} \|\mathcal{D}_{\theta^*}(x_i + \varepsilon_i) - x_i\|_2^2$$

$$\underset{n\to\infty}{\longrightarrow} \operatorname*{argmin}_{\theta} \mathbb{E}_{E,X}(\|\mathcal{D}_\theta(Y) - X\|_2^2 \,|\, Y = X + E).$$

That is the MMSE (minimum mean square error)!

$$\mathcal{D}_{\theta^*}(Y) \simeq \mathbb{E}_{E,X}(X|Y = X + E) = \hat{x}_{\mathrm{MMSE}}.$$



What we have so far...

What we have so far...

- A clue y with lacking information

What we have so far...

- A clue y with lacking information
  - A direct model: $y = Ax + \varepsilon$

What we have so far...

- A clue y with lacking information
  - A direct model: $y = Ax + \varepsilon$
    - A face database $(x_i)_i$

What we have so far...

- A clue y with lacking information
  - A direct model: $y = Ax + \varepsilon$
    - A face database $(x_i)_i$
- A denoiser $\mathcal{D}_{\theta^*} \simeq \hat{x}_{\mathrm{MMSE}}$ for $A = \mathrm{Id}$

What we have so far...

- A clue y with lacking information
  - A direct model: $y = Ax + \varepsilon$
    - A face database $(x_i)_i$
- A denoiser $\mathcal{D}_{\theta^*} \simeq \hat{x}_{\text{MMSE}}$ for $A = \mathrm{Id}$
  - And so what ? ...

What we have so far...

- A clue y with lacking information
- A direct model: $y = Ax + \varepsilon$
- A face database $(x_i)_i$
- A denoiser $\mathcal{D}_{\theta^*} \simeq \hat{x}_{\text{MMSE}}$ for $A = \mathrm{Id}$
- And so what ? ...

The trick is to use Tweedy's formula:

$$\hat{x}_{\text{MMSE}} = \mathbb{E}\,(X|Y) = Y + ...$$

- A clue y with lacking information
- A direct model: $y = Ax + \varepsilon$
- A face database $(x_i)_i$
- A denoiser $\mathcal{D}_{\theta^*} \simeq \hat{x}_{\text{MMSE}}$ for $A = \mathrm{Id}$
- And so what ? ...

The trick is to use Tweedy's formula:

$$\hat{x}_{\text{MMSE}} = \mathbb{E}\left(X|Y\right) = Y + \sigma^2 \nabla \log p_X(Y).$$

What we have so far...

- A clue y with lacking information
- A direct model: $y = Ax + \varepsilon$
- A face database $(x_i)_i$
- A denoiser $\mathcal{D}_{\theta^*} \simeq \hat{x}_{\mathrm{MMSE}}$ for $A = \mathrm{Id}$
- And so what ? ...

The trick is to use Tweedy's formula:

$$\hat{x}_{\mathrm{MMSE}} = \mathbb{E}\left(X|Y\right) = Y + \sigma^2 \nabla \log p_X(Y).$$

In particular

$$\mathcal{D}_{\theta^*}(y) \simeq \mathbb{E}\left(X|Y = y\right) = y + \sigma^2 \nabla \log p_X(y).$$

What we have so far...

- A clue y with lacking information
- A direct model: $y = Ax + \varepsilon$
- A face database $(x_i)_i$
- A denoiser $\mathcal{D}_{\theta^*} \simeq \hat{x}_{\mathrm{MMSE}}$ for $A = \mathrm{Id}$
- And so what ? ...

The trick is to use Tweedy's formula:

$$\hat{x}_{\mathrm{MMSE}} = \mathbb{E}(X|Y) = Y + \sigma^2 \nabla \log p_X(Y).$$

In particular

$$\mathcal{D}_{\theta^*}(y) \simeq \mathbb{E}(X|Y = y) = y + \sigma^2 \nabla \log p_X(y).$$

Or more explicitly

$$\boxed{\nabla \log p_X(y) \simeq \frac{\mathcal{D}_{\theta^*}(y) - y}{\sigma^2}}.$$

$$\hat{x}_{\text{MAP}} = \underset{x}{\operatorname{argmax}} \; p(x|y)$$

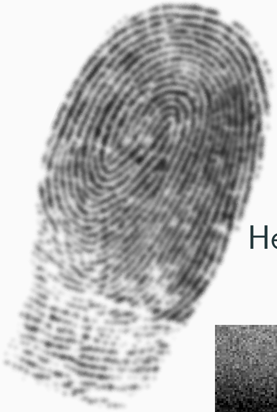
$$\begin{aligned}\hat{x}_{\mathrm{MAP}} &= \operatorname*{argmax}_x \; p(x|y) \\ &= \operatorname*{argmin}_x \; \frac{1}{2}\|Ax - y\|_2^2 - \log p_X(x)\end{aligned}$$

$$\hat{x}_{\mathsf{MAP}} = \operatorname*{argmax}_x \; p(x|y)$$
$$= \operatorname*{argmin}_x \; \frac{1}{2} \|Ax - y\|_2^2 - \log p_X(x)$$

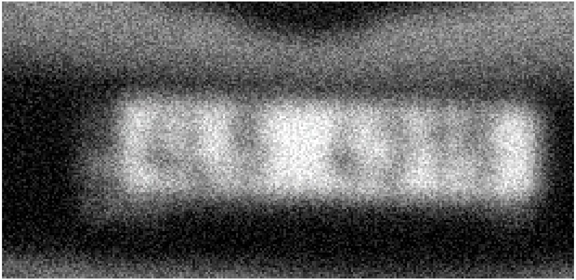
Now we have access to

$$\nabla \log p_X(y) \simeq \frac{\mathcal{D}_{\theta^*}(y) - y}{\sigma^2}$$

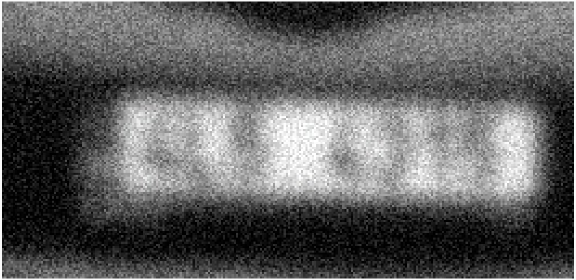
and are able to compute $\hat{x}_{\mathsf{MAP}}$ thanks to an optimization algorithm!!



Let's unveil
the results!
Here are the $\hat{x}_{\text{MAP}}$:

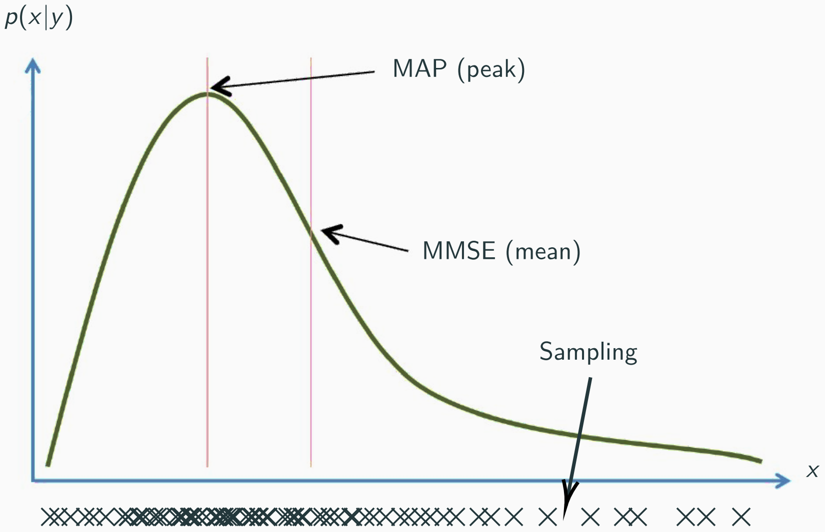


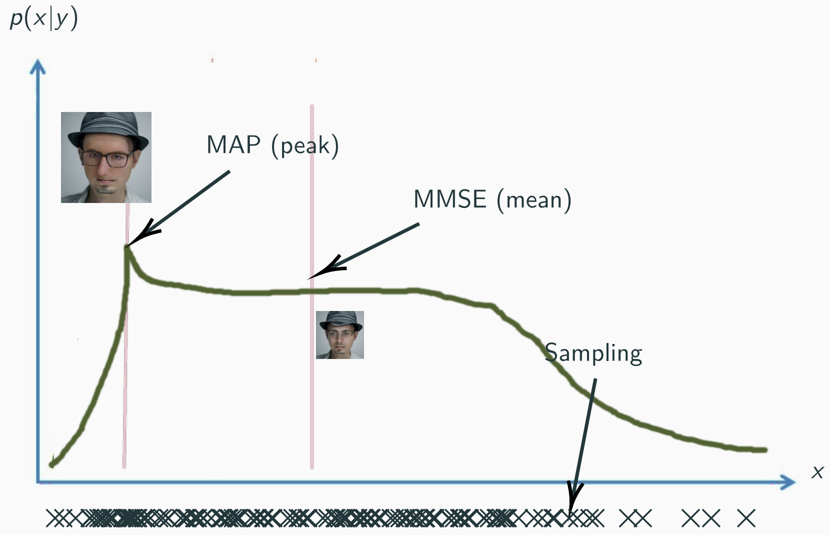
Let's unveil the results!
Here are the $\hat{x}_{\mathrm{MAP}}$:



Let's unveil
the results!
Here are the $\hat{x}_{\text{MAP}}$:

Let's unveil
the results!
Here are the $\hat{x}_{\mathrm{MAP}}$:


$p(x|y)$
MAP (peak)
MMSE (mean)
Sampling
x


$p(x|y)$
MAP (peak)
MMSE (mean)
Sampling
x

Sampling and diffusion model[1][2]

For $f = e^{-v}$ a probability distribution function, the SDE (over-damped Langevin stochastic differential equation)

$$\mathrm{d}X_t = \nabla v(X_t)\mathrm{d}t + \sqrt{2}\mathrm{d}B_t$$

gives a way to sample the distribution f.

Here B_t is the standard Brownian motion.

[1] Andre Wibisono. "Sampling as optimization in the space of measures: The Langevin dynamics as a composite optimization problem". In: 2018.

[2] William Feller. "On the Theory of Stochastic Processes, with Particular Reference to Applications". In: (1949).

Sampling and diffusion model

$$\mathrm{d}X_t = \nabla v(X_t)\mathrm{d}t + \sqrt{2}\mathrm{d}B_t$$

The only required knowledge on Brownian motion is

$$\mathrm{d}B_t \simeq \frac{1}{\delta}\left(B_{t+\delta} - B_t\right) \sim \mathcal{N}(0, \mathrm{Id}).$$

We can use the Euler-Maruyama (EM) discretization of SDE

$$X_{k+1} = X_k + \delta \nabla v(X_k) + \sqrt{2\delta} Z_{k+1}$$

where $Z_k \sim \mathcal{N}(0, \mathrm{Id})$ and $\delta > 0$.

We can use the Euler-Maruyama (EM) discretization of SDE

$$X_{k+1} = X_k + \delta \nabla v(X_k) + \sqrt{2\delta} Z_{k+1}$$

where $Z_k \sim \mathcal{N}(0, \mathrm{Id})$ and $\delta > 0$.

Sampling and diffusion model

We can use the Euler-Maruyama (EM) discretization of SDE

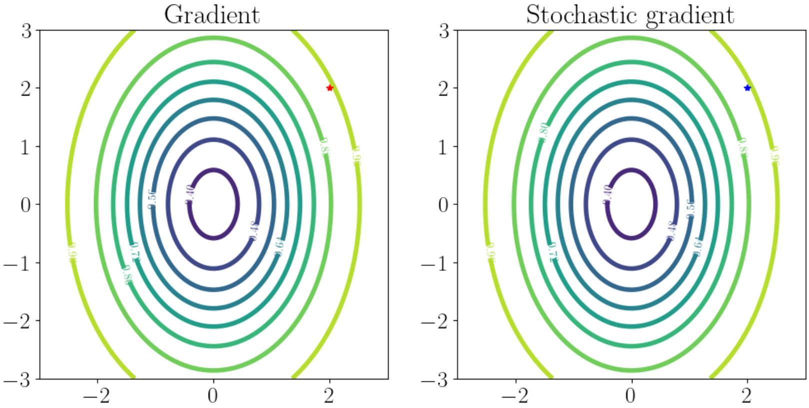
$$X_{k+1} = X_k + \delta \nabla v(X_k) + \sqrt{2\delta} Z_{k+1}$$

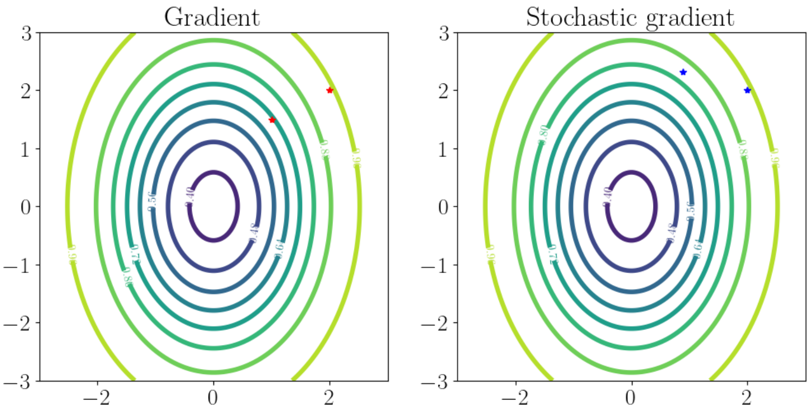
where $Z_k \sim \mathcal{N}(0, \mathrm{Id})$ and $\delta > 0$.

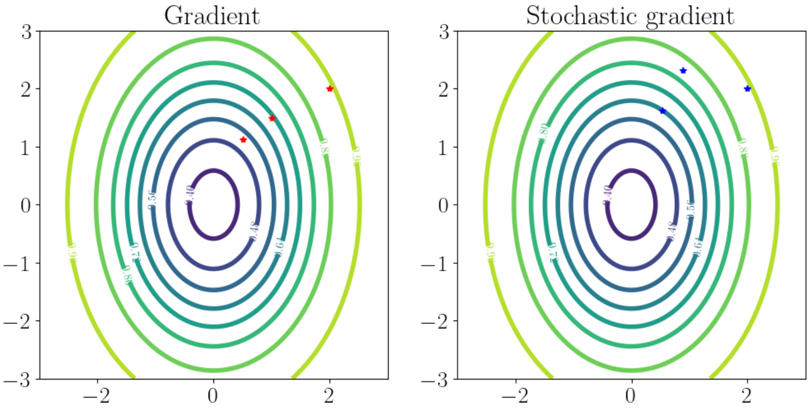
Finally, applying it to $f = p_{X|Y}$

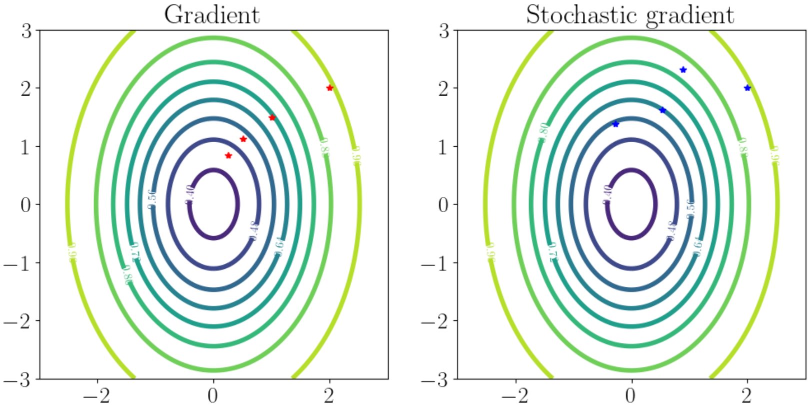
$$\begin{aligned} X_{k+1} &= X_k + \delta \nabla \log p(X_k|y) + \sqrt{2\delta} Z_{k+1} \\ &= X_k + \delta \nabla \log p(y|X_k) + \delta \boxed{\nabla \log p_X(X_k)} + \sqrt{2\delta} Z_{k+1} \end{aligned}$$

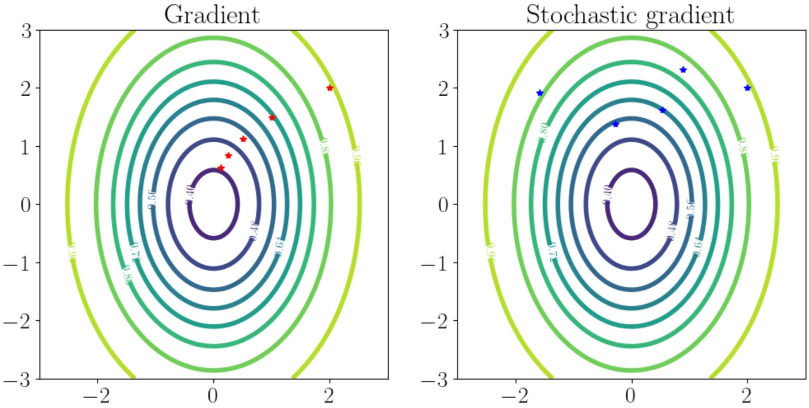
where $Z_k \sim \mathcal{N}(0, \mathrm{Id})$ and $\delta > 0$.

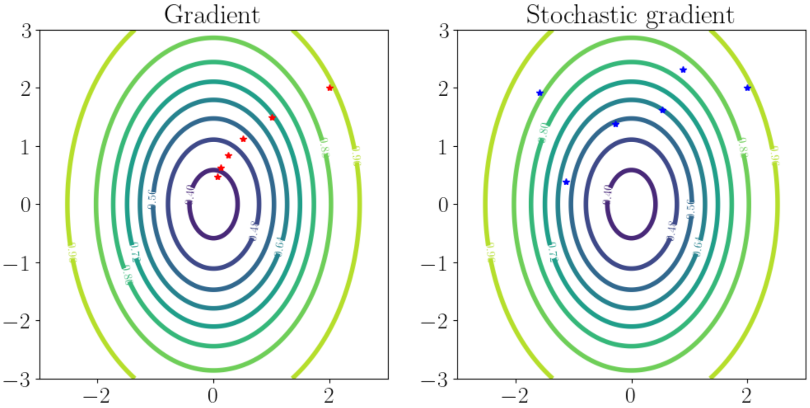

Gradient
Stochastic gradient


Gradient
Stochastic gradient
3
2
1
0
−1
−2
−3
−2
0
2
0.40
0.48
0.56
0.64
0.72
0.80
0.88
0.96

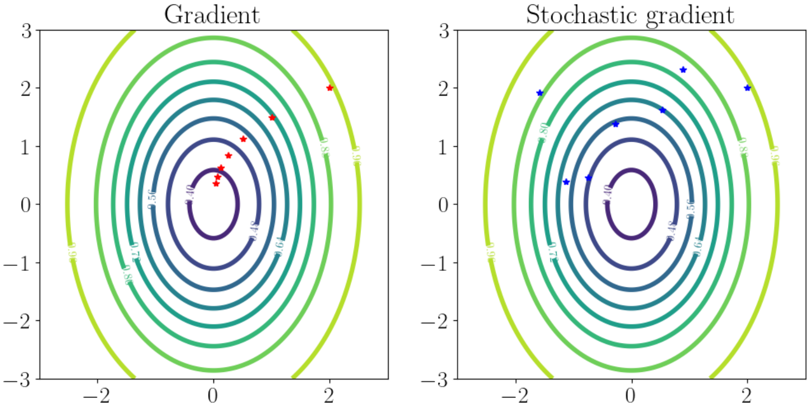

Gradient
Stochastic gradient
3
2
1
0
−1
−2
−3
−2
0
2

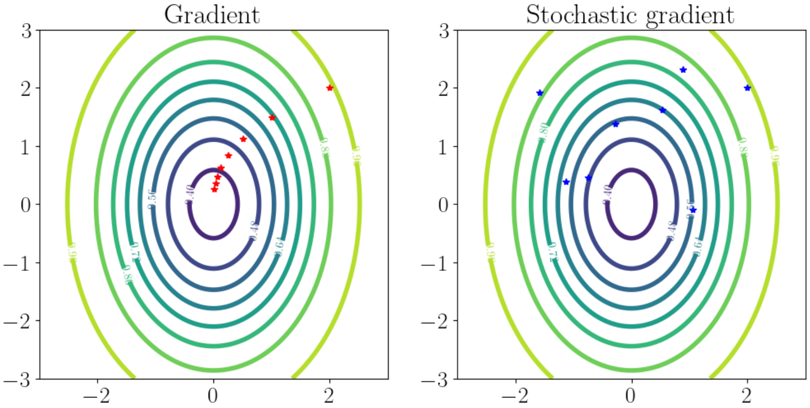

Gradient
Stochastic gradient
3
2
1
0
−1
−2
−3
−2
0
2


Gradient
Stochastic gradient
3
2
1
0
−1
−2
−3
−2
0
2
0.40
0.48
0.56
0.64
0.72
0.80
0.88
0.96

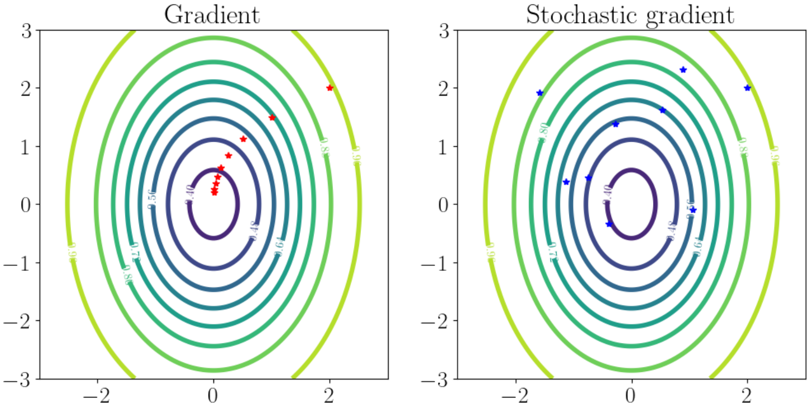


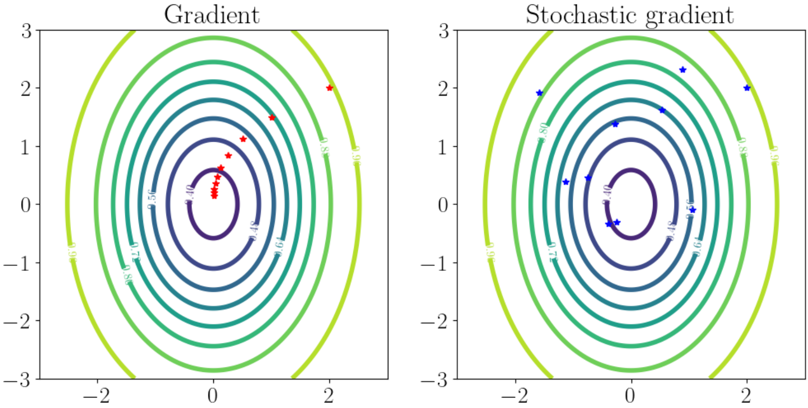
Gradient
Stochastic gradient
3
2
1
0
−1
−2
−3
−2
0
2
0.40
0.48
0.56
0.64
0.72
0.80
0.88
0.96

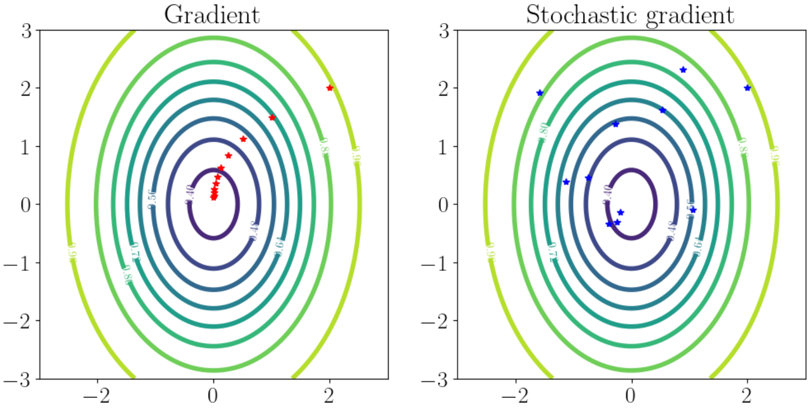

Gradient
Stochastic gradient
3
2
1
0
−1
−2
−3
−2
0
2
0.40
0.48
0.56
0.64
0.72
0.80
0.88
0.96

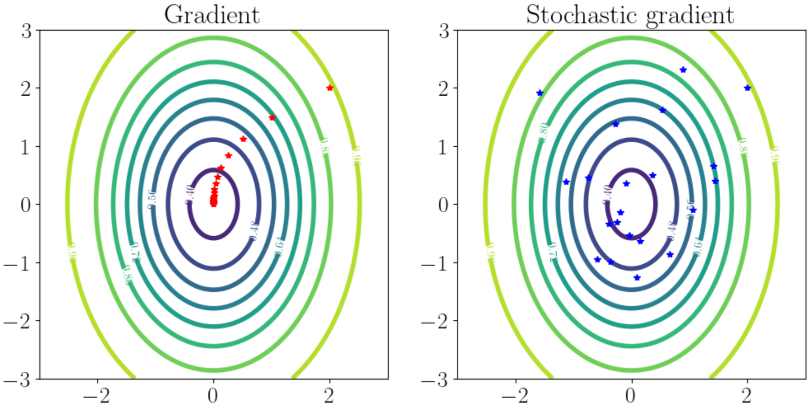


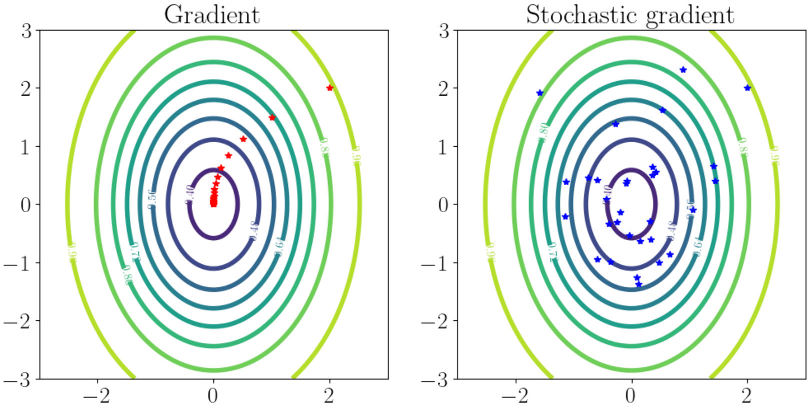
Gradient
Stochastic gradient

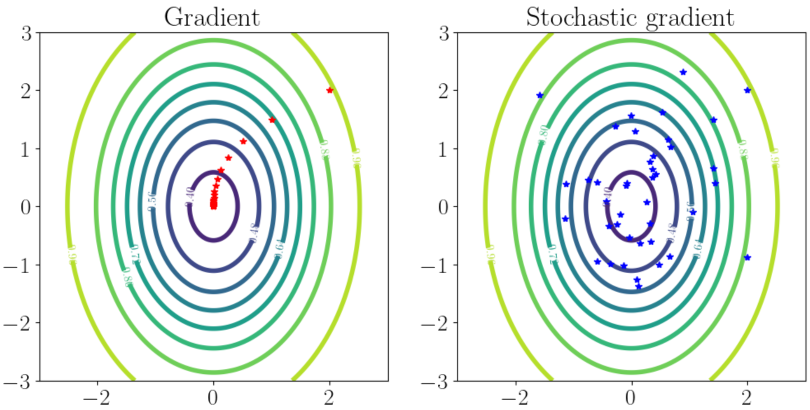

Gradient
Stochastic gradient
3
2
1
0
−1
−2
−3
−2
0
2


Gradient
Stochastic gradient
0.40
0.48
0.56
0.64
0.72
0.80
0.88
0.96
−3
−2
−1
0
1
2
3

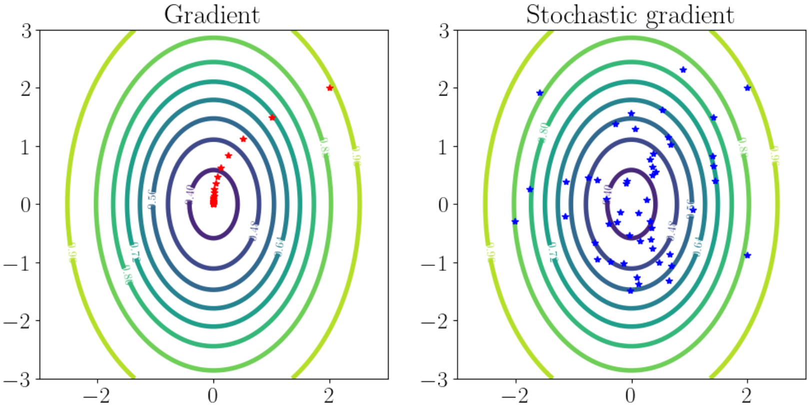

Gradient
Stochastic gradient
3
2
1
0
−1
−2
−3
−2
0
2


Gradient
Stochastic gradient

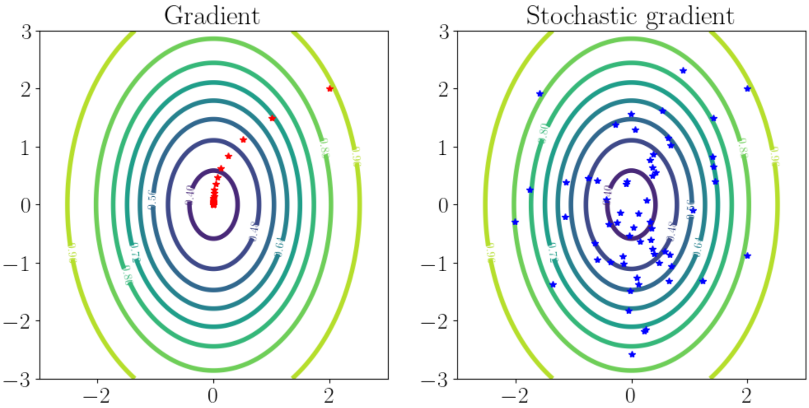

Gradient
Stochastic gradient
3
2
1
0
−1
−2
−3
−2
0
2

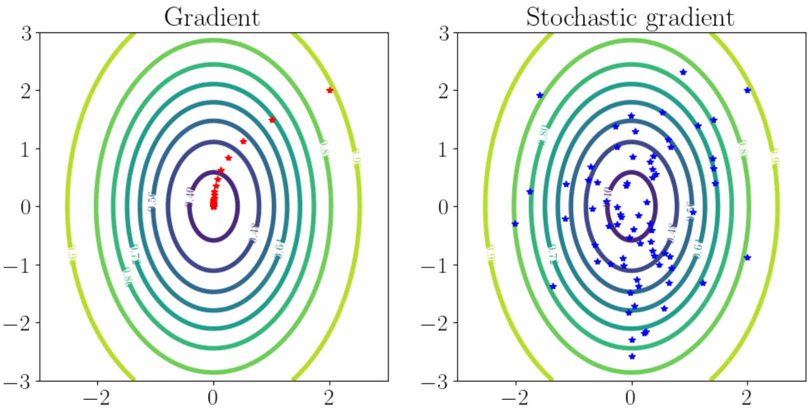


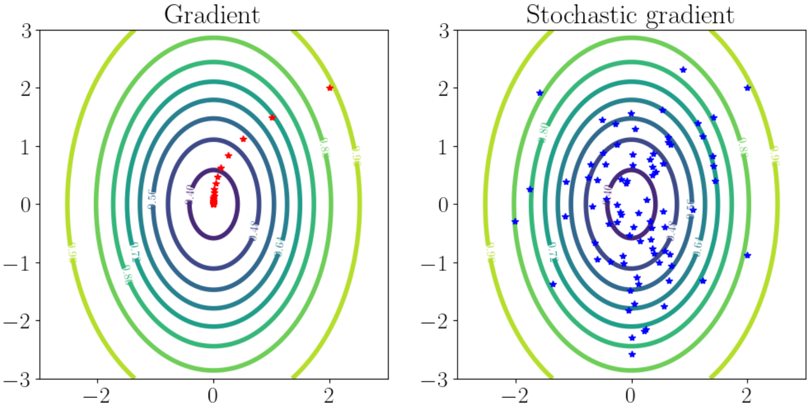
Gradient
Stochastic gradient

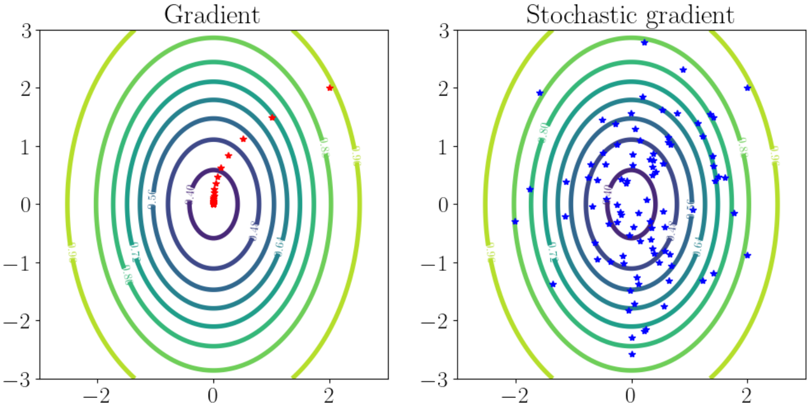


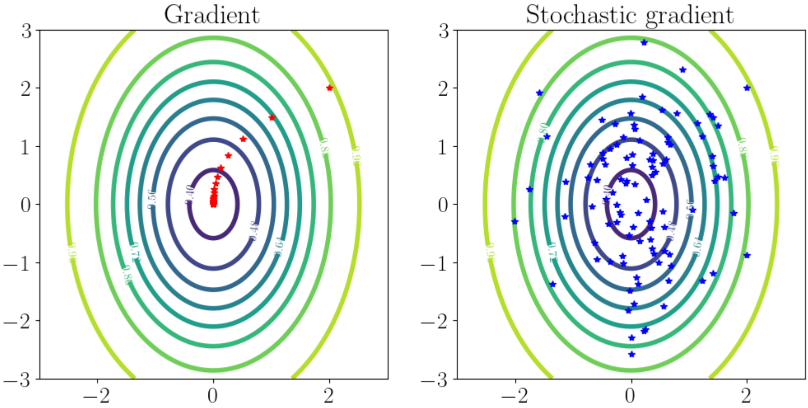
Gradient
Stochastic gradient
3
2
1
0
−1
−2
−3
−2
0
2
0.40
0.48
0.56
0.64
0.72
0.80
0.88
0.96

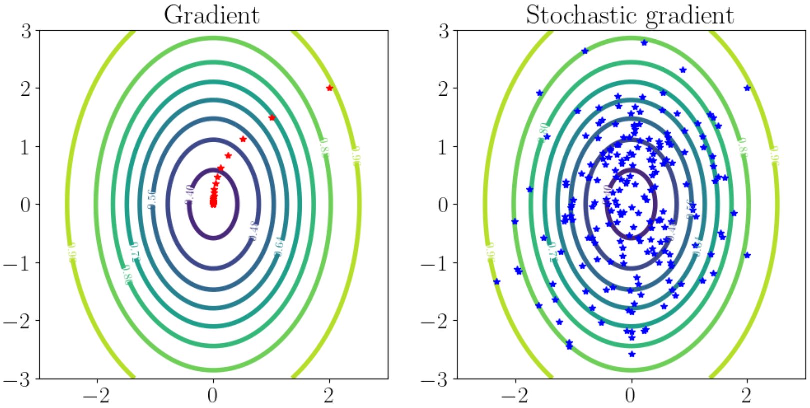

Gradient
Stochastic gradient
3
2
1
0
−1
−2
−3
−2
0
2

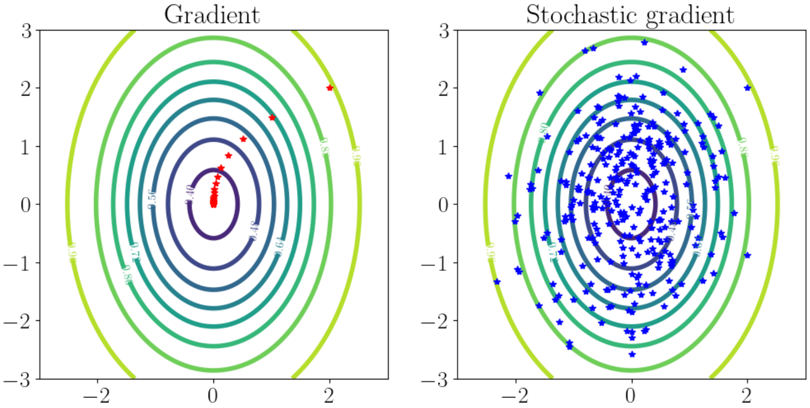

Gradient
Stochastic gradient
3
2
1
0
−1
−2
−3
−2
0
2

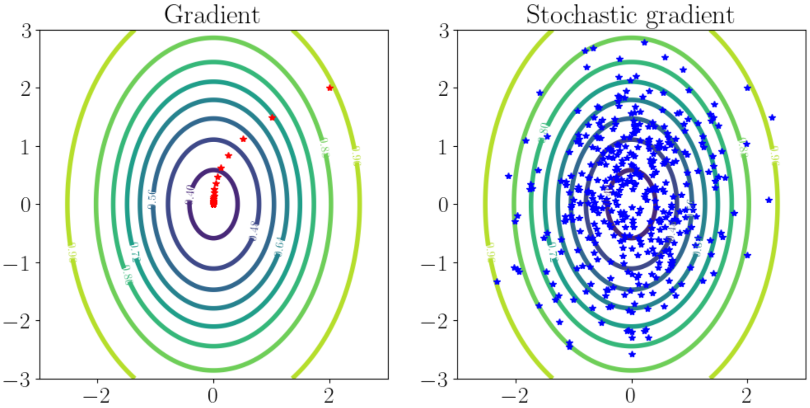

Gradient
Stochastic gradient
3
2
1
0
−1
−2
−3
−2
0
2
0.40
0.48
0.56
0.64
0.72
0.80
0.88
0.96

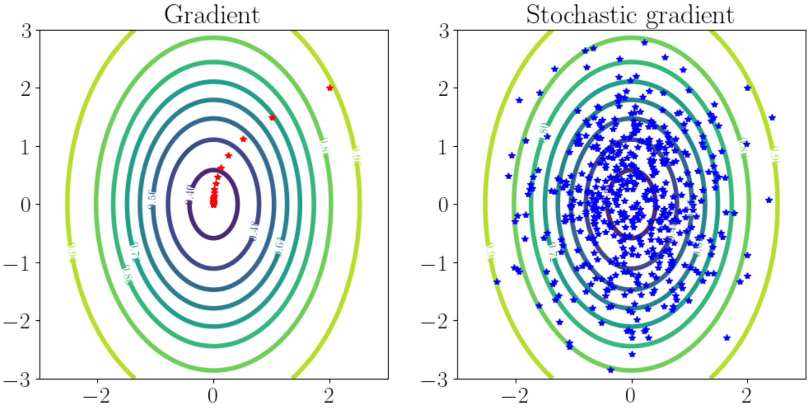

Gradient
Stochastic gradient
3
2
1
0
−1
−2
−3
−2
0
2
0.40
0.48
0.56
0.64
0.72
0.80
0.88
0.96


Gradient
Stochastic gradient
3
2
1
0
−1
−2
−3
−2
0
2

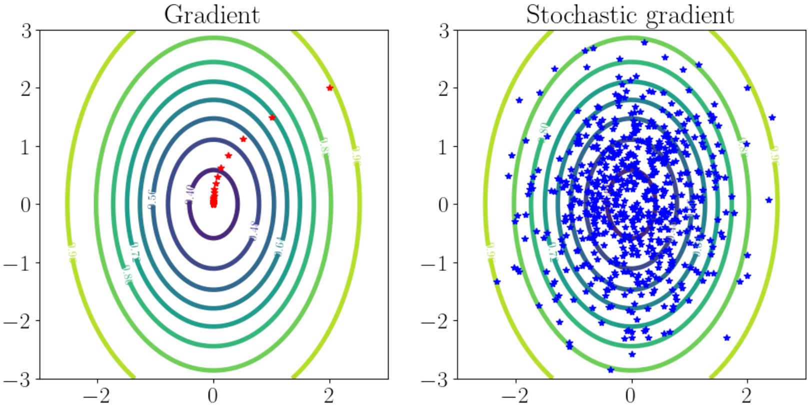

Gradient
Stochastic gradient
3
2
1
0
−1
−2
−3
−2
0
2

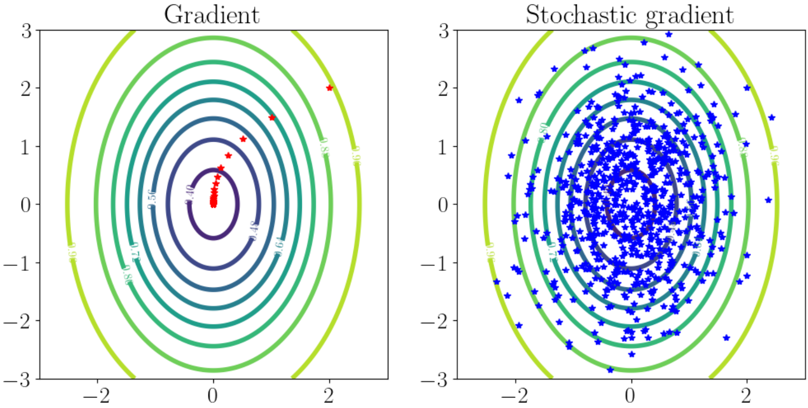

Gradient
Stochastic gradient

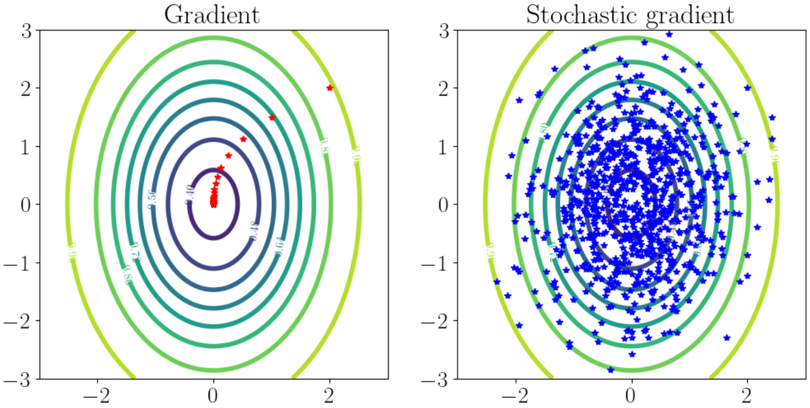

Gradient
Stochastic gradient
3
2
1
0
−1
−2
−3
−2
0
2
0.40
0.48
0.56
0.64
0.72
0.80
0.88
0.96



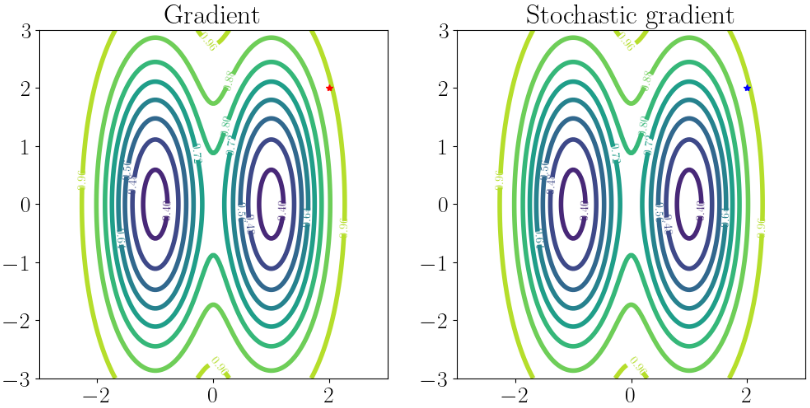
Gradient
Stochastic gradient
3
2
1
0
−1
−2
−3
−2
0
2


Gradient
Stochastic gradient

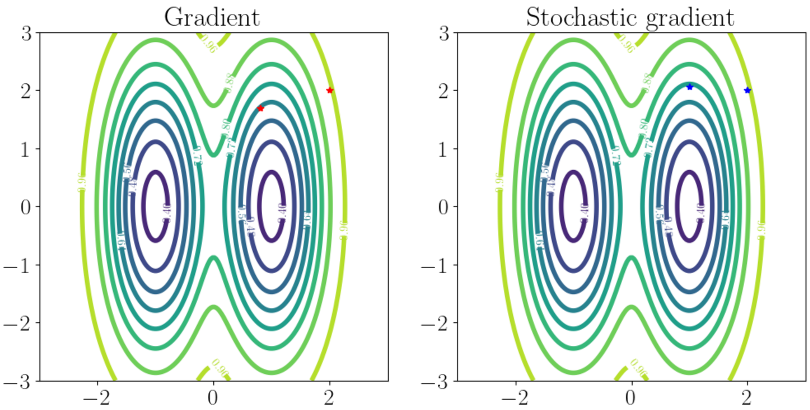

Gradient
Stochastic gradient



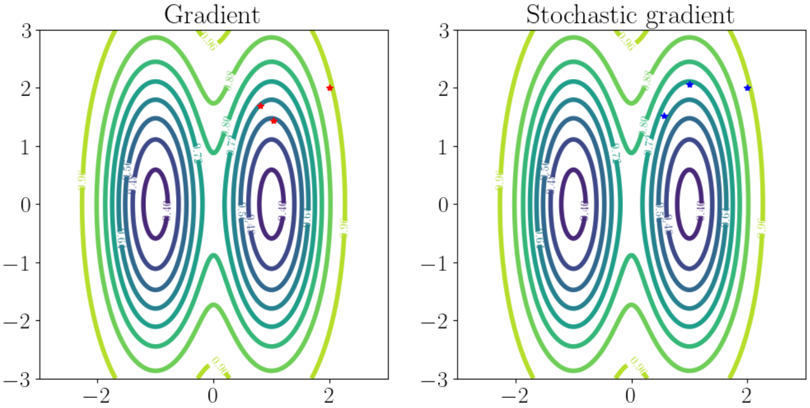
Gradient
Stochastic gradient

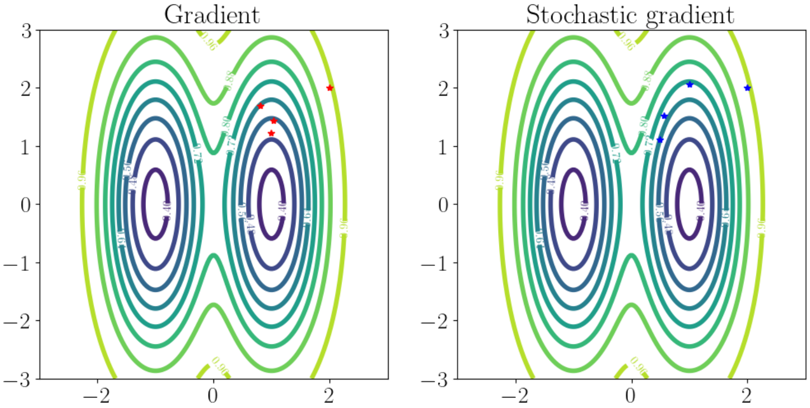


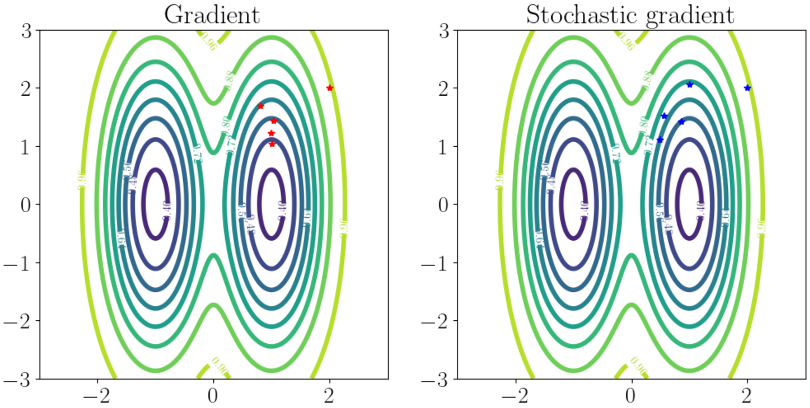
Gradient
Stochastic gradient
3
2
1
0
−1
−2
−3
−2
0
2



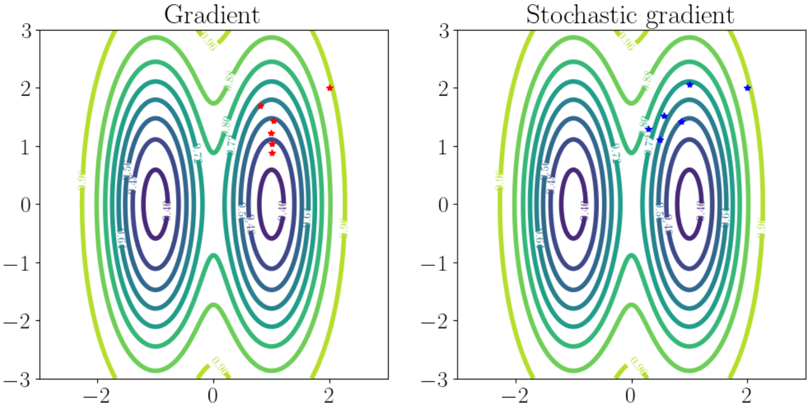
Gradient
Stochastic gradient


Gradient
Stochastic gradient
3
2
1
0
−1
−2
−3
−2
0
2
0.96
0.88
0.80
0.72
0.56
0.48
0.40
0.64

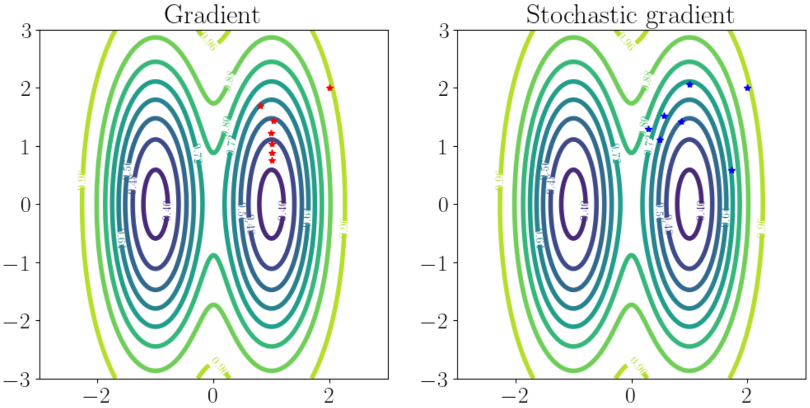

Gradient
Stochastic gradient
3
2
1
0
−1
−2
−3
−2
0
2



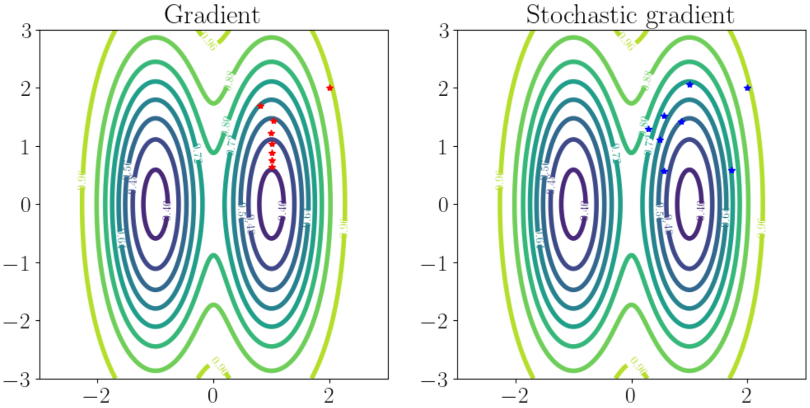
Gradient
Stochastic gradient
3
2
1
0
−1
−2
−3
−2
0
2
0.96
0.88
0.80
0.72
0.56
0.48
0.40
0.64

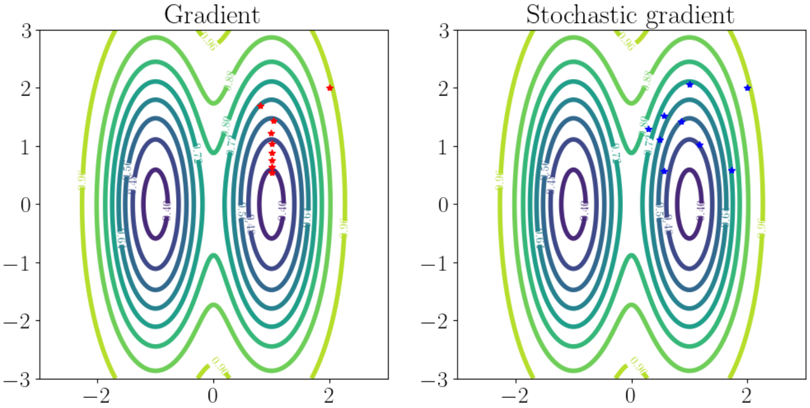


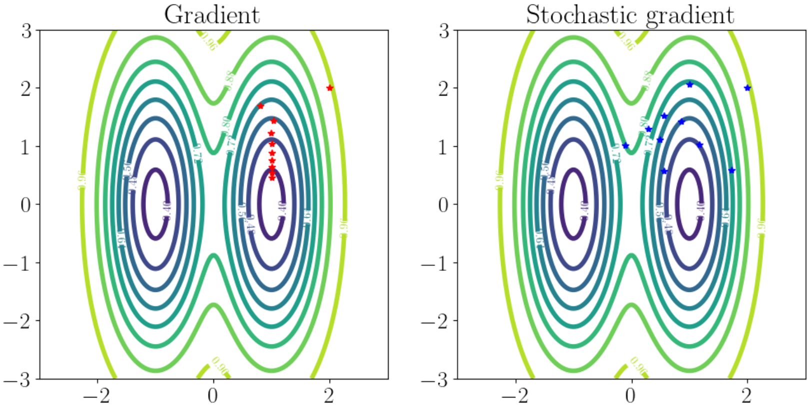
Gradient
Stochastic gradient
3
2
1
0
−1
−2
−3
−2
0
2
0.96
0.88
0.80
0.72
0.64
0.56
0.48
0.40

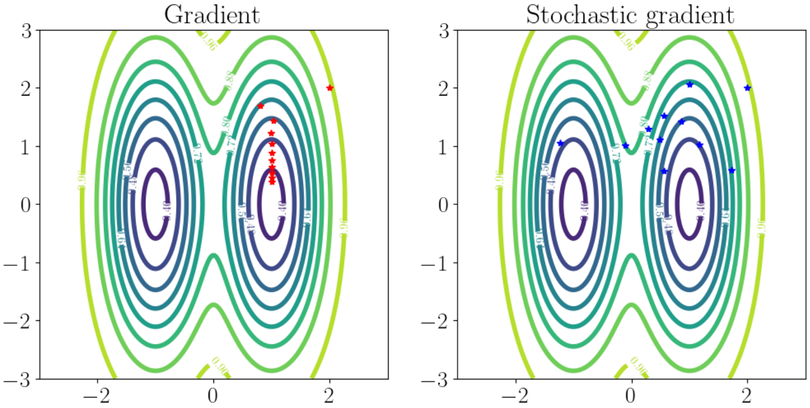


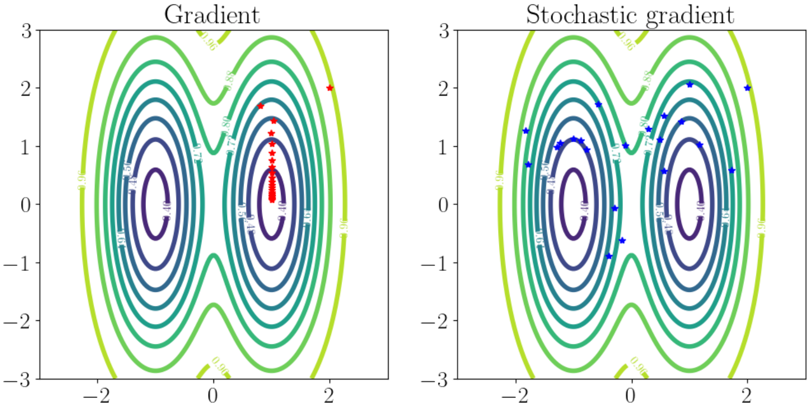
Gradient
Stochastic gradient
3
2
1
0
−1
−2
−3
−2
0
2

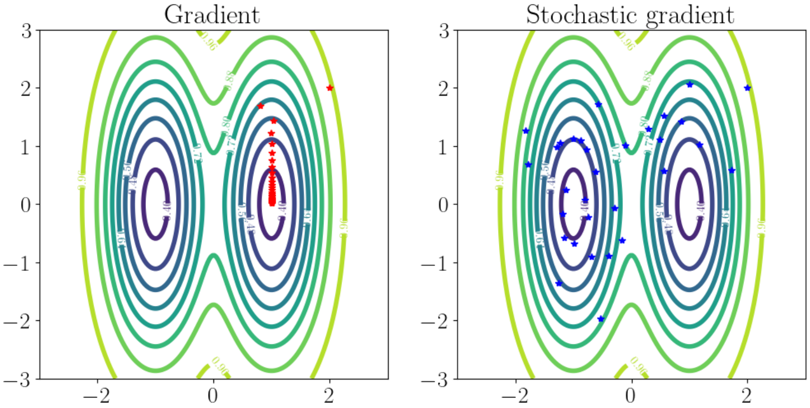

Gradient
Stochastic gradient
3
2
1
0
−1
−2
−3
−2
0
2


Gradient
Stochastic gradient
3
2
1
0
−1
−2
−3
−2
0
2

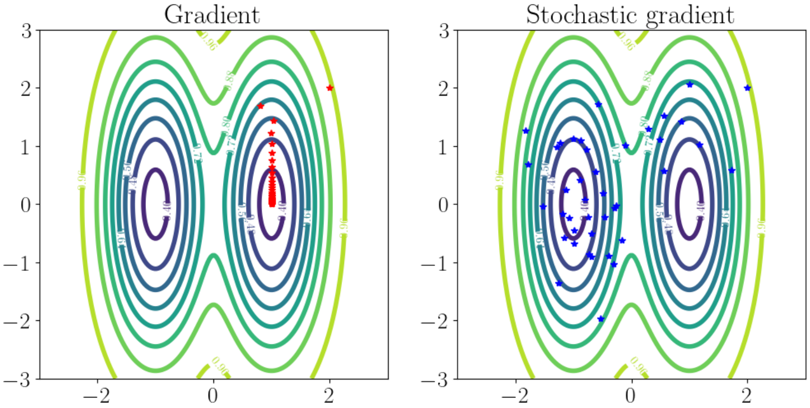

Gradient
Stochastic gradient
3
2
1
0
−1
−2
−3
−2
0
2

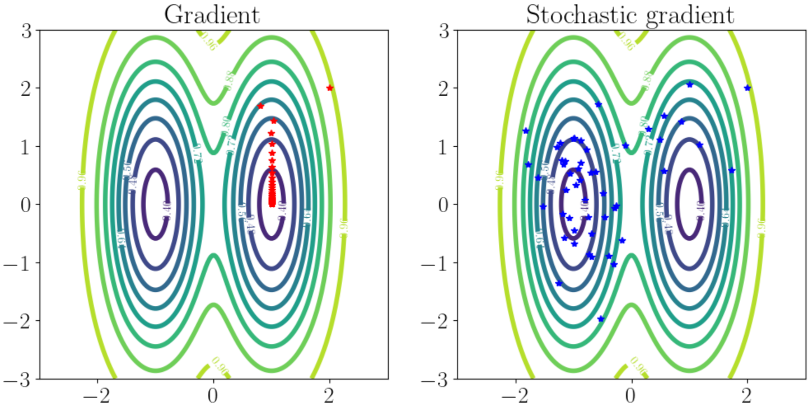

Gradient
Stochastic gradient
3
2
1
0
−1
−2
−3
−2
0
2


Gradient
Stochastic gradient
3
2
1
0
−1
−2
−3
−2
0
2
0.96
0.88
0.80
0.72
0.40

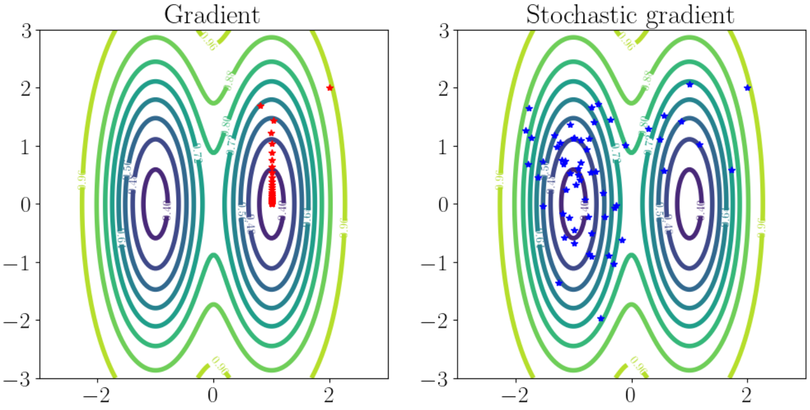


Gradient
Stochastic gradient
3
2
1
0
−1
−2
−3
−2
0
2

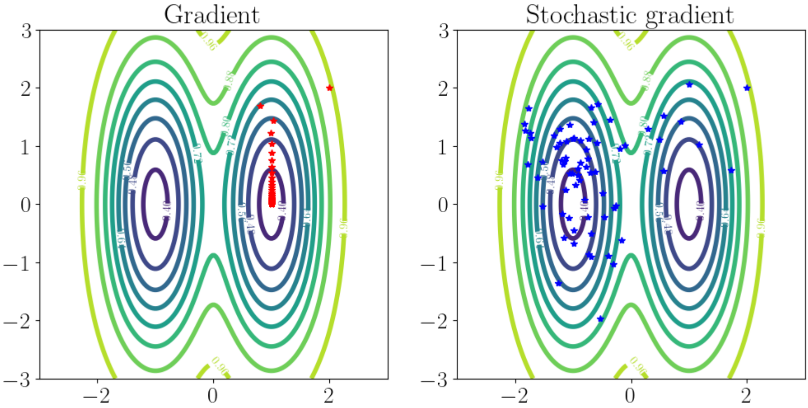

Gradient
Stochastic gradient
3
2
1
0
−1
−2
−3
−2
0
2



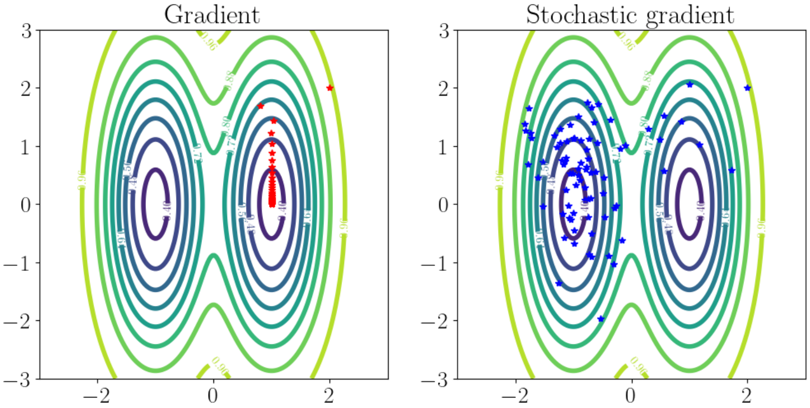
Gradient
Stochastic gradient



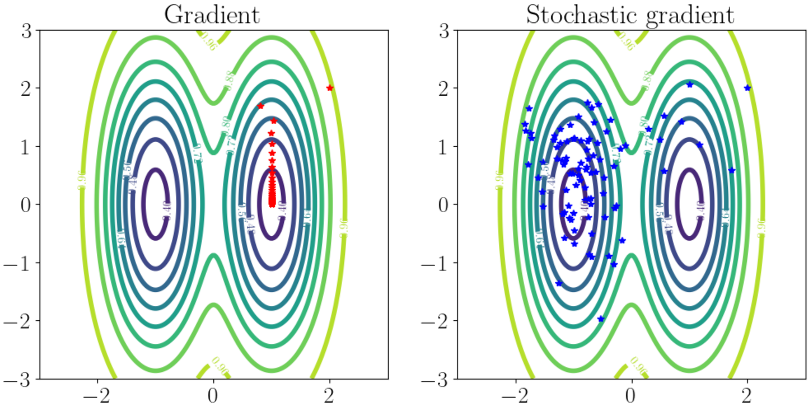
Gradient
Stochastic gradient

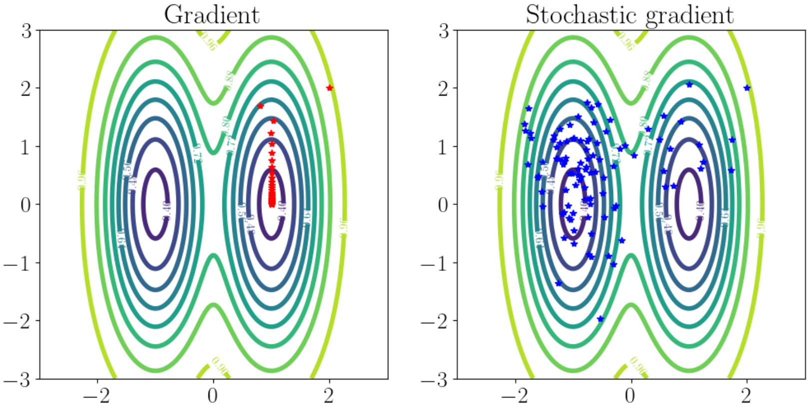


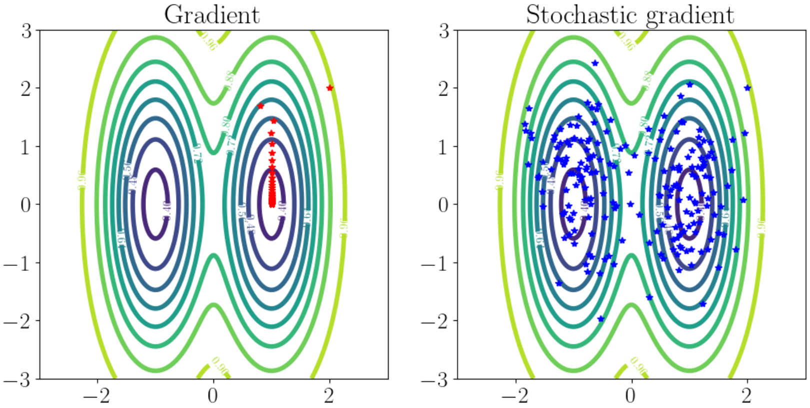
Gradient
Stochastic gradient
3
2
1
0
−1
−2
−3
−2
0
2



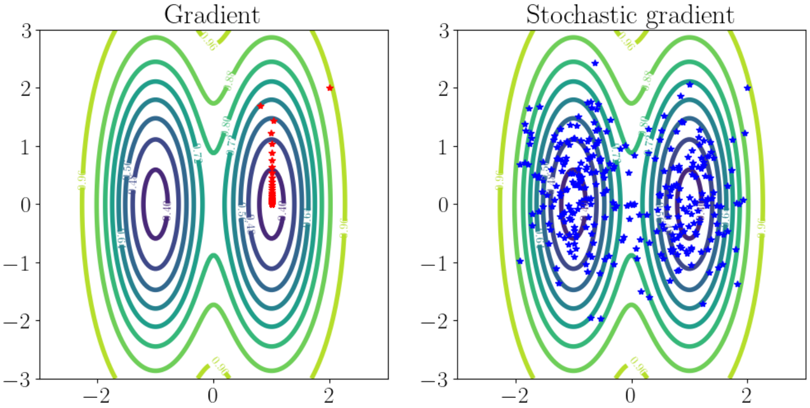
Gradient
Stochastic gradient
3
2
1
0
−1
−2
−3
−2
0
2

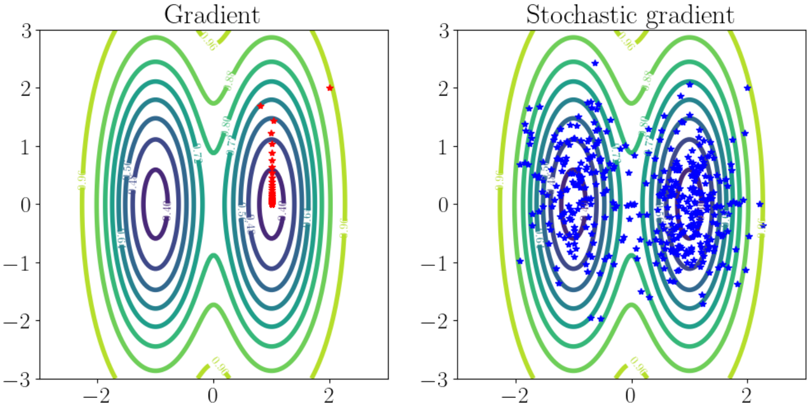

Gradient
Stochastic gradient
3
2
1
0
−1
−2
−3
−2
0
2


Gradient
Stochastic gradient
3
2
1
0
−1
−2
−3
−2
0
2

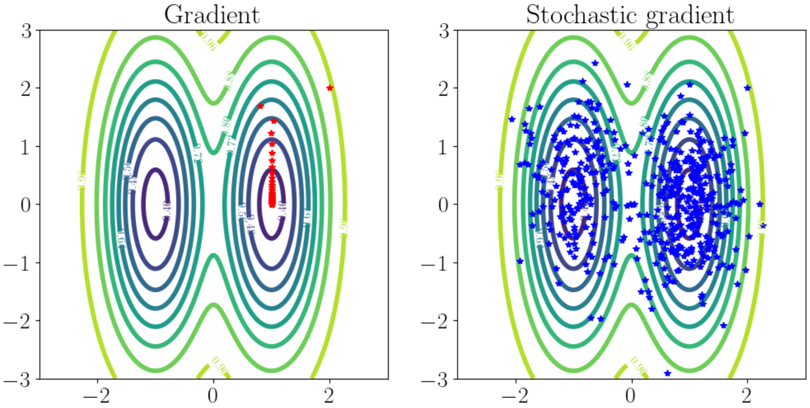


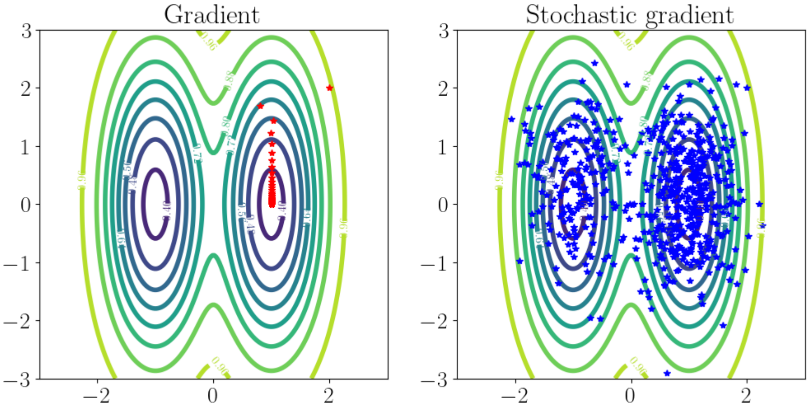
Gradient
Stochastic gradient

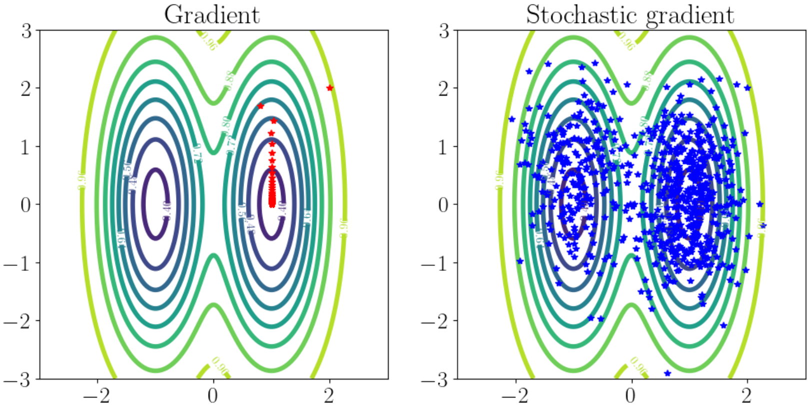

Gradient
Stochastic gradient
3
2
1
0
−1
−2
−3
−2
0
2

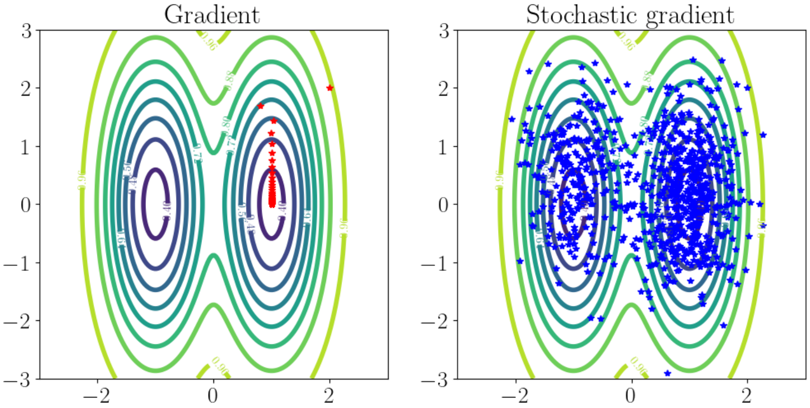


Gradient
Stochastic gradient

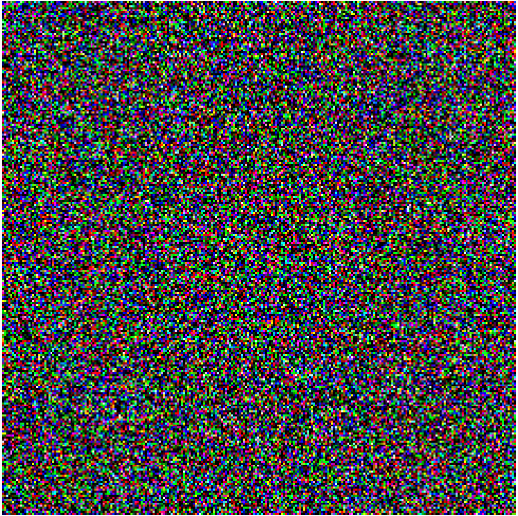

Gradient
Stochastic gradient
3
2
1
0
−1
−2
−3
−2
0
2


Gradient
Stochastic gradient
3
2
1
0
−1
−2
−3
−2
0
2

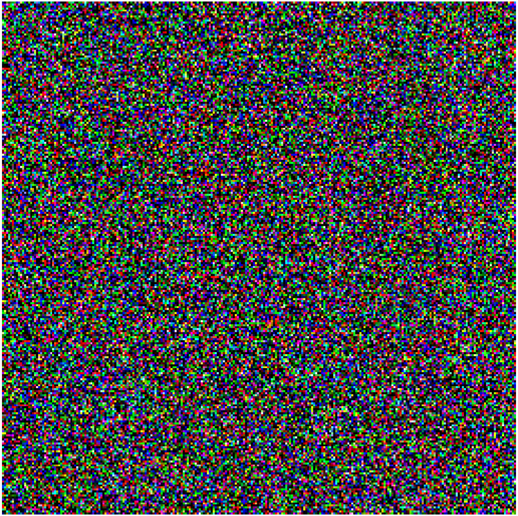

Gradient
Stochastic gradient
3
2
1
0
−1
−2
−3
−2
0
2


Gradient
Stochastic gradient
3
2
1
0
−1
−2
−3
−2
0
2

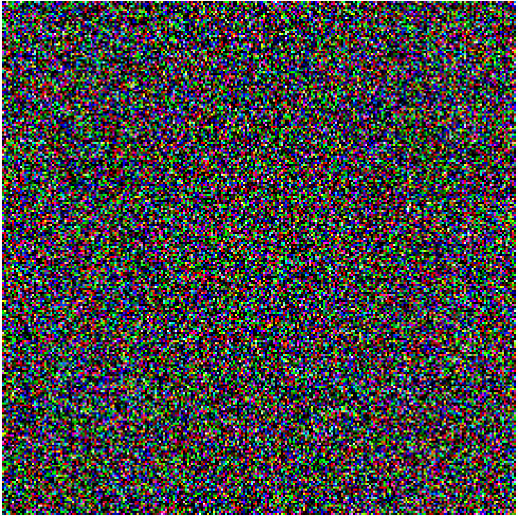
iteration 1



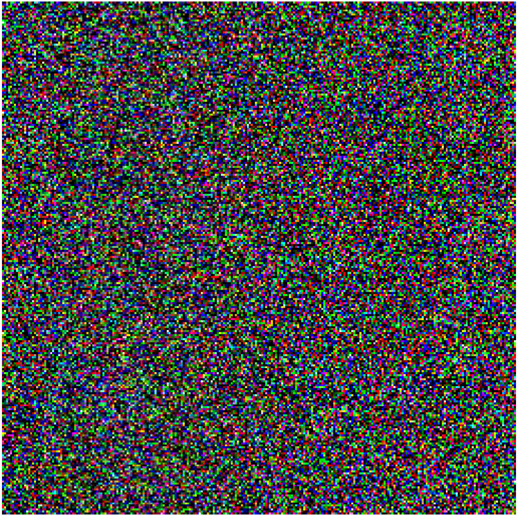
iteration 2



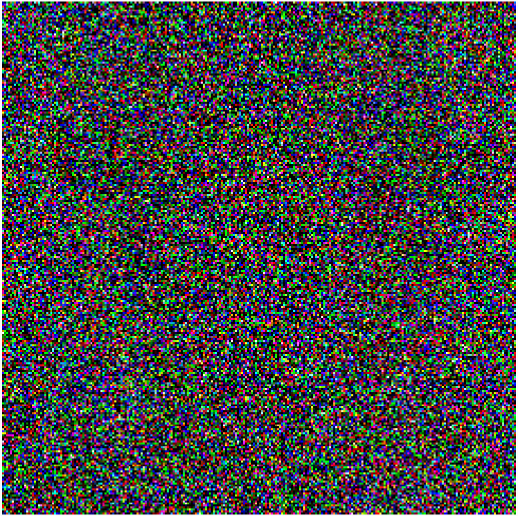
iteration 3



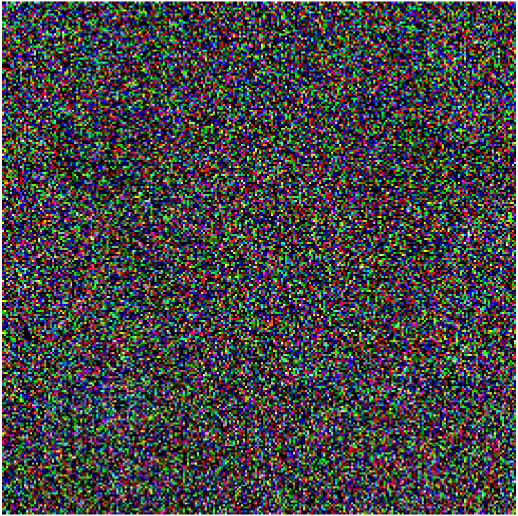
iteration 4



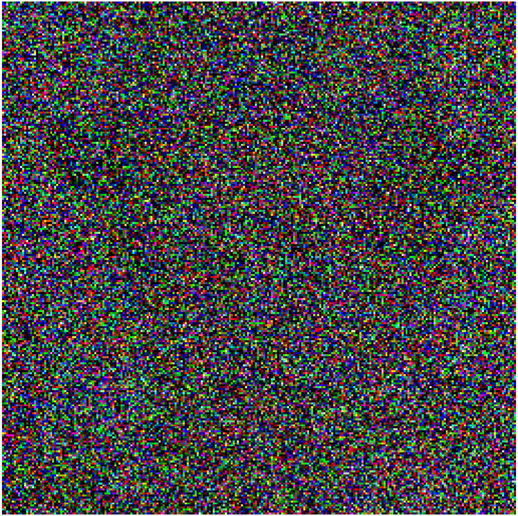
iteration 5



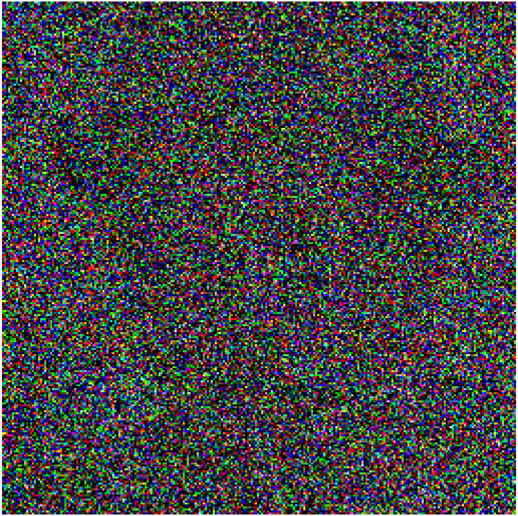
iteration 6



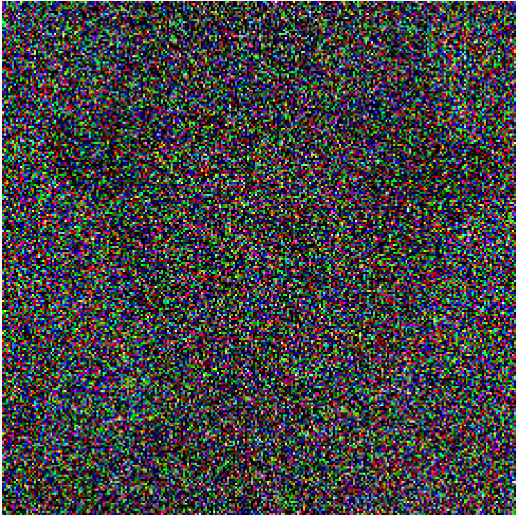
iteration 7



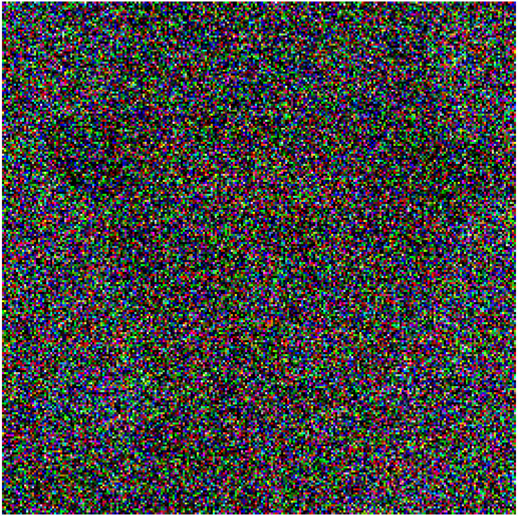
iteration 8



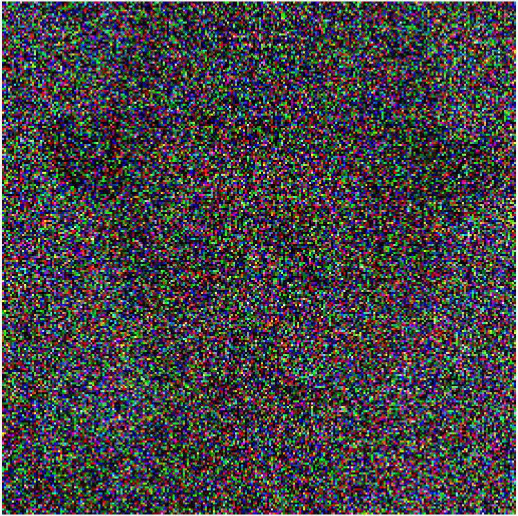
iteration 9



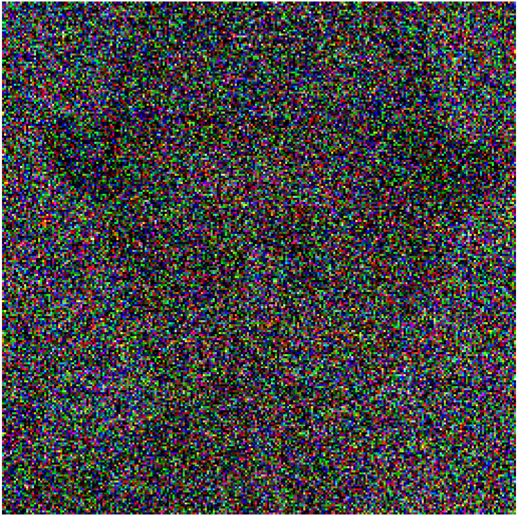
iteration 10



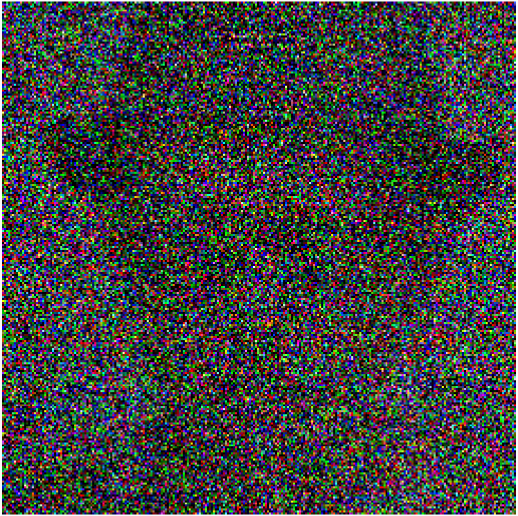
iteration 11



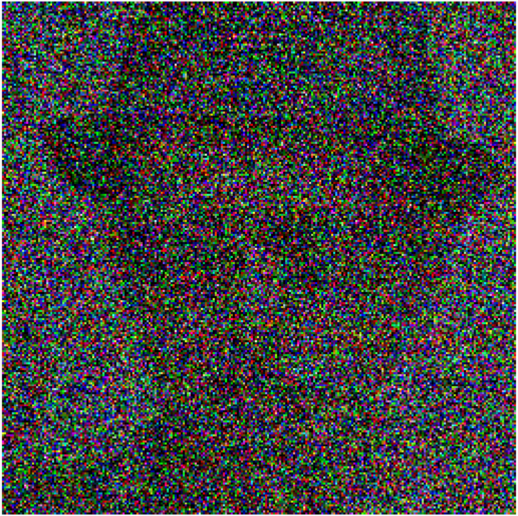
iteration 12



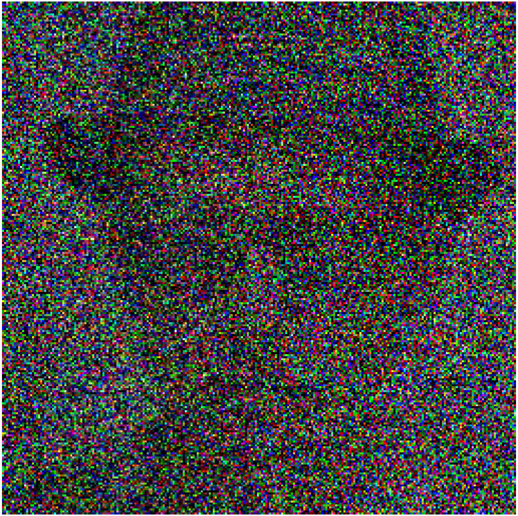
iteration 13



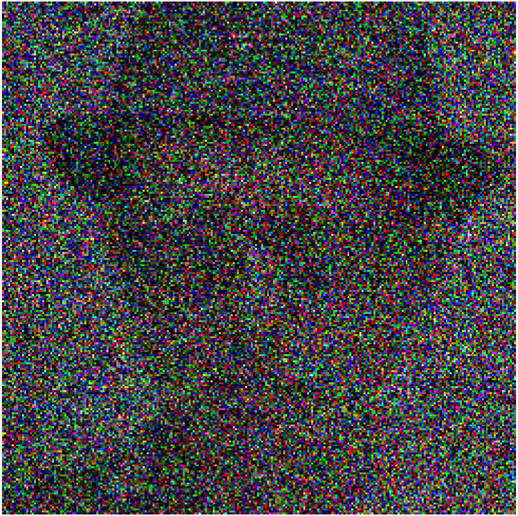
iteration 14



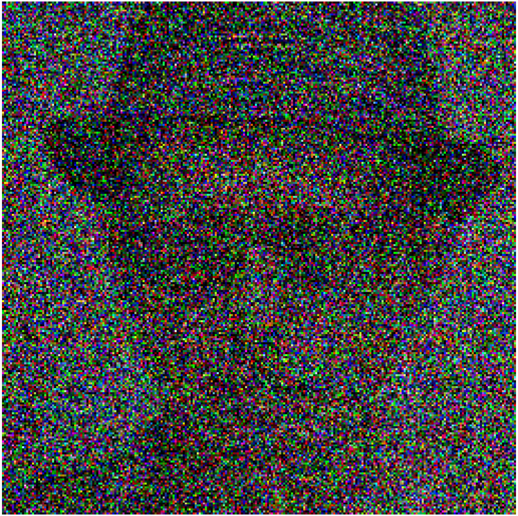
iteration 15



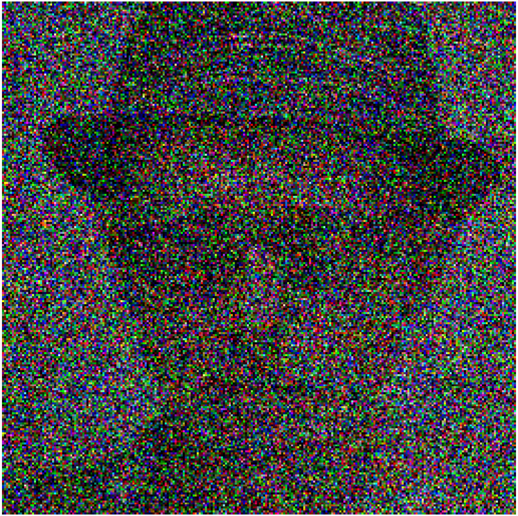
iteration 16



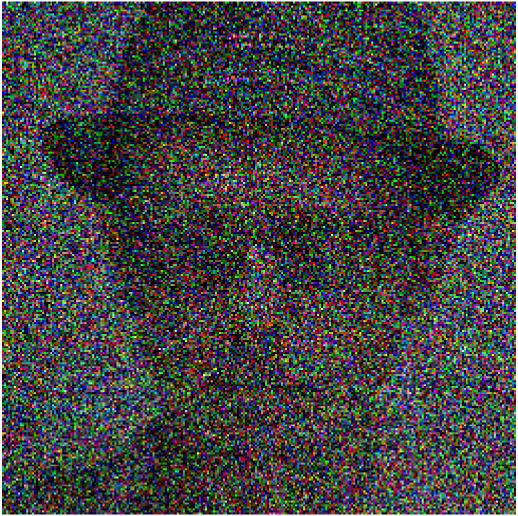
iteration 17



iteration 18



iteration 19



iteration 20

iteration 21

iteration 22

iteration 23

iteration 24

iteration 25

iteration 26

iteration 27

iteration 28

iteration 29

iteration 30

iteration 31

iteration 32

iteration 33

iteration 34

iteration 35

iteration 36

iteration 37

iteration 38

iteration 39

iteration 40

iteration 41

iteration 42

iteration 43

iteration 44

iteration 45

iteration 46

iteration 47

iteration 48

iteration 49

Here are the results!

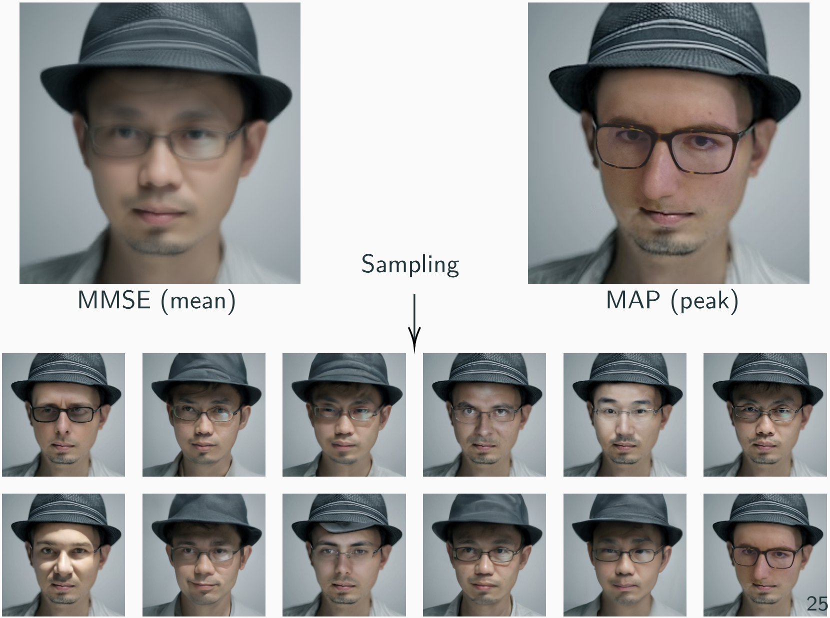
MAP (peak)

Here are the results!

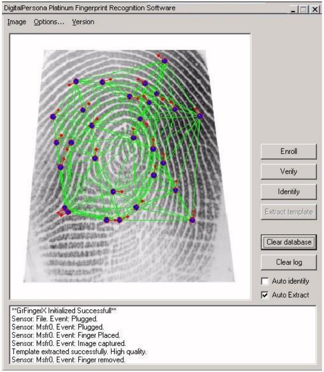
MAP (peak)

Sampling

Here are the results!

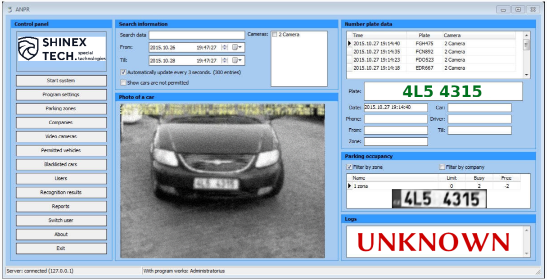


Here are the results!



A random IMT agent.

Maybe the guilty ?



An unknown car ....

The investigation is continuing ...

Thanks for your attention

Thanks for your attention