Likelihood and Bayesian inference

Michael Matschiner

NHM, University of Oslo

Likelihood and Bayesian inference

Michael Matschiner
NHM, University of Oslo

Likelihood and Bayesian inference

Michael Matschiner
NHM, University of Oslo

Inference methods

1. Maximum likelihood

Inference methods

1. Maximum likelihood
2. Bayesian inference

Inference methods

1. Maximum likelihood

2. Bayesian inference

3. Approximate Bayesian computation

Inference methods

1. Maximum likelihood
2. Bayesian inference
3. Approximate Bayesian computation
4. Machine learning

Inference methods

1. Maximum likelihood
2. Bayesian inference
3. Approximate Bayesian computation
4. Machine learning

Vitor Sousa

Inference methods

1. Maximum likelihood
2. Bayesian inference
3. Approximate Bayesian computation
4. Machine learning

Vitor Sousa

Flora Jay,
Matteo Fumagalli

Likelihood

Probability

Probability

Probability

$$P = {}^{1}/_{6}$$

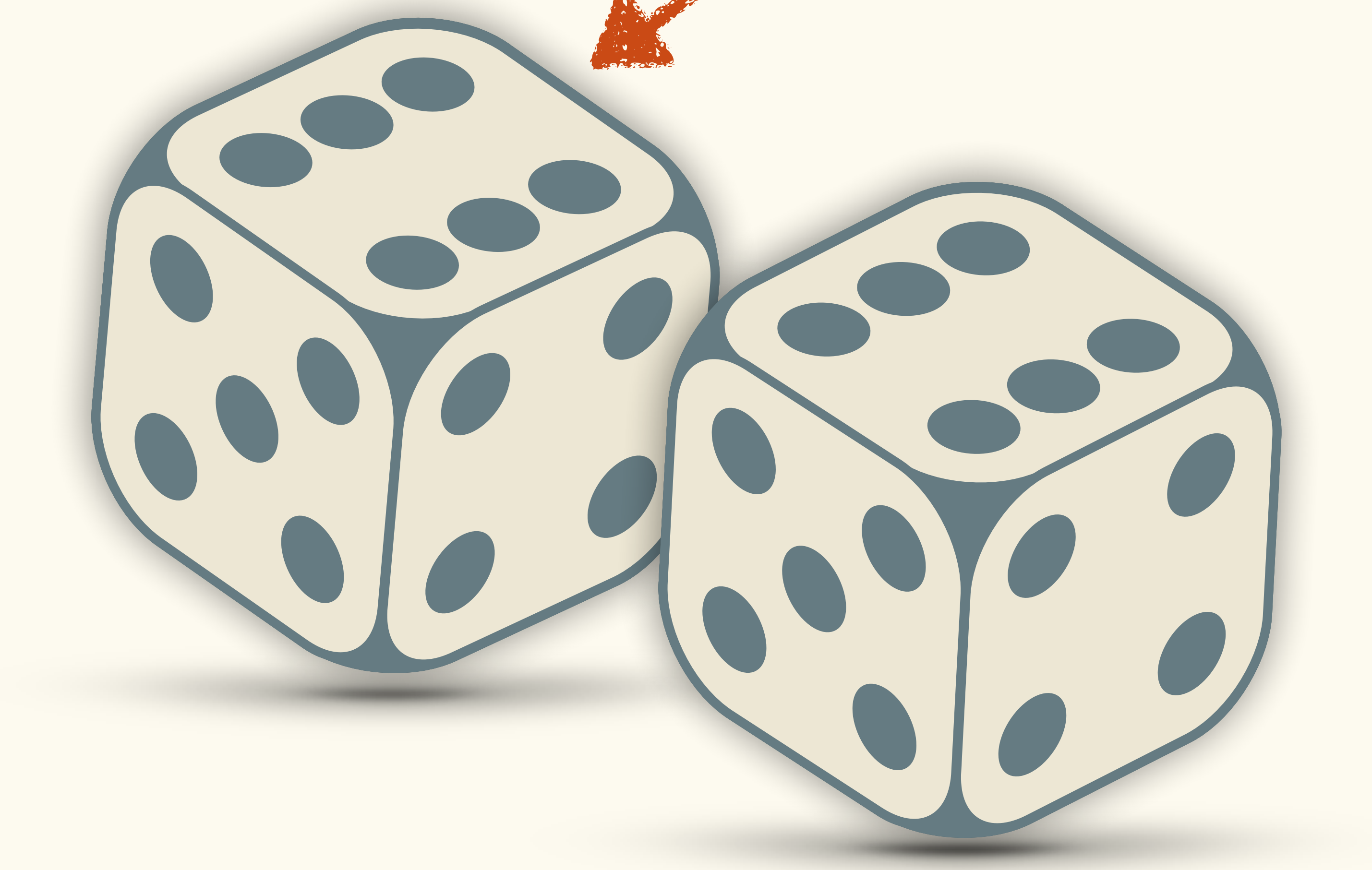
Probability

$P(\text{⚅}) = {}^{1}/_{6}$

Probability

Probability

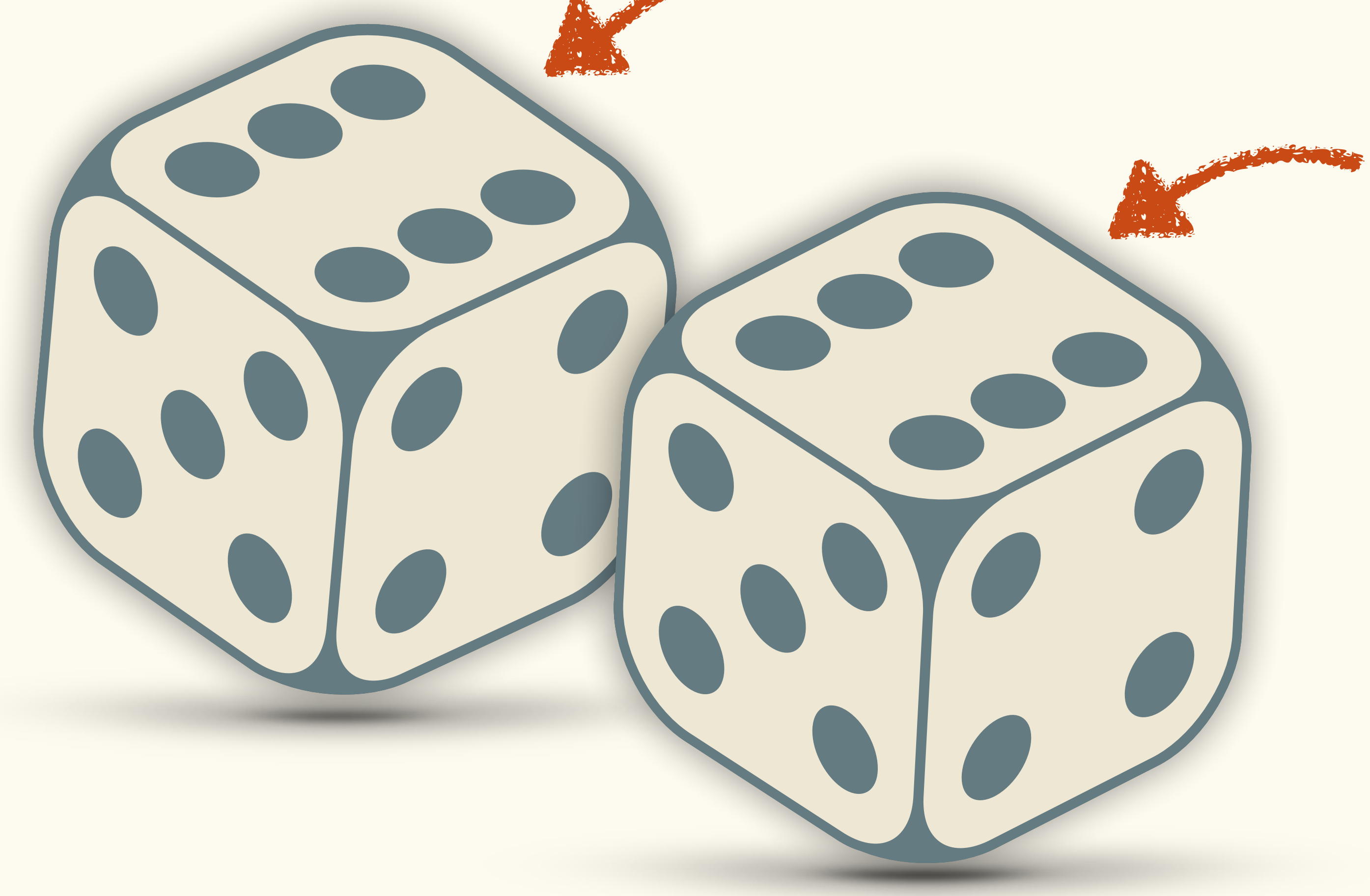
$P(⚅) = {}^{1}/_{6}$



Probability

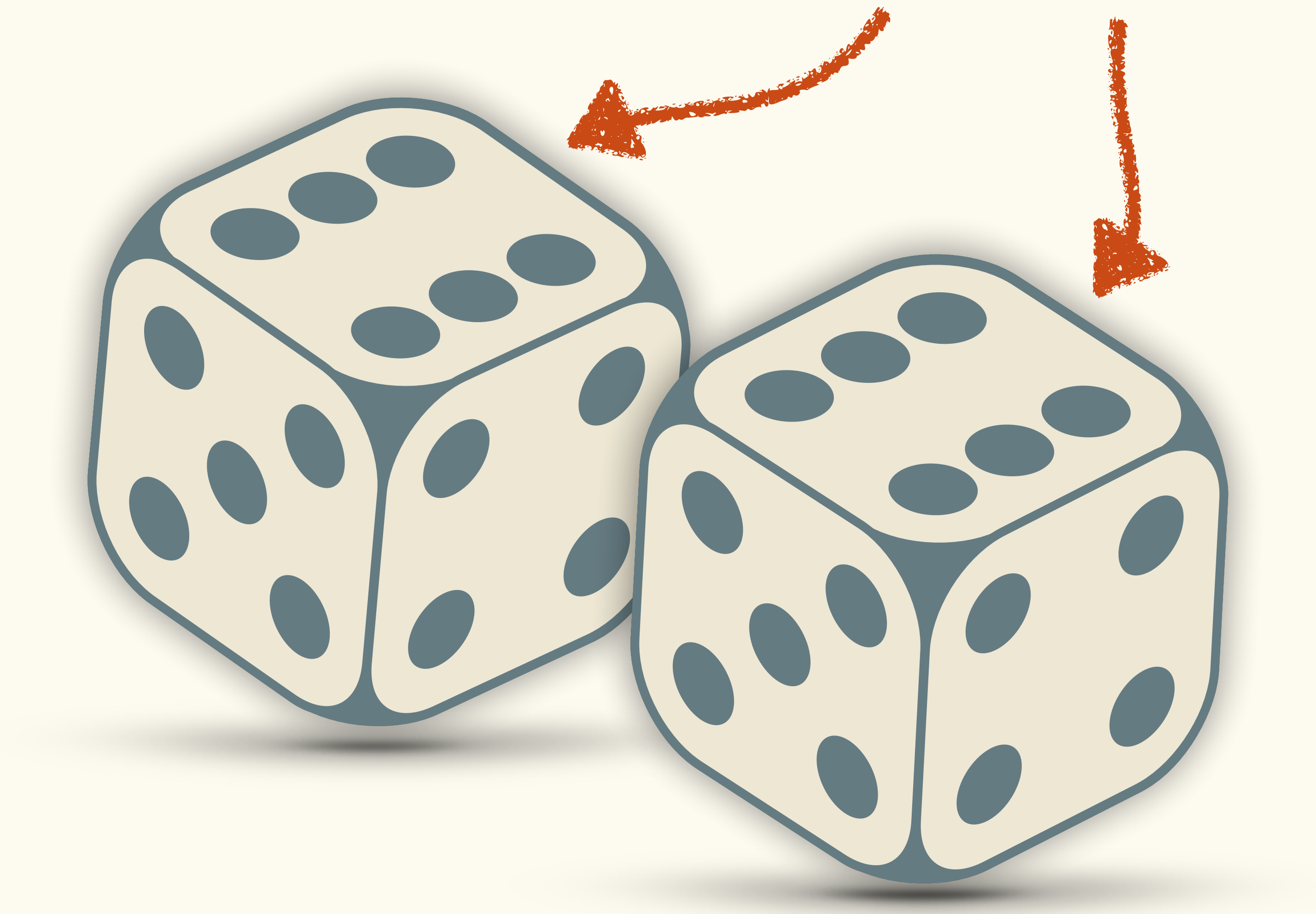
$P(⚅) = {}^{1}/_{6}$

$P(⚅) = {}^{1}/_{6}$



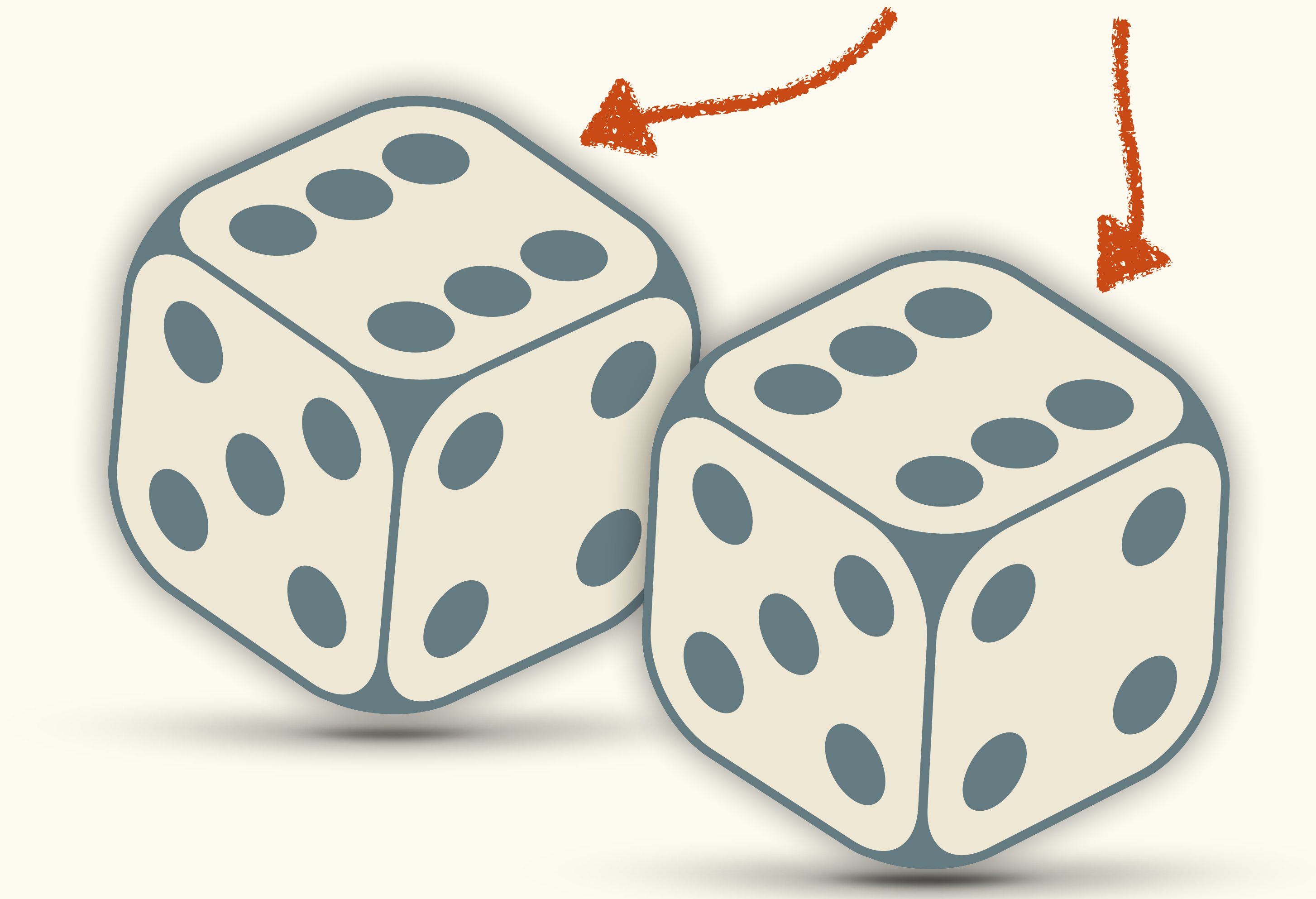
Probability

P(⚅ & ⚅) = $^1/_6 \times {}^1/_6$



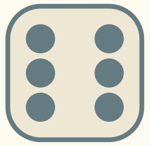
Probability

$$P(⚅ \,\&\, ⚅) = {}^{1}/_{6} \times {}^{1}/_{6} = {}^{1}/_{36}$$

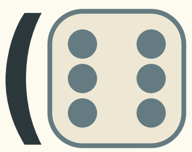


Probability

Probability

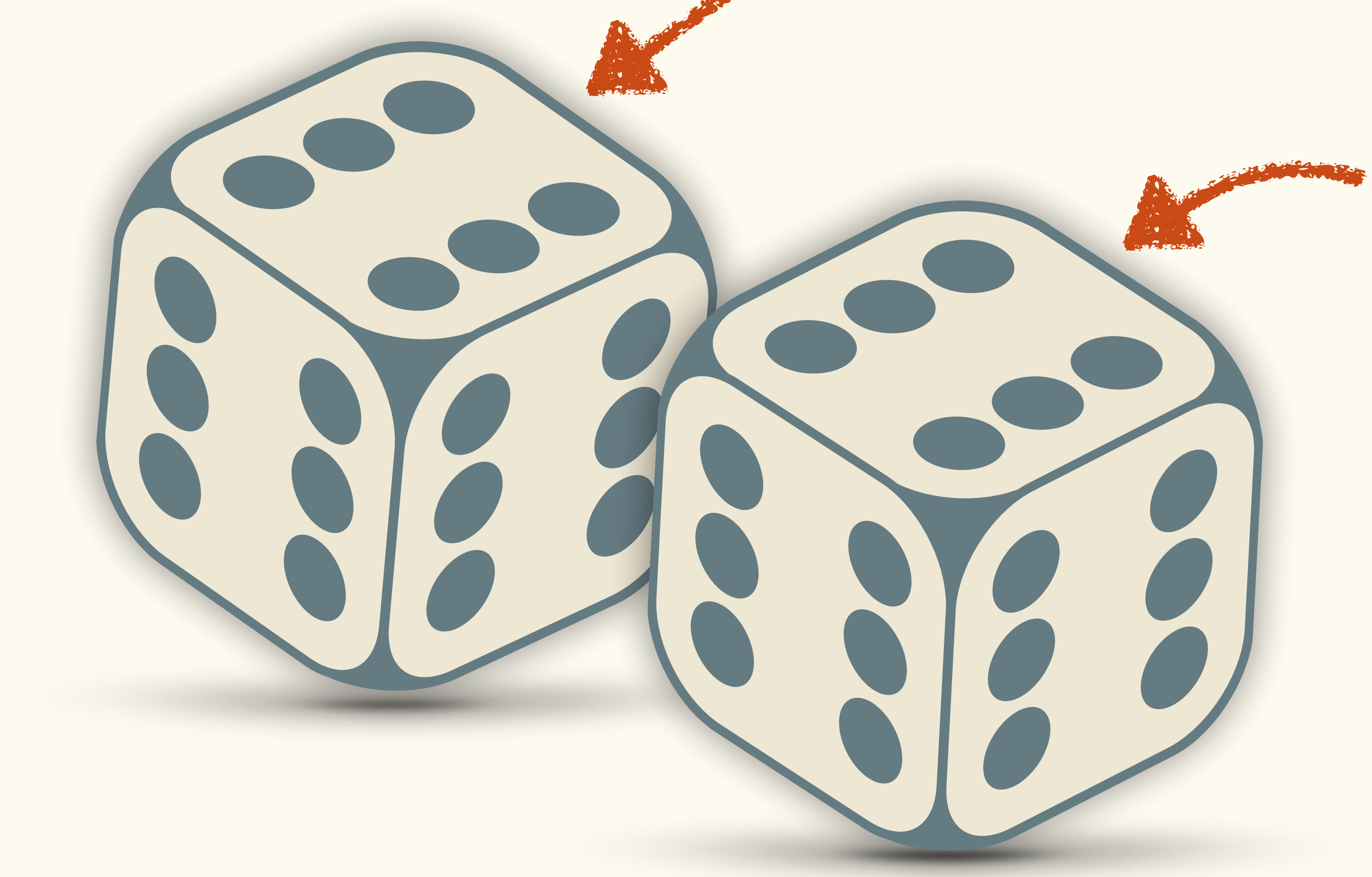
P(**) = ?**

Probability

P() = 1

Probability

Probability

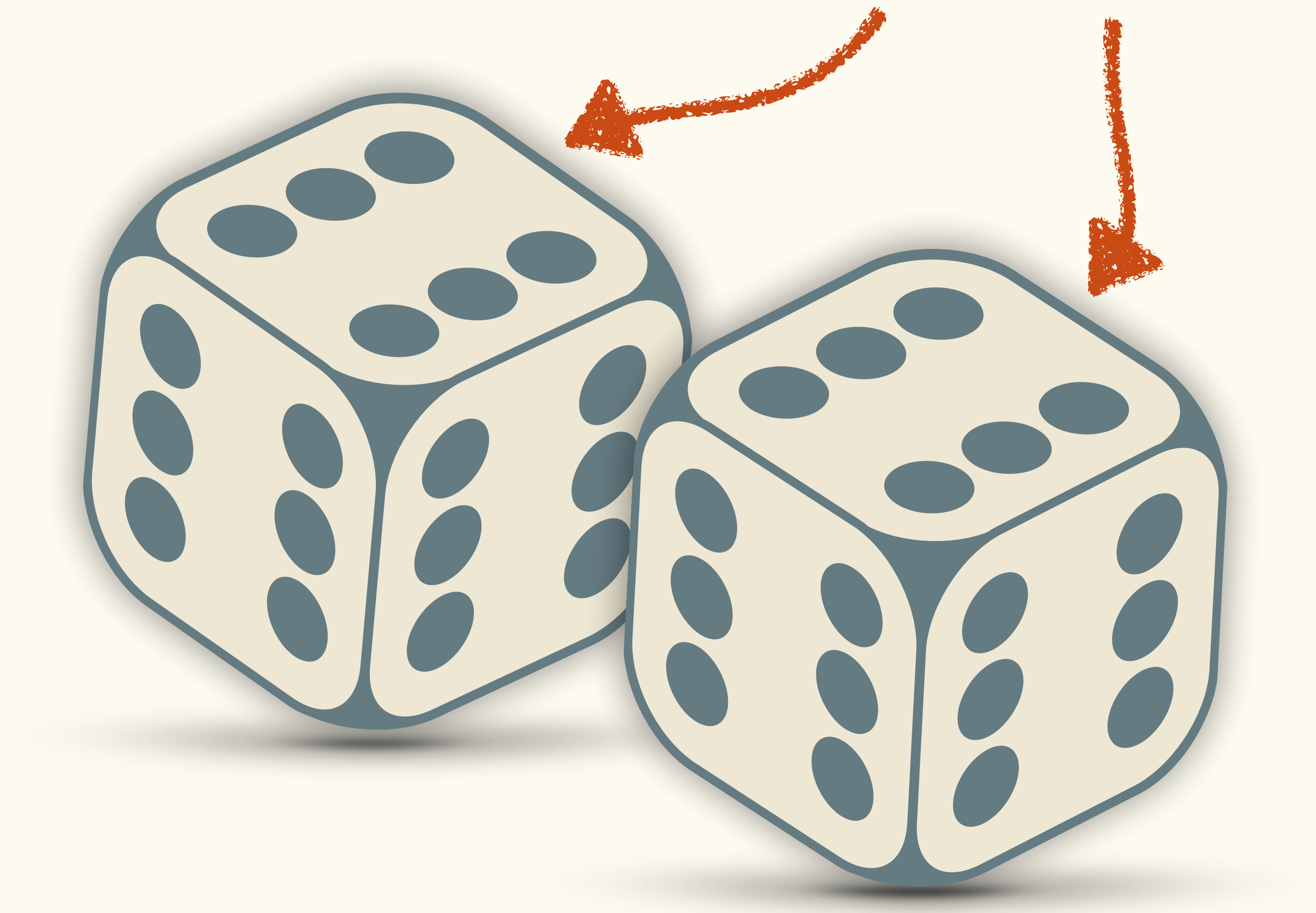


$P($⚅$) = 1$

$P($⚅$) = 1$

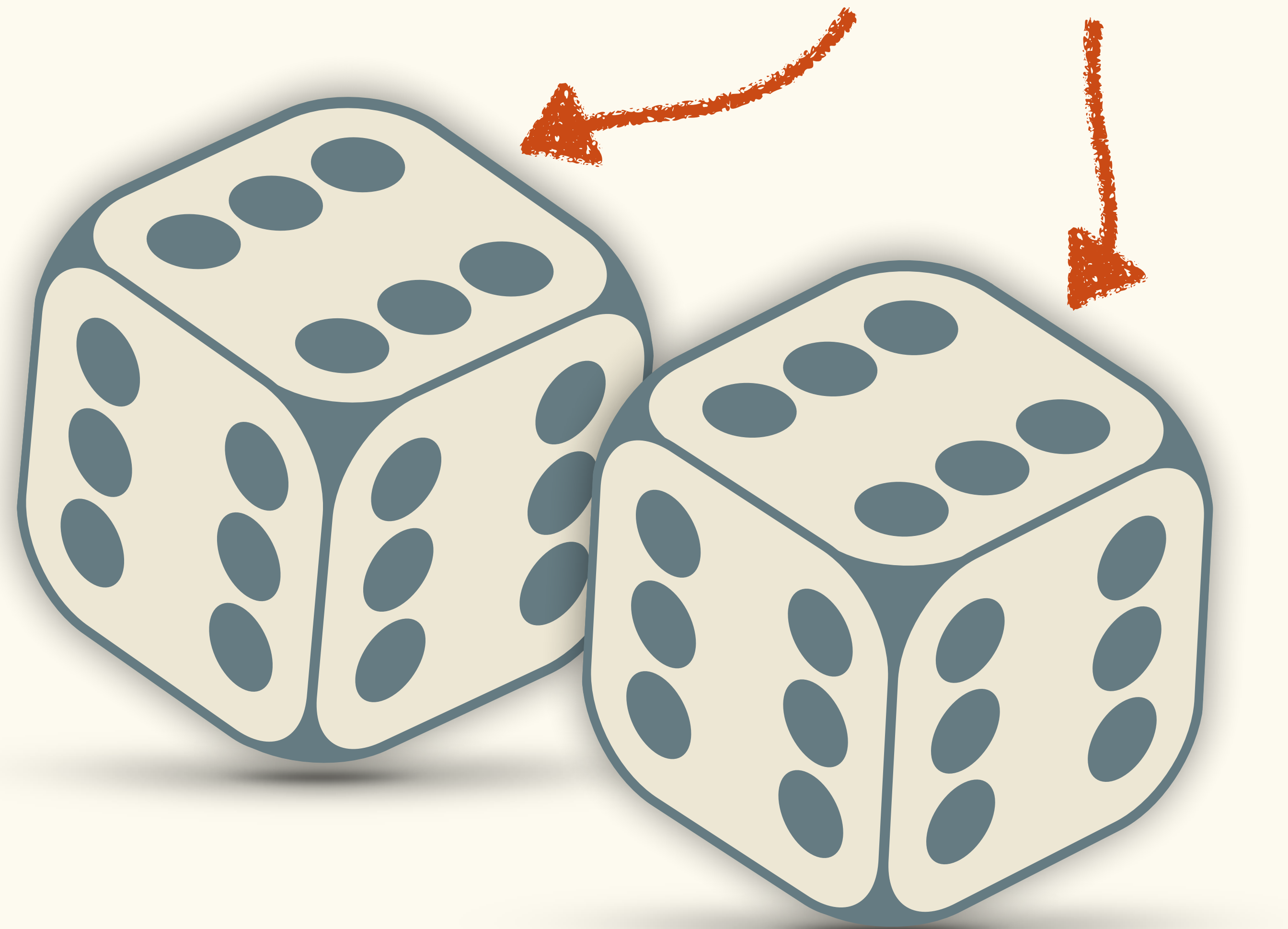
Probability

P(⚅ & ⚅) = 1 × 1



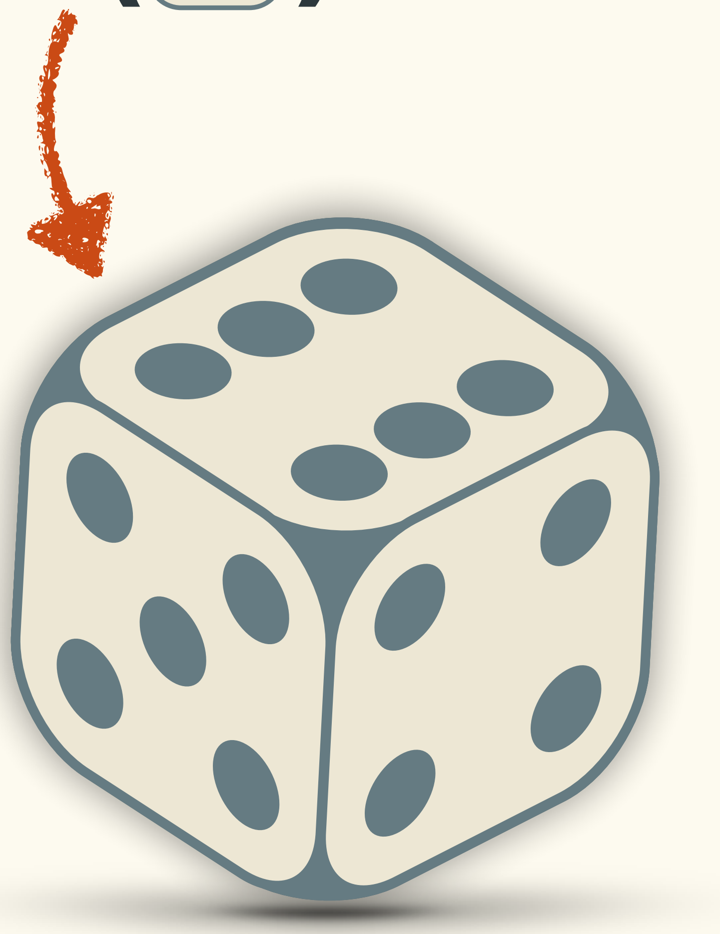
Probability

P(⚅ & ⚅) = 1 × 1 = 1

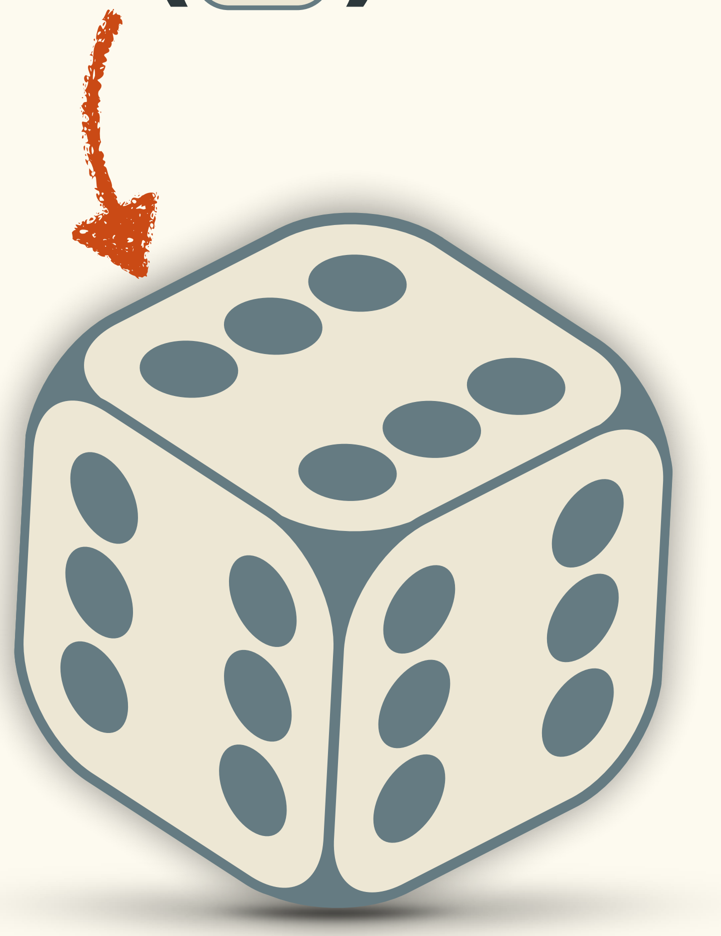


Probability

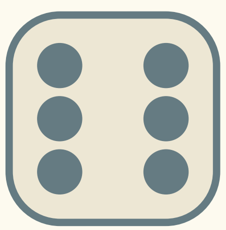
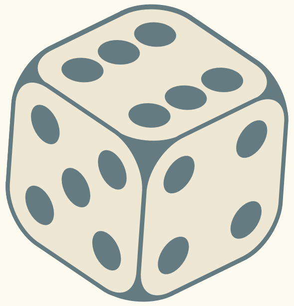
$P(⚅) = {}^{1}/_{6}$



$P(⚅) = 1$



Probability

P( | ) = 1/6

Observation

Probability

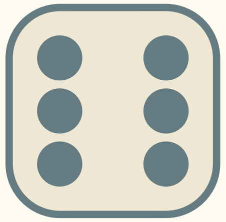
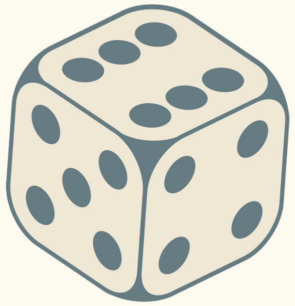
P(⚅ | 🎲) = $1/6$

Result

Probability

“under the assumption of”

“given”

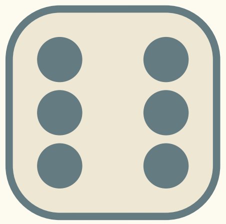
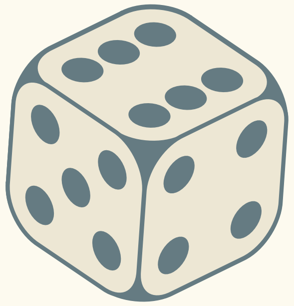
P( | ) = $^1/_6$

Result

Probability

“under the assumption of”

“given”

$P($  $|$  $) = {}^{1}/_{6}$

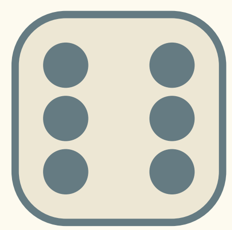
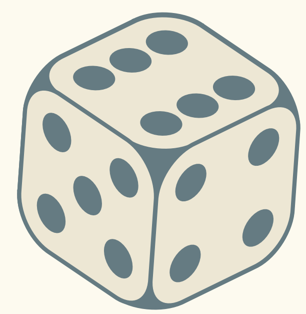
Result

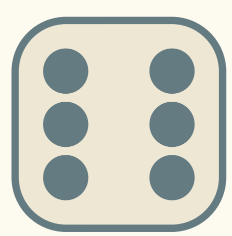
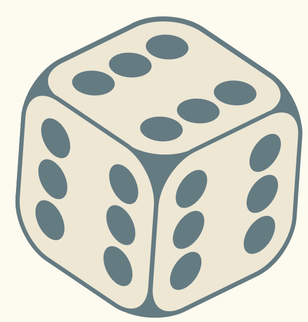
Model

(fair dice)

Probability

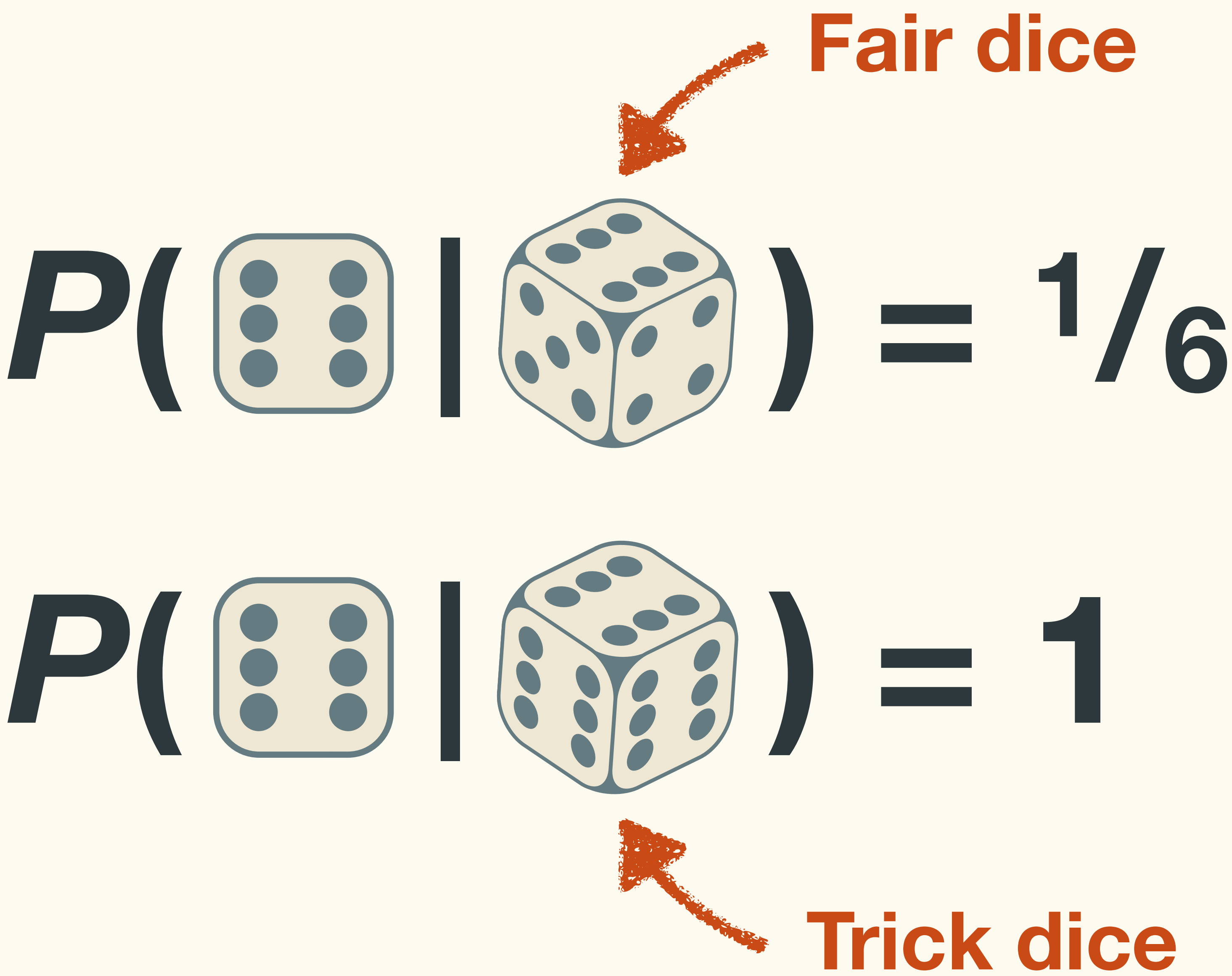
Fair dice

$P($  $|$  $) = 1/6$

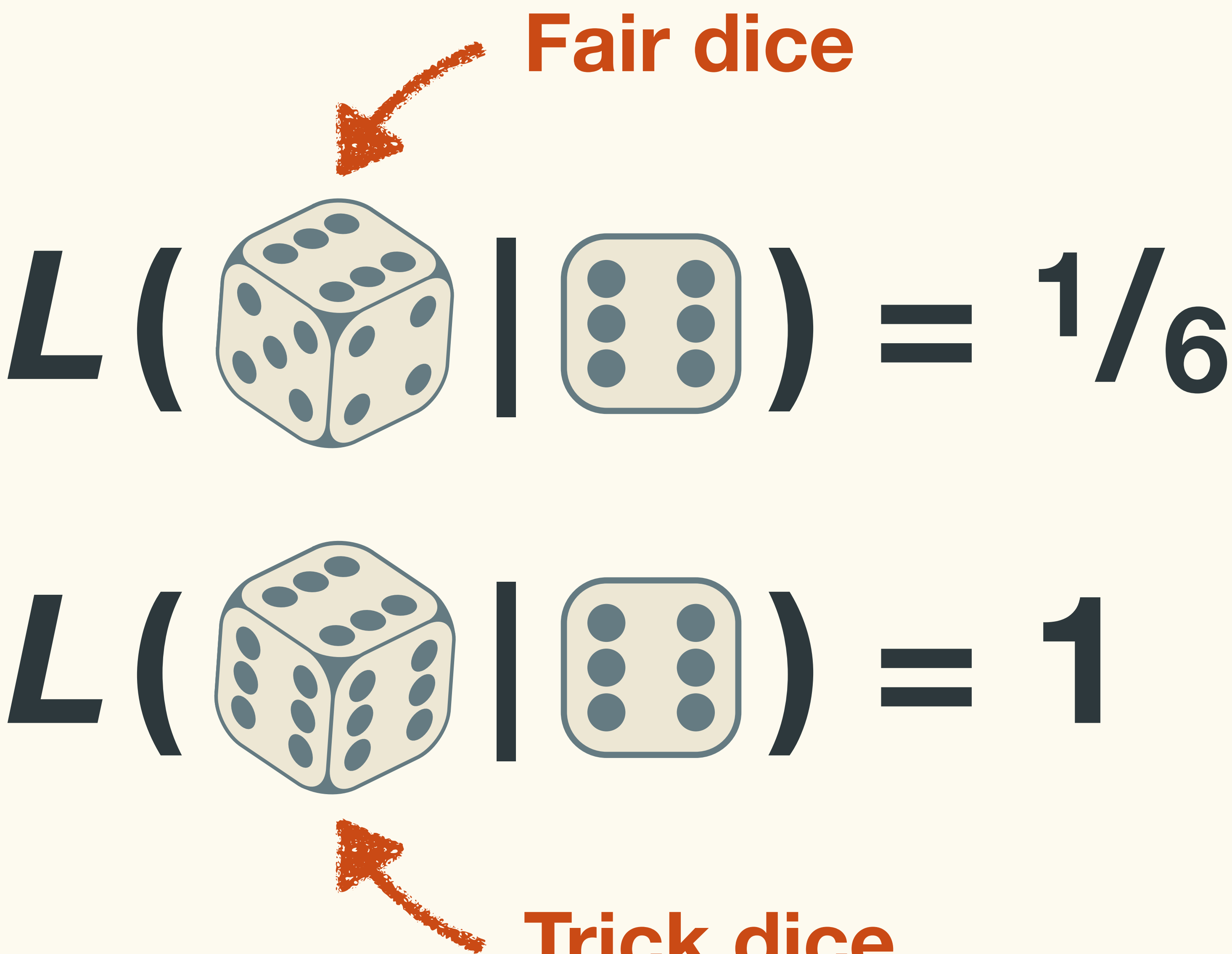
$P($  $|$  $) = 1$

Trick dice

Probability



Likelihood



Likelihood

L(🎲 | ⚅) = P(⚅ | 🎲)

Probability

Probability

$P(A) = ?$

Probability

$P(A) = 1/4$

$P(A) = 1$

Probability

A, C, G, T

$P($ A $|$ 

$) = 1/4$

$P($ A $|$ 

$) = 1$

Only A

Probability

A, C, G, T

$P($ A $\mid$ 

$) = 1/4$

$P($ A $\mid$ 

$) = 1$

Only A

Likelihood

A, C, G, T

$L($ 

$|\,A) = 1/4$

$L($ 

$|\,A) = 1$

Only A

Probability

$$P(\mathrm{ACTTG}) = ?$$

Probability

A, C, G, T

$$P(ACTTG \mid A) = ?$$

Probability

A, C, G, T

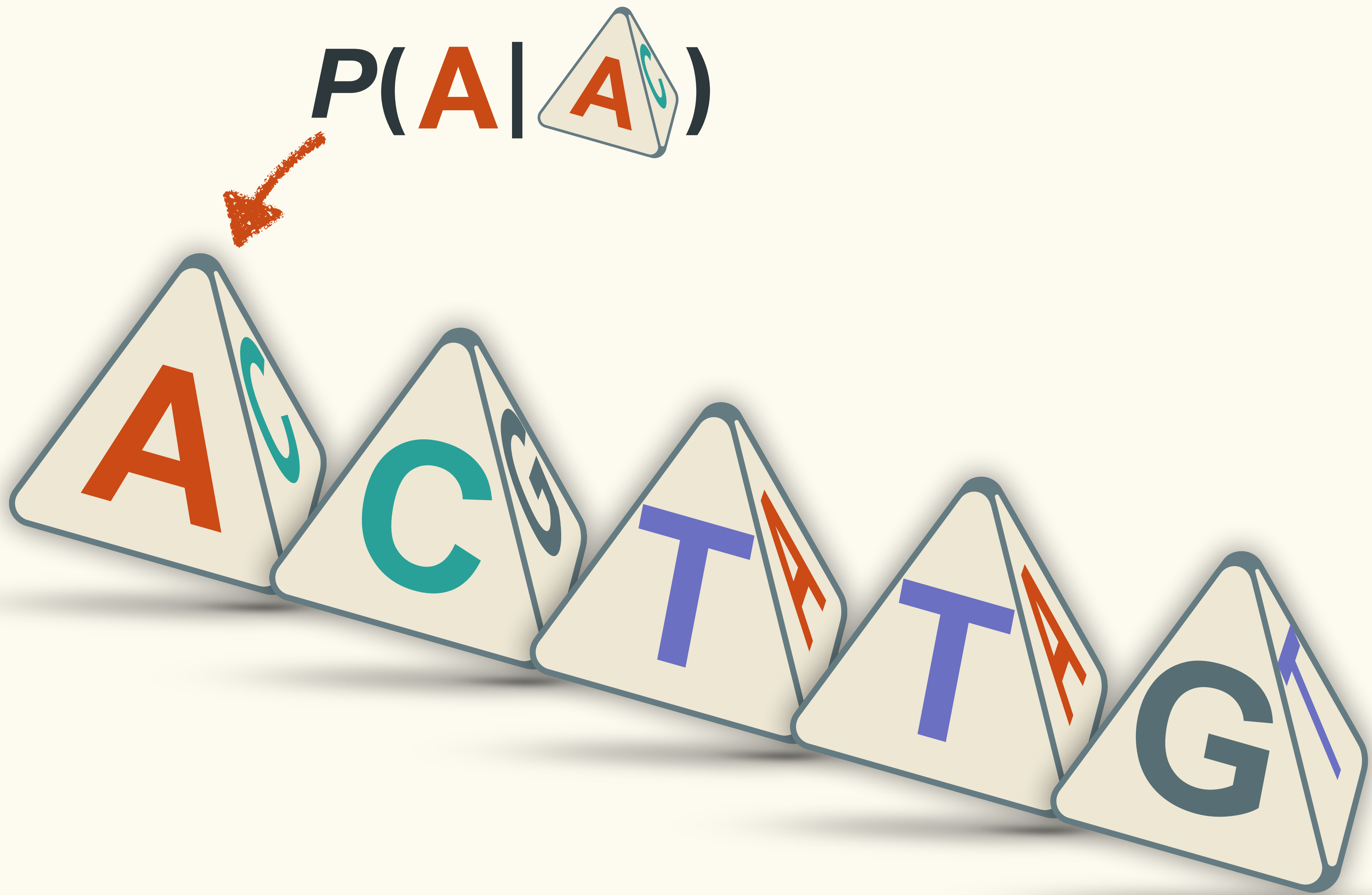
$$P(ACTTG \mid A) = ?$$

Probability

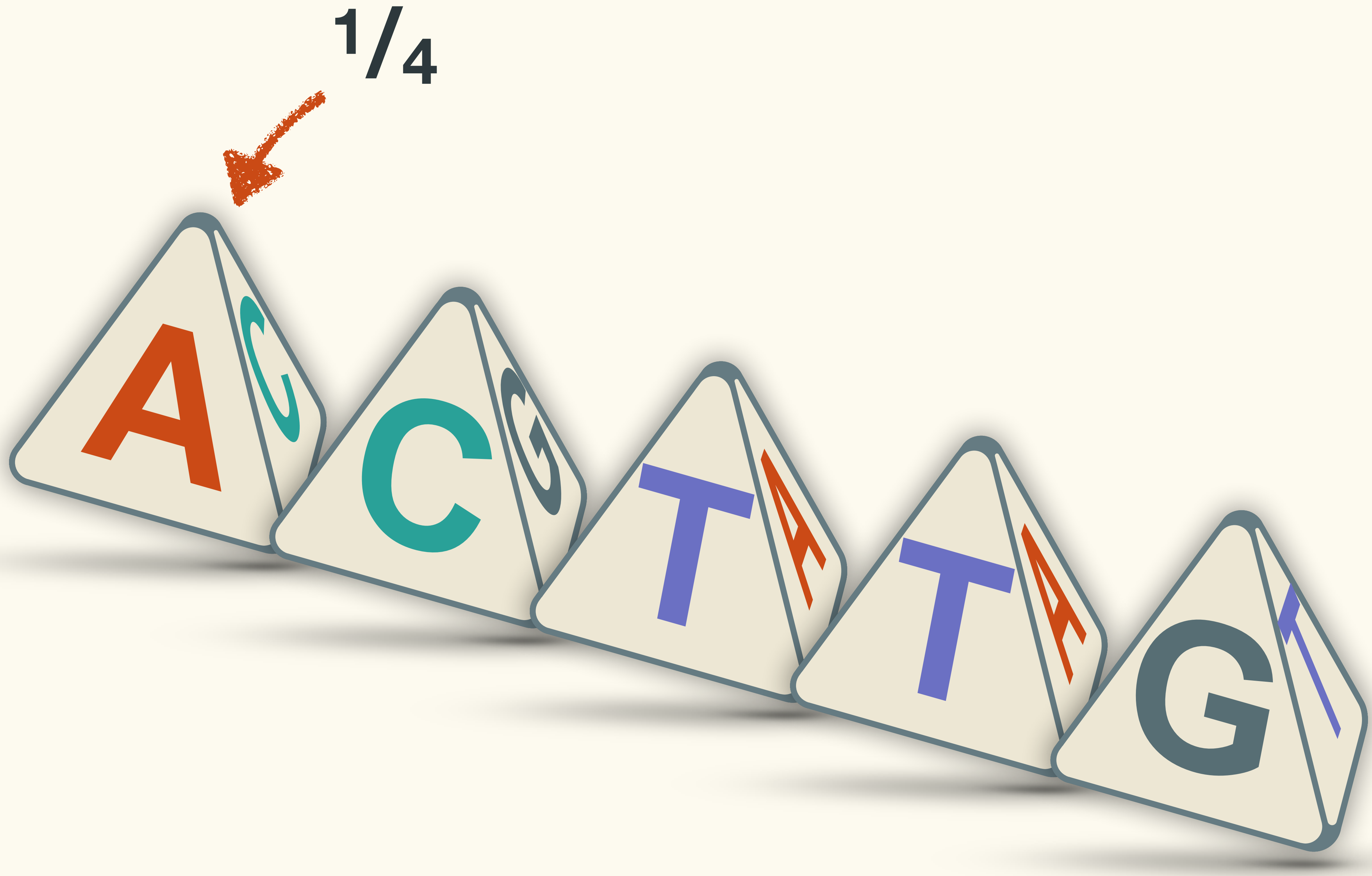
A, C, G, T

$P(A \,\&\, C \,\&\, T \,\&\, T \,\&\, G \mid$  $) = ?$

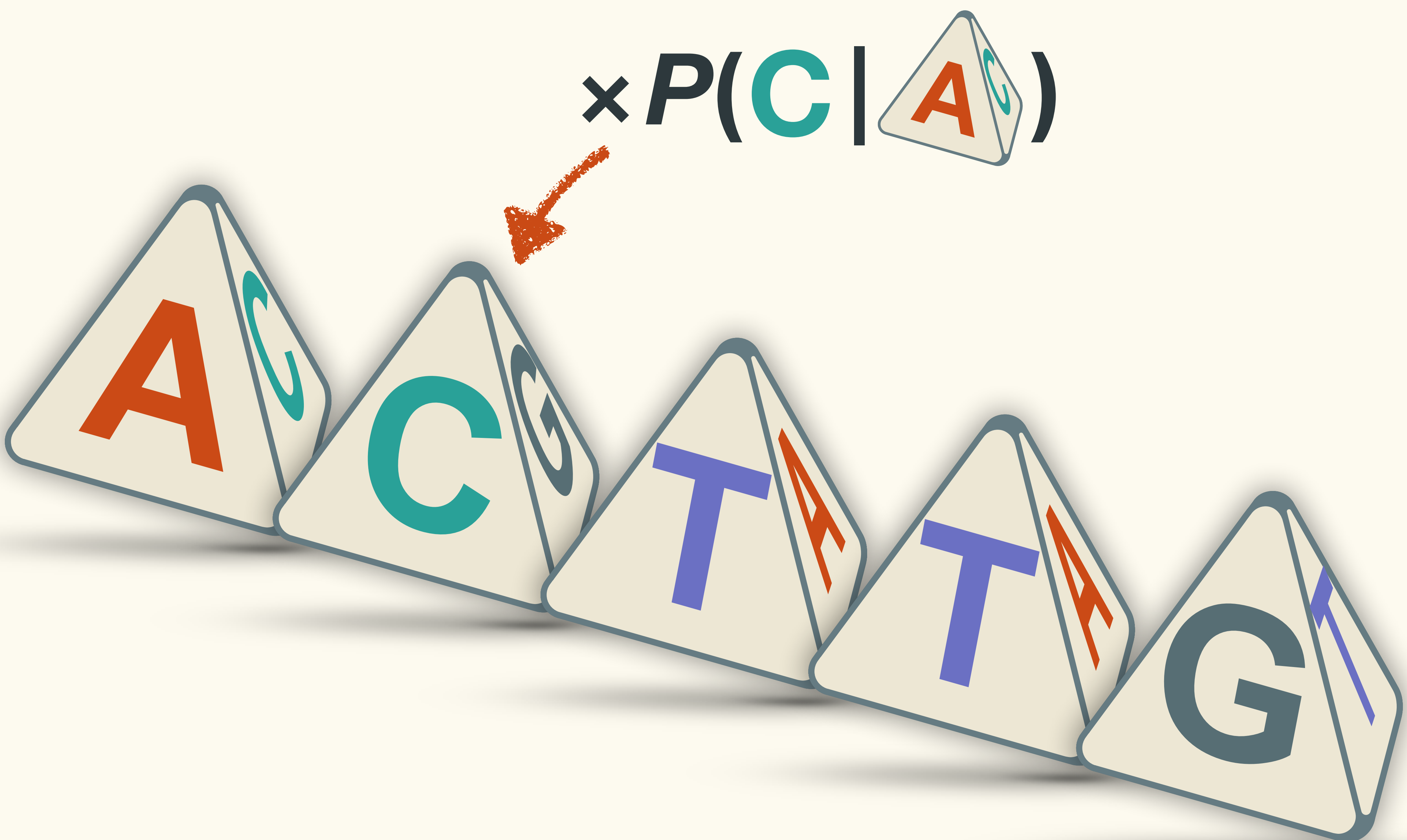
Probability



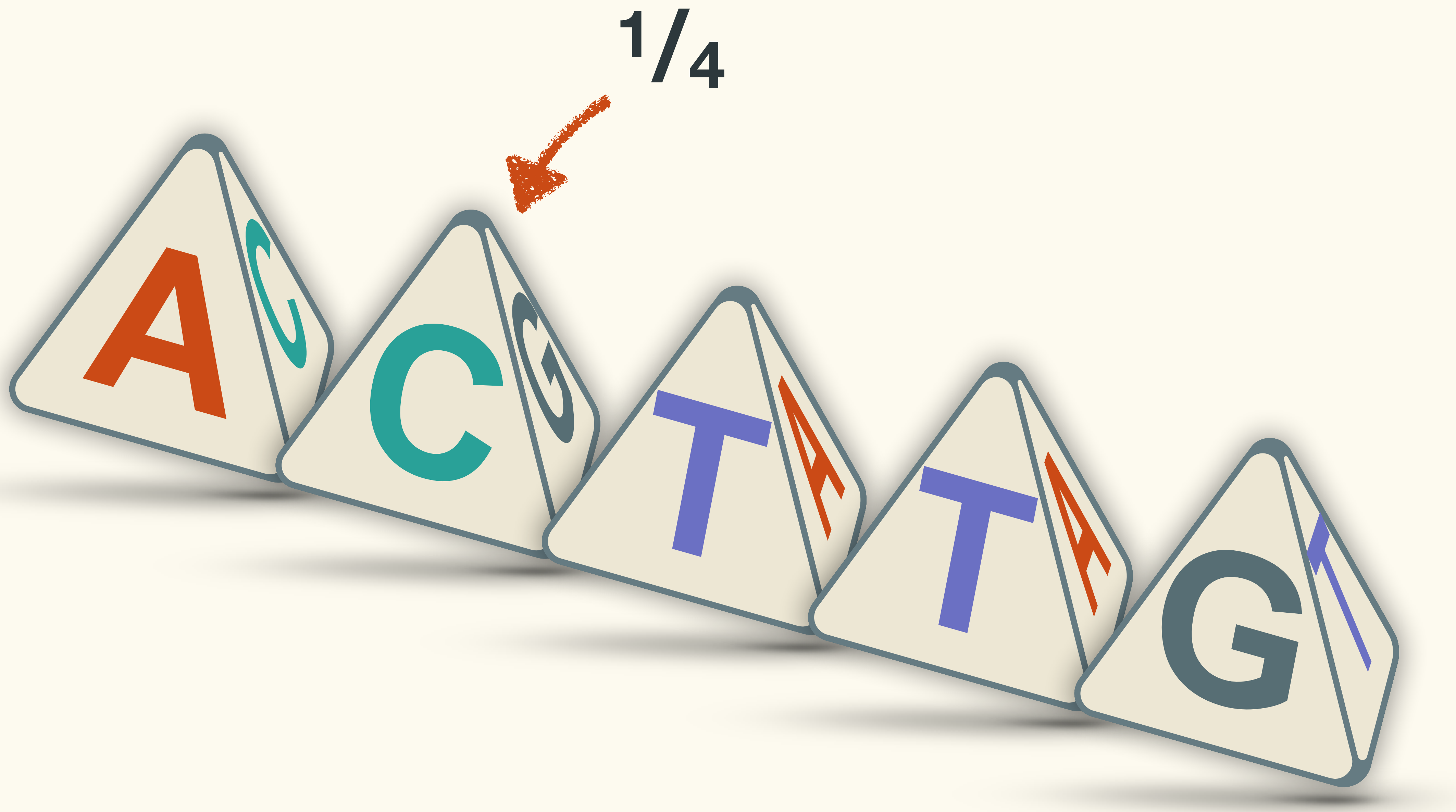
Probability



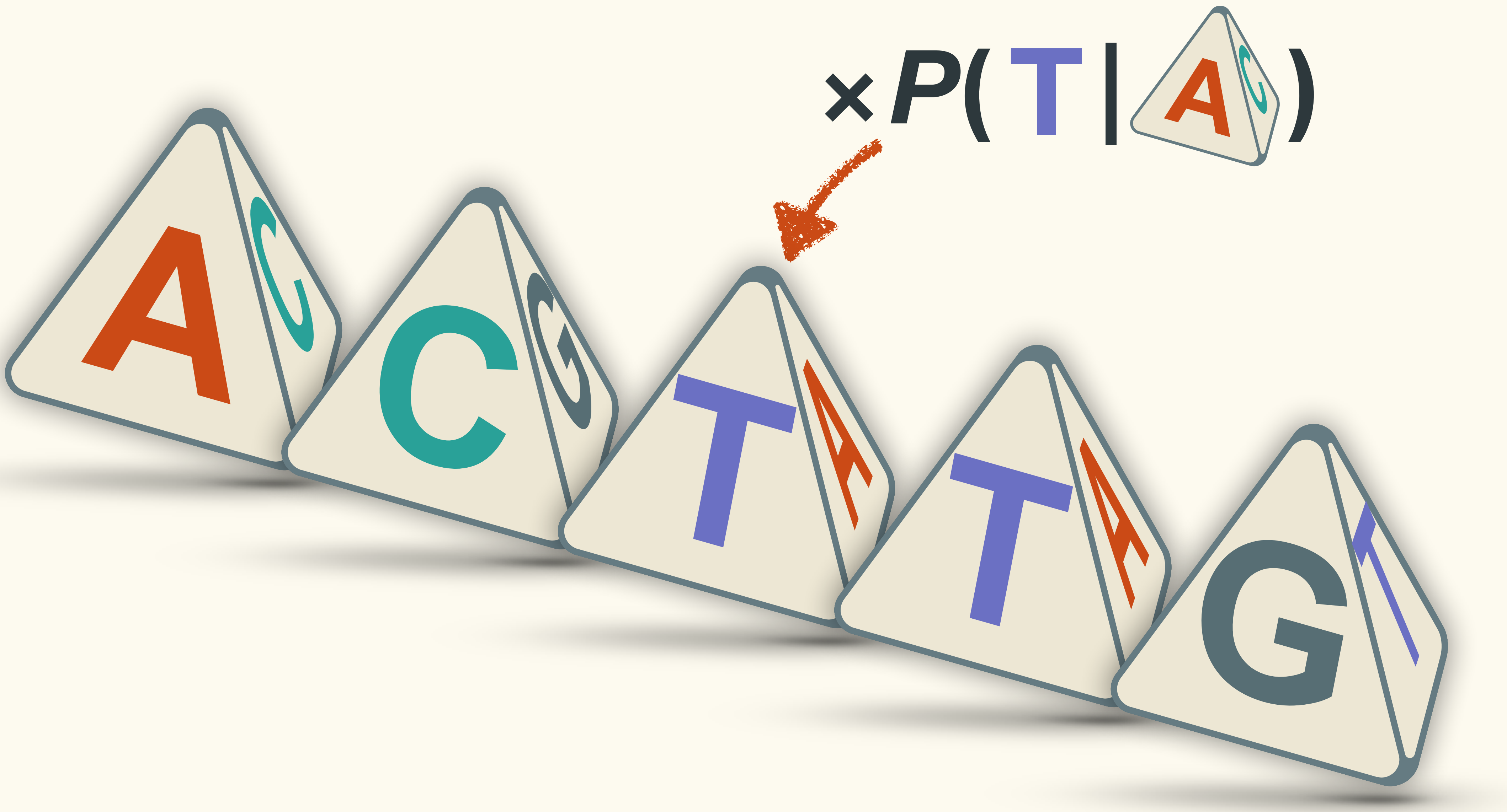
Probability



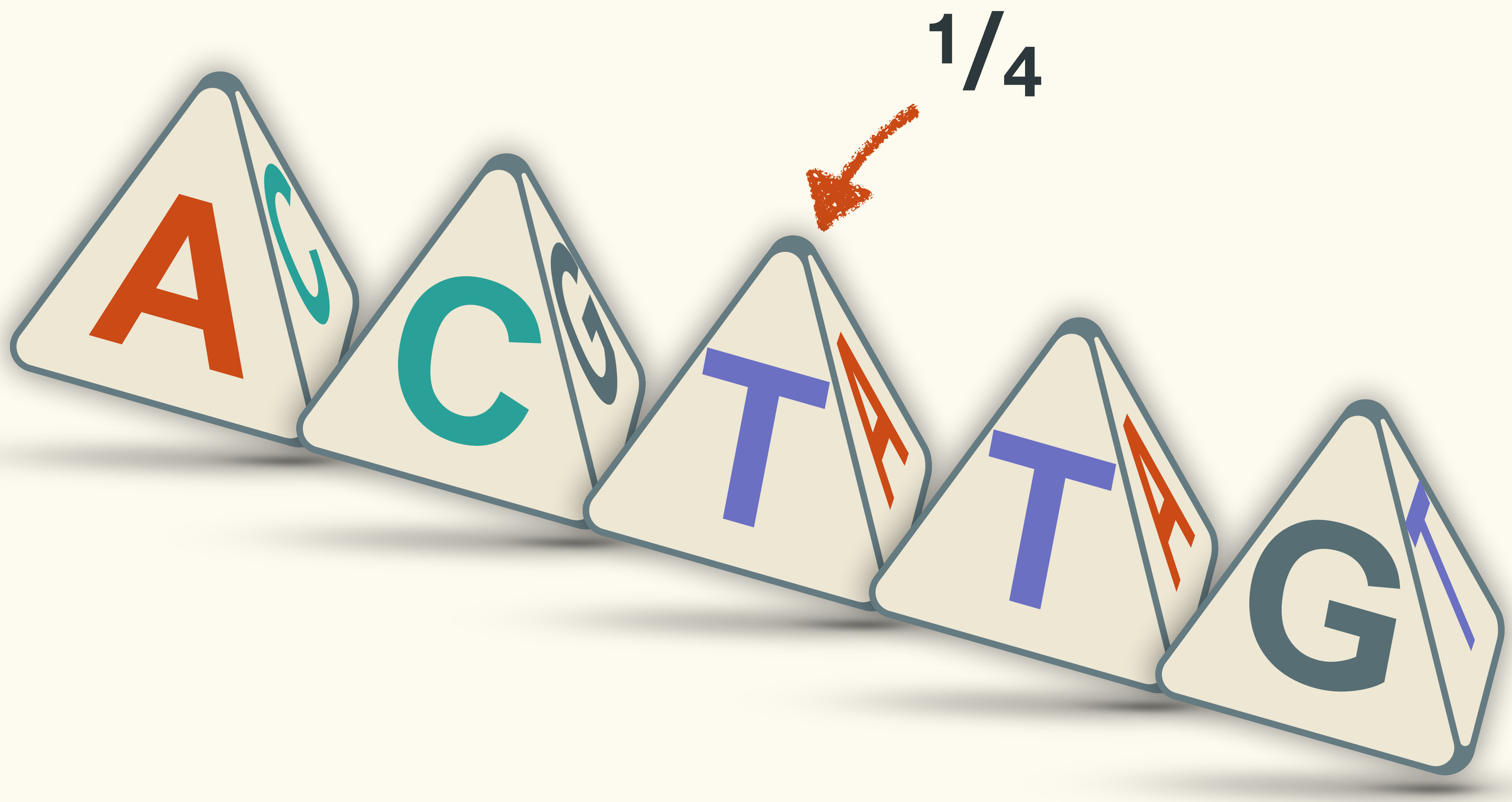
Probability



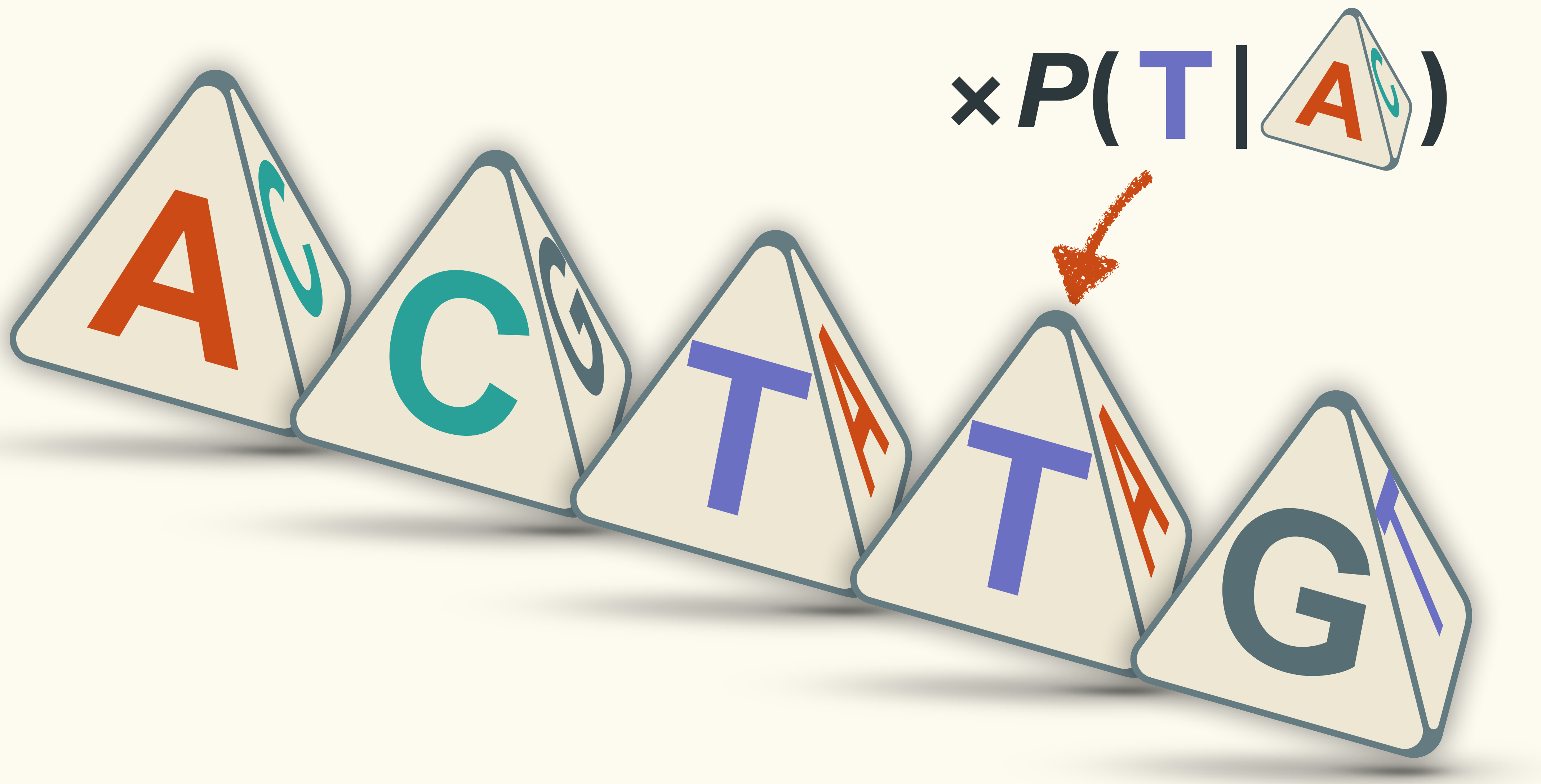
Probability



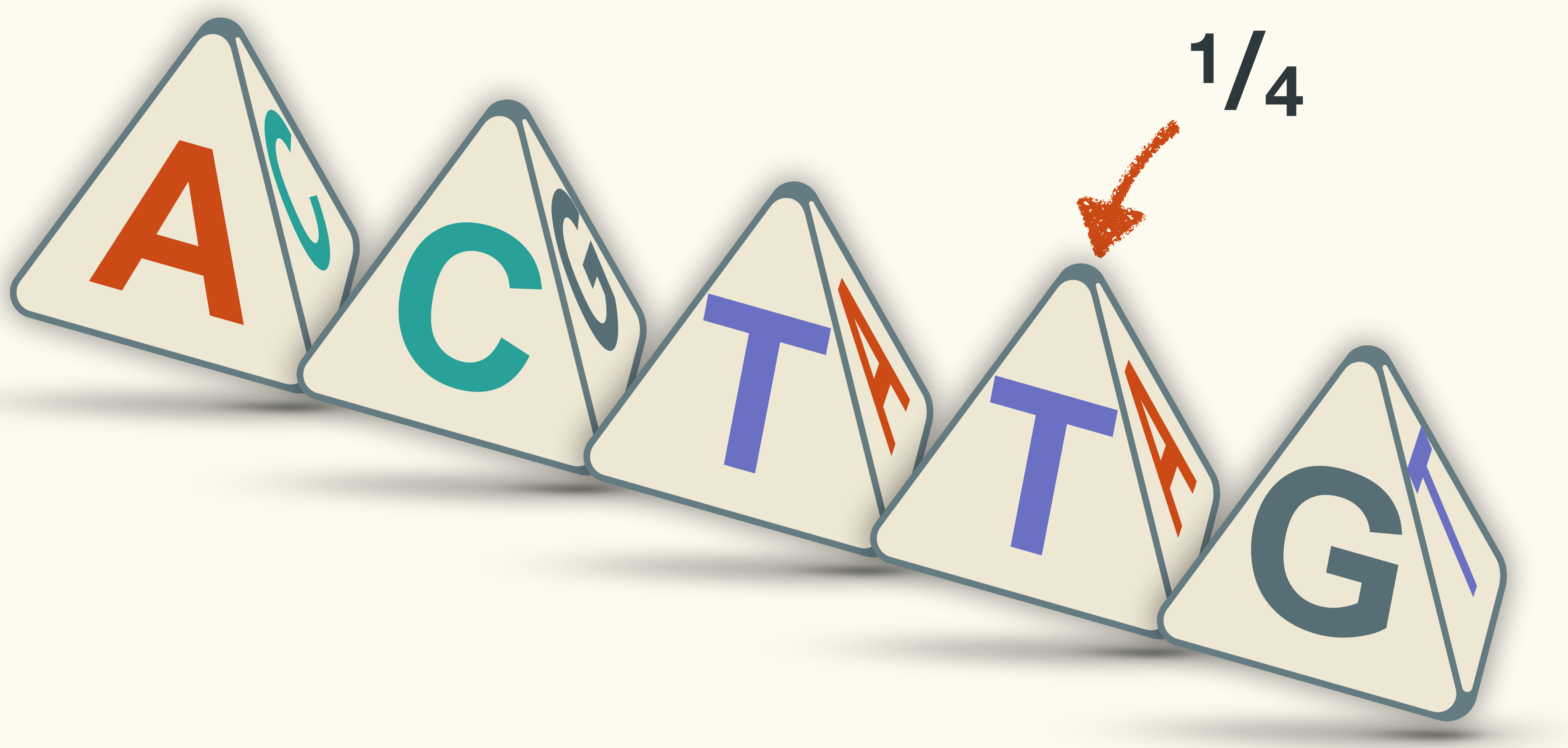
Probability



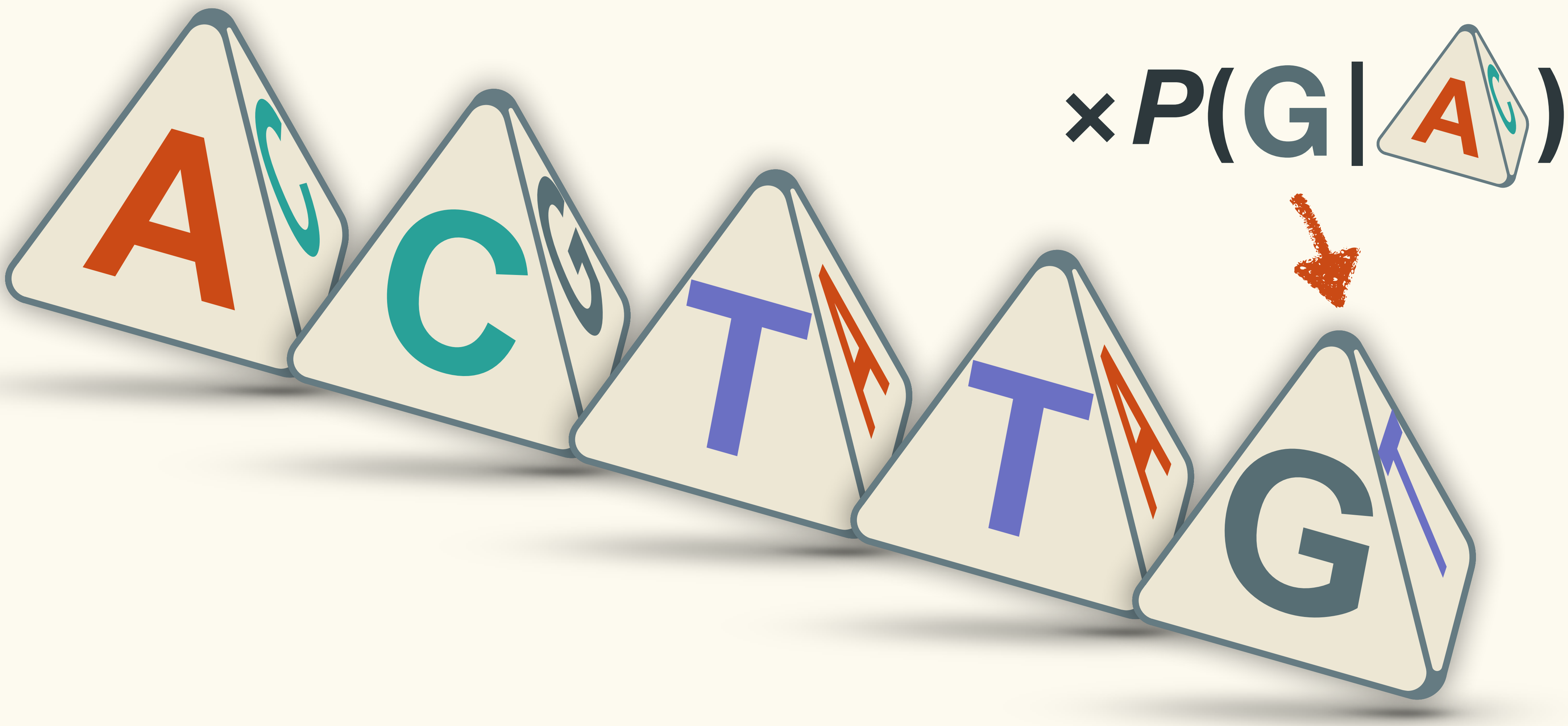
Probability



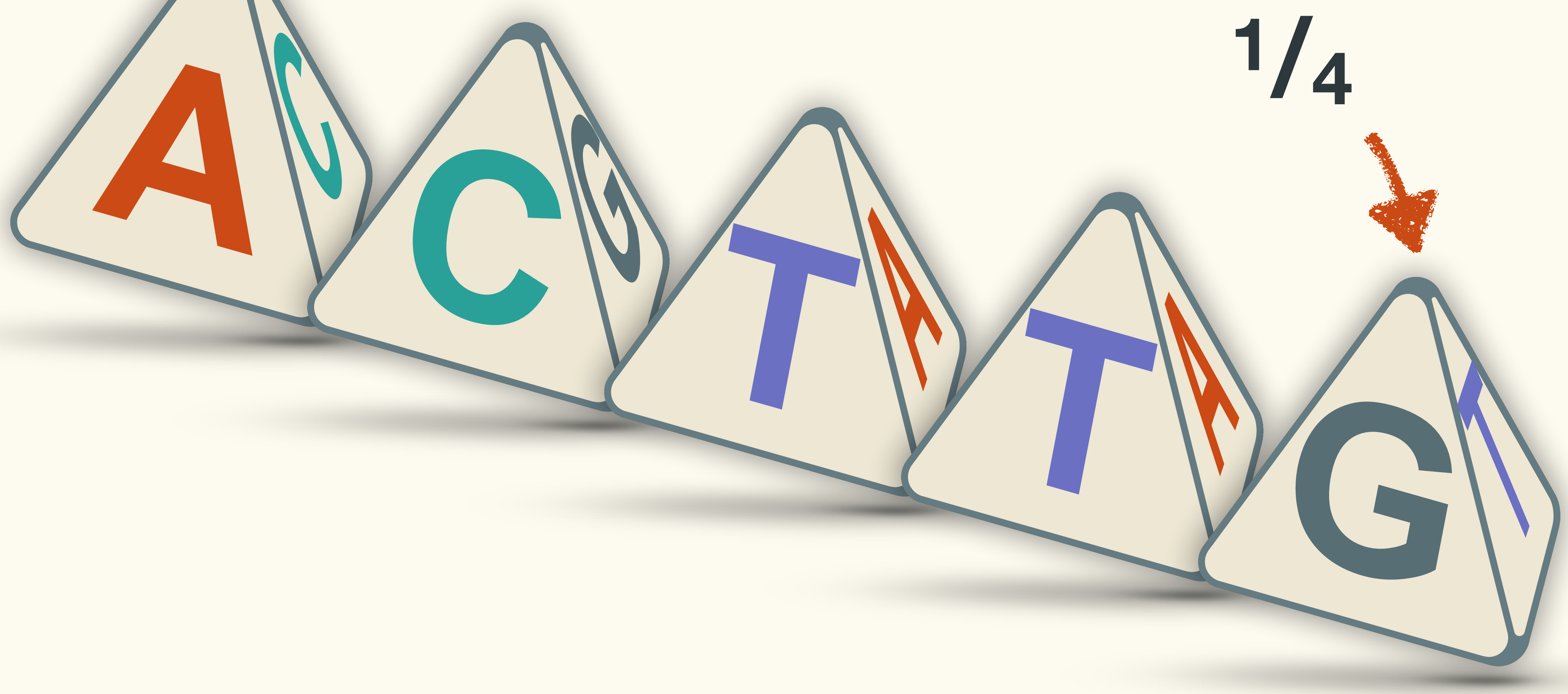
Probability



Probability

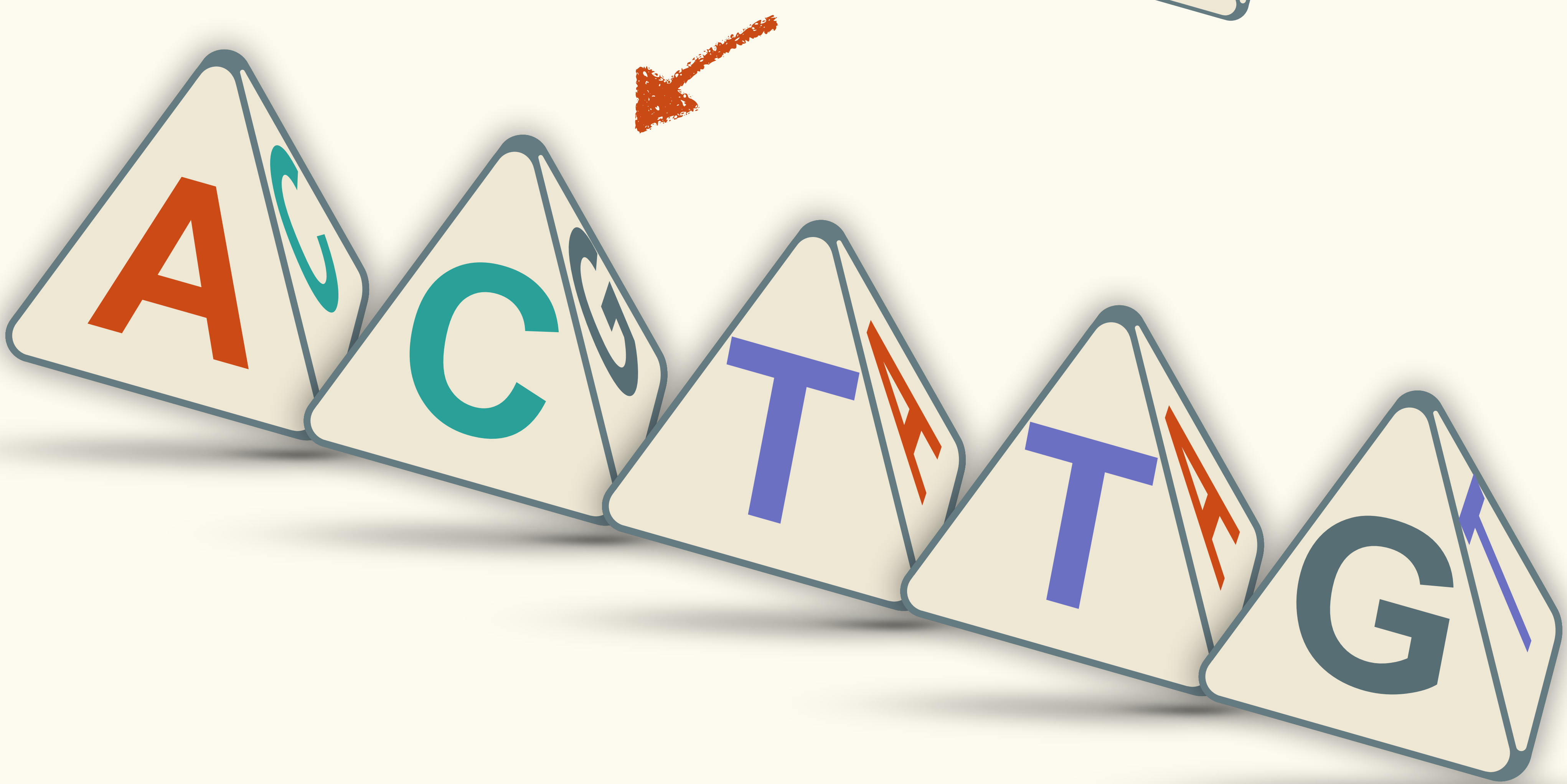


Probability

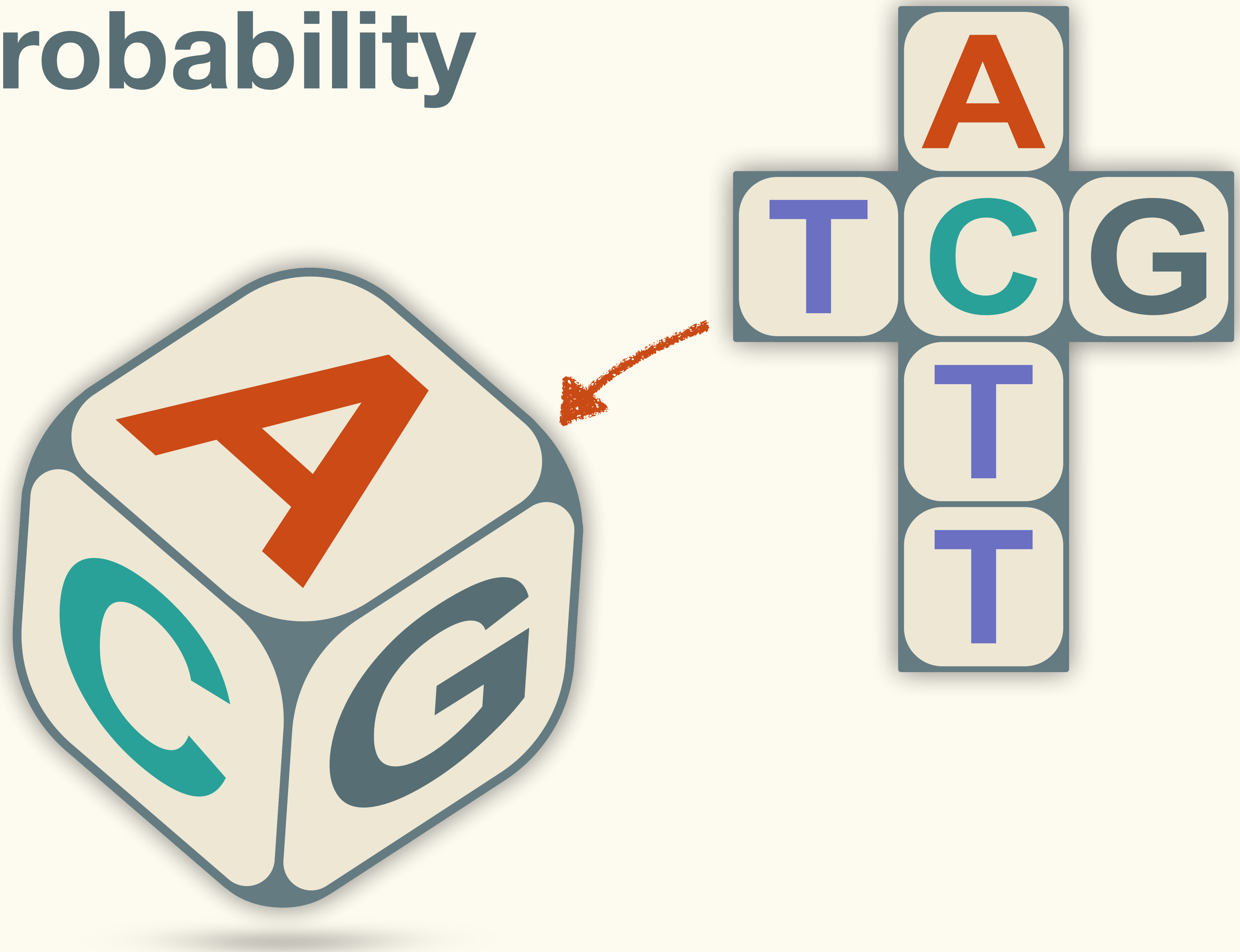


Probability

$P(\mathrm{A} \,\&\, \mathrm{C} \,\&\, \mathrm{T} \,\&\, \mathrm{T} \,\&\, \mathrm{G} \mid$  $) = ({}^{1}/_{4})^{5}$



Probability



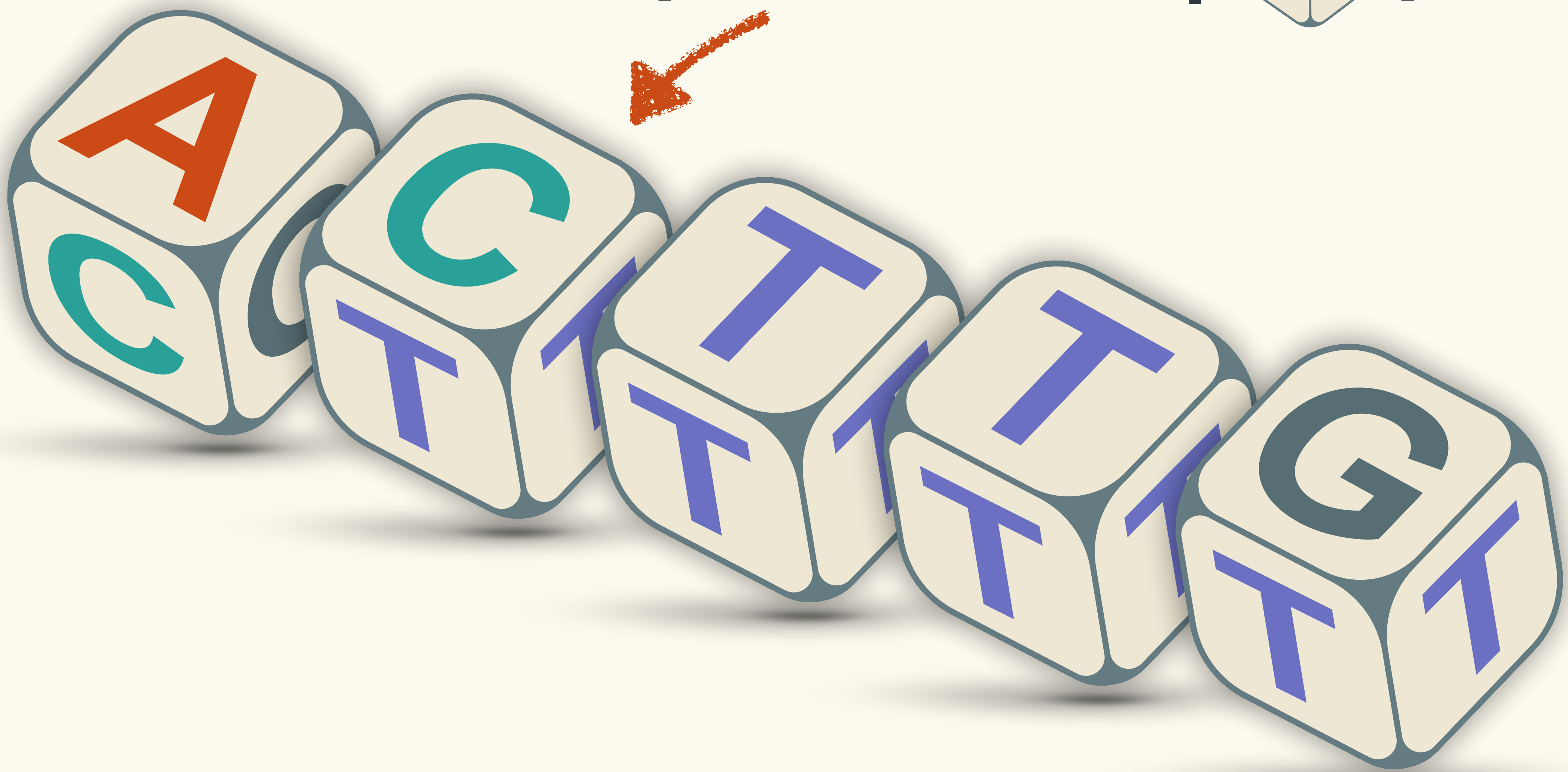
Probability

$P(\text{ACTTG}) = ?$

Probability

$f(T) > f(A), f(C), f(G)$

$P($ACTTG $|$  $) = ?$



Probability

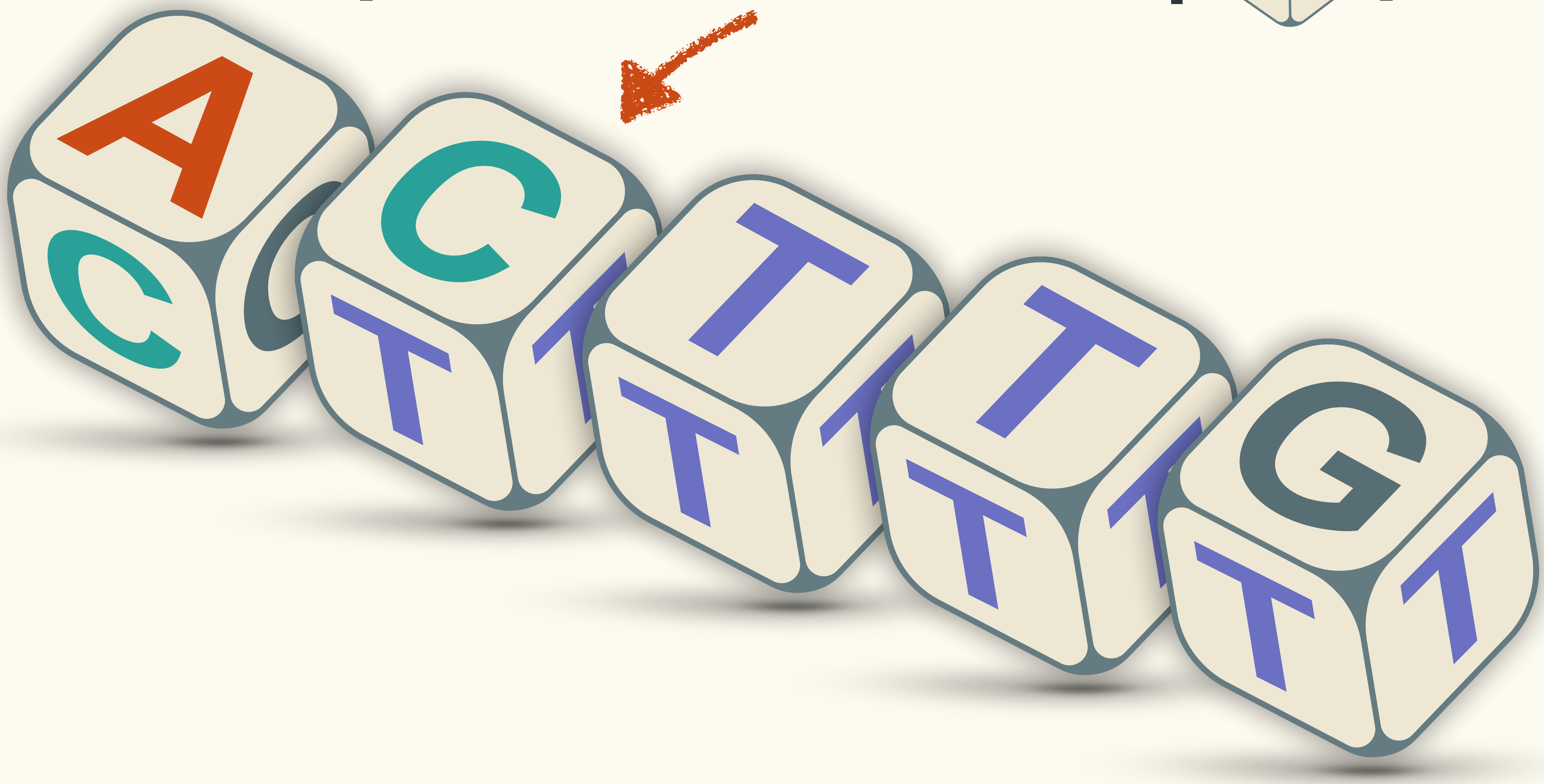
f(T) > f(A), f(C), f(G)

$P($ACTTG $|$  $) = ?$

Probability

f(T) > f(A), f(C), f(G)

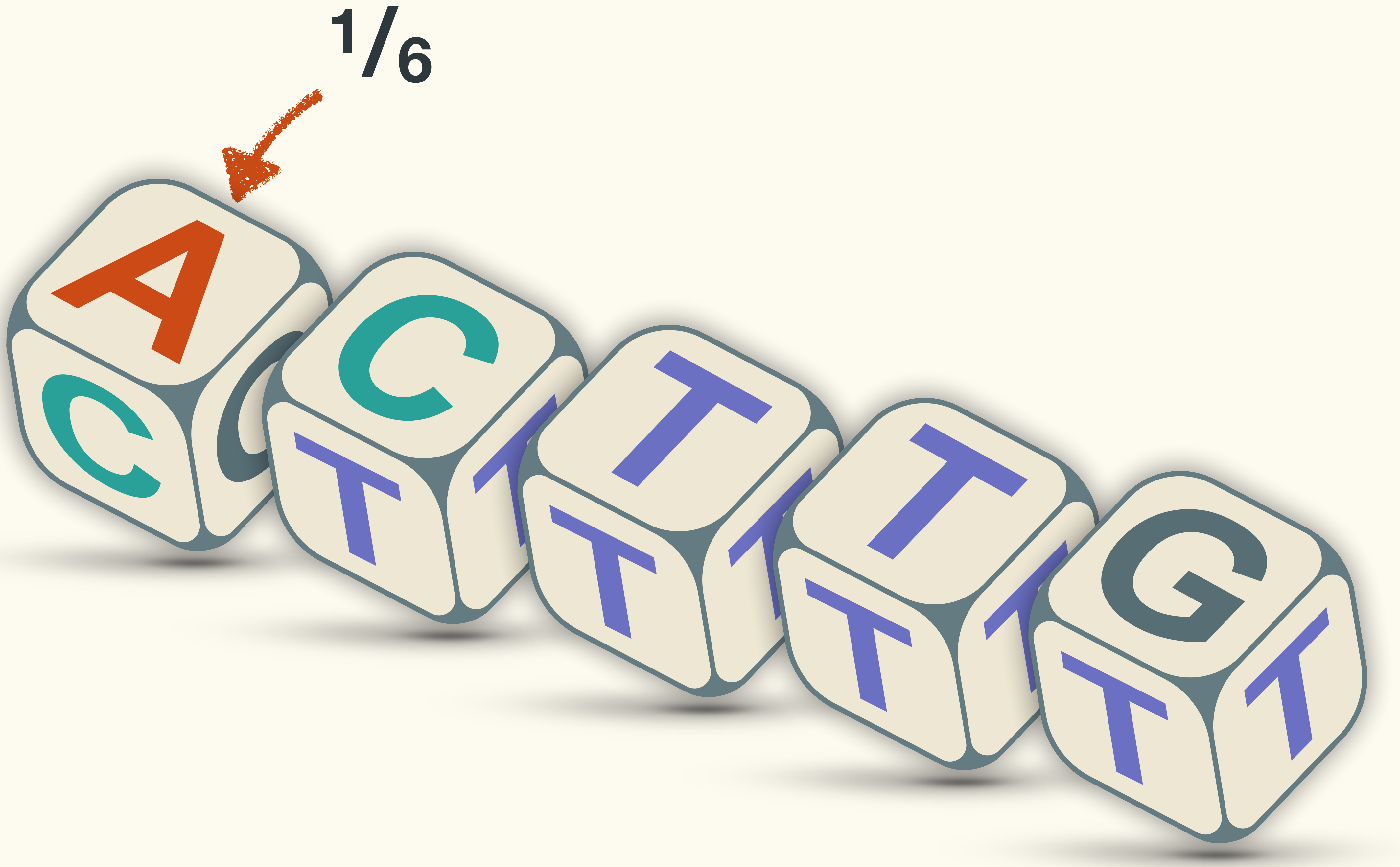
P(A & C & T & T & G | ) = ?



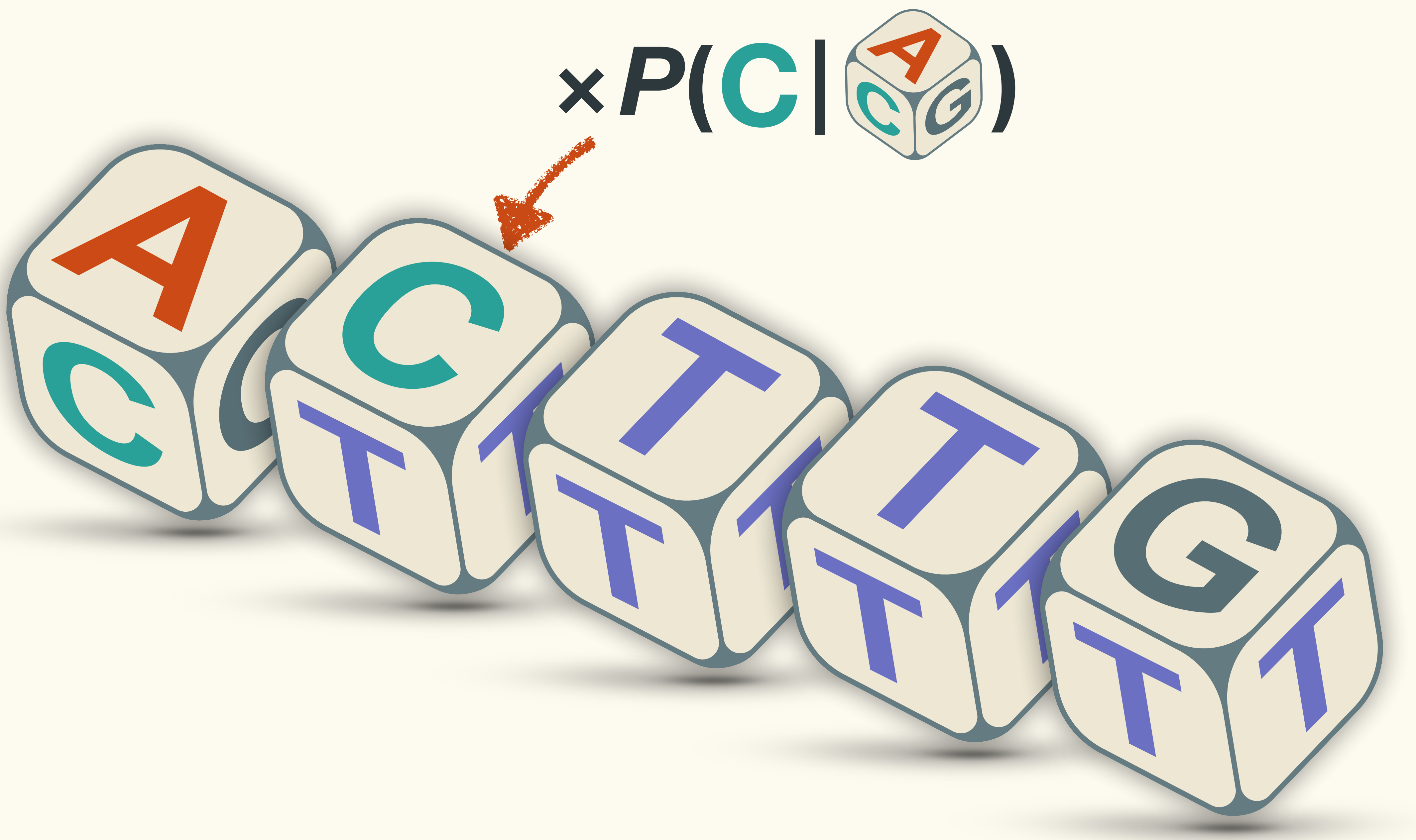
Probability

P(A | )

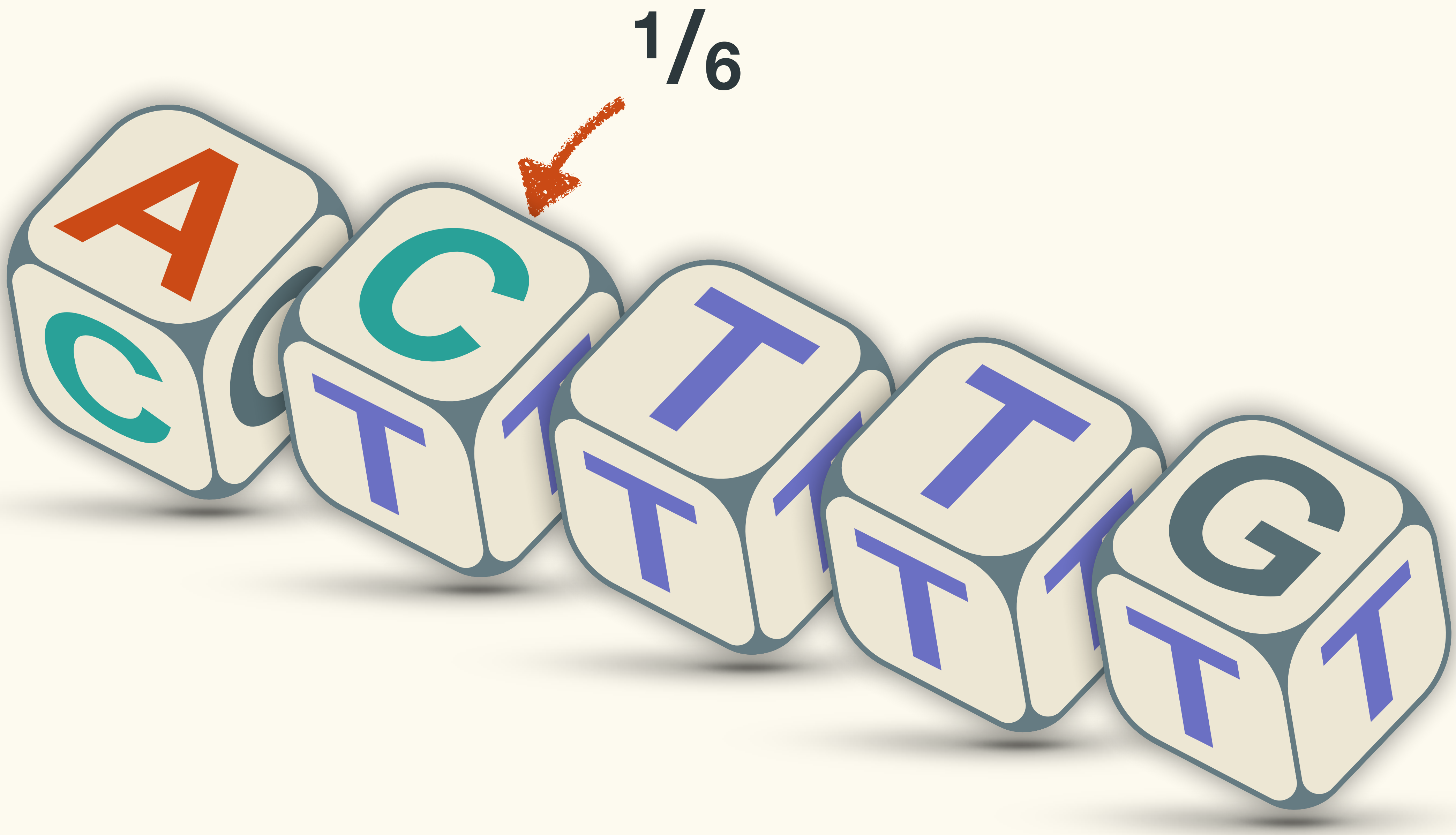
Probability



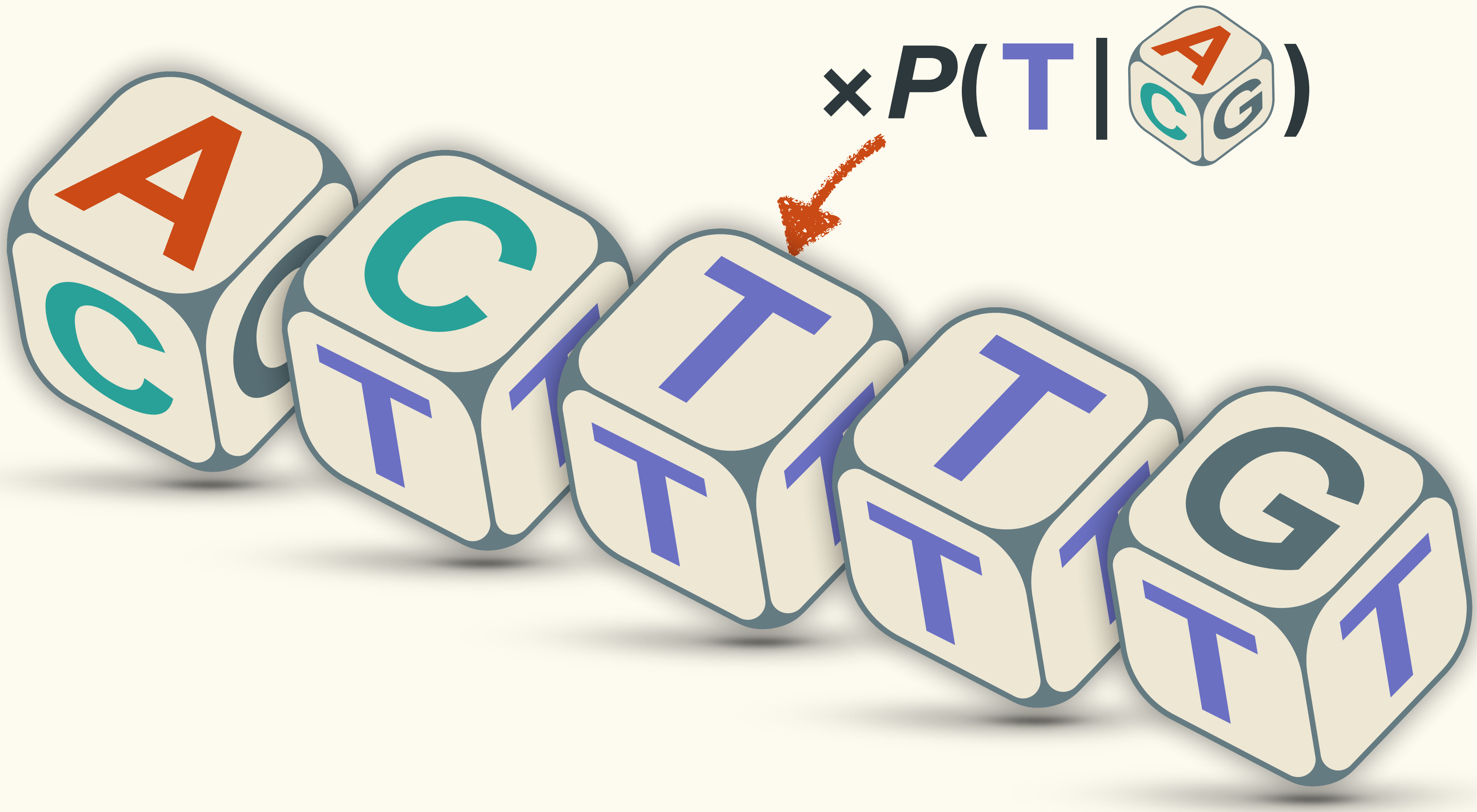
Probability



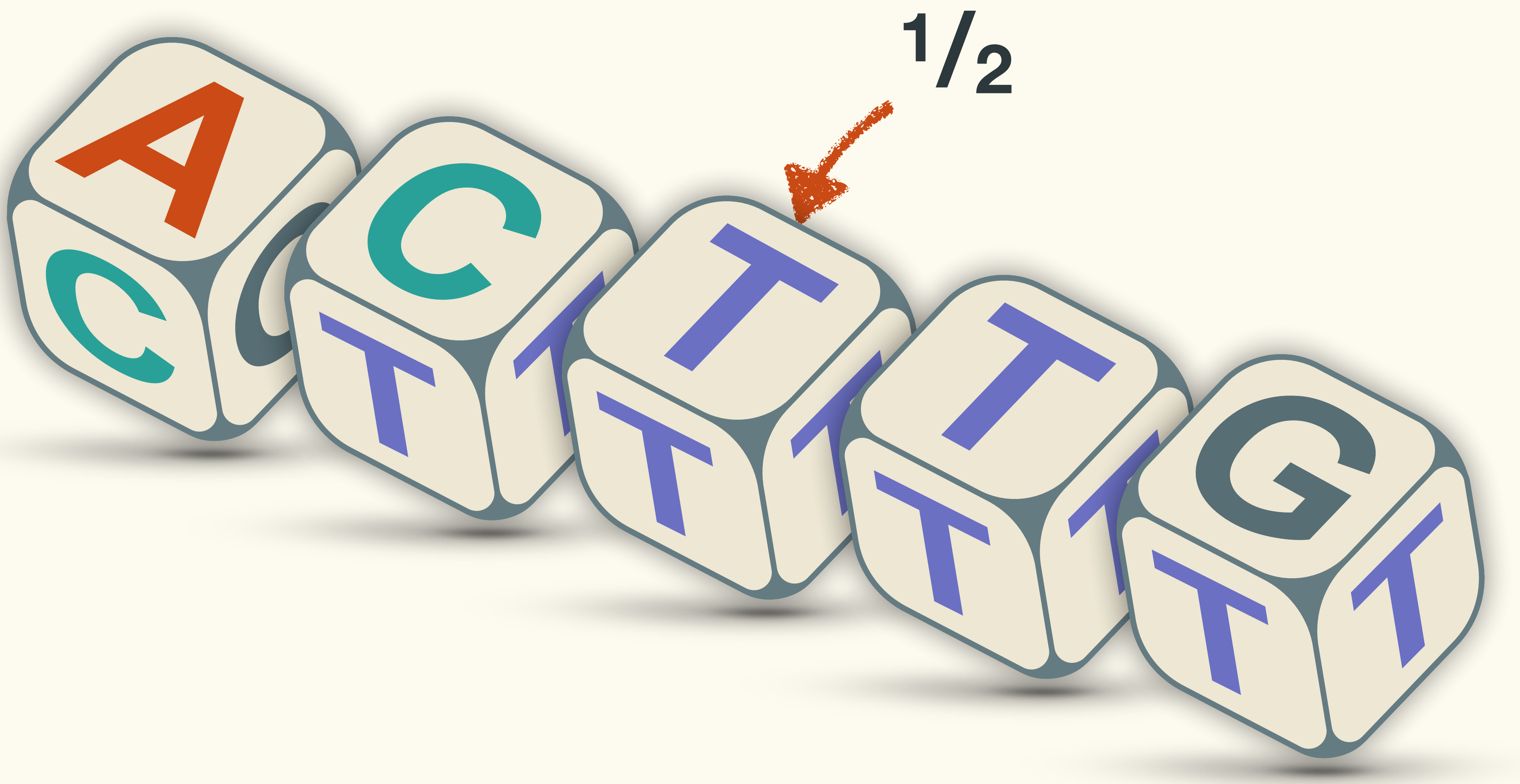
Probability



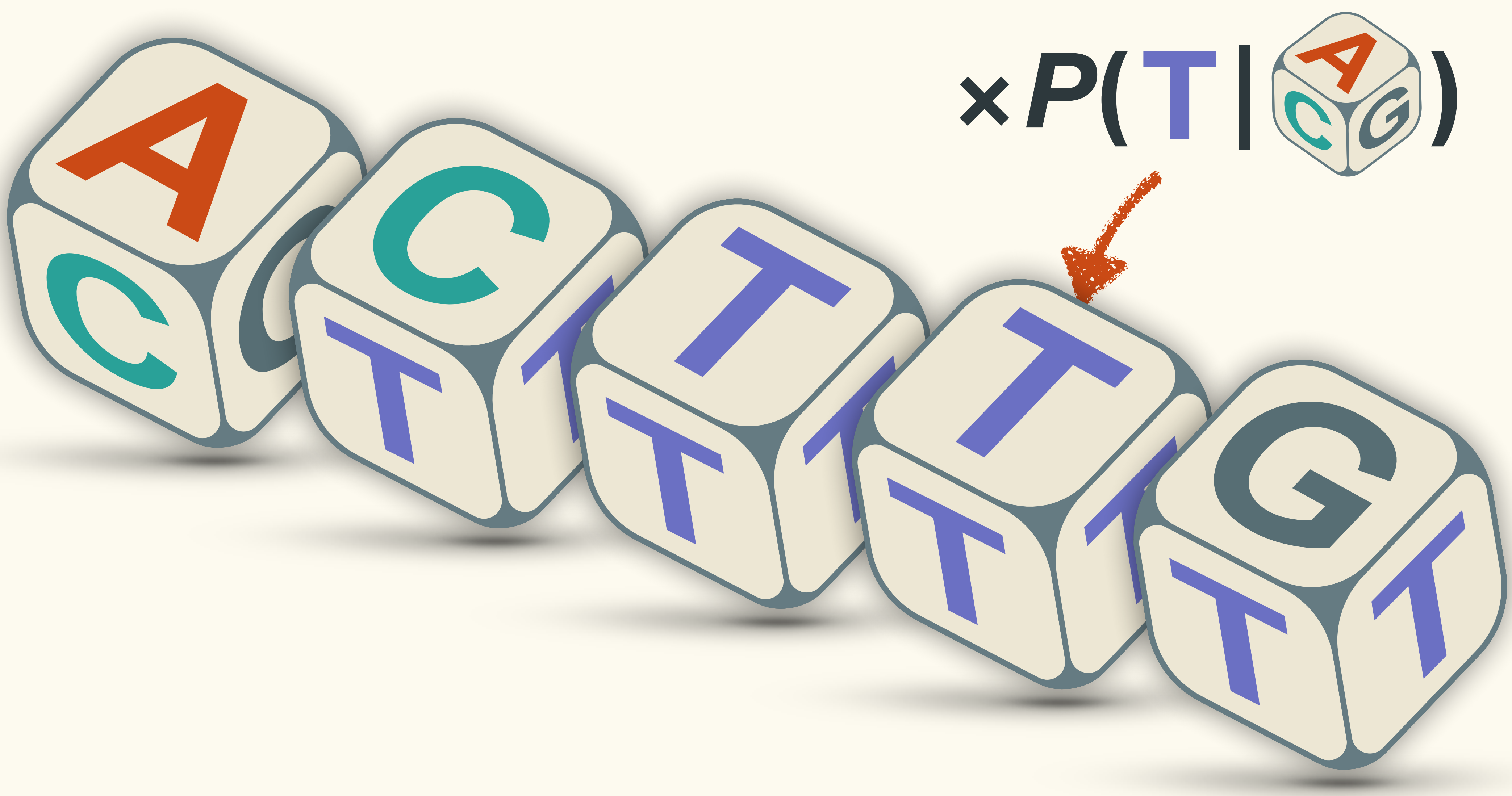
Probability



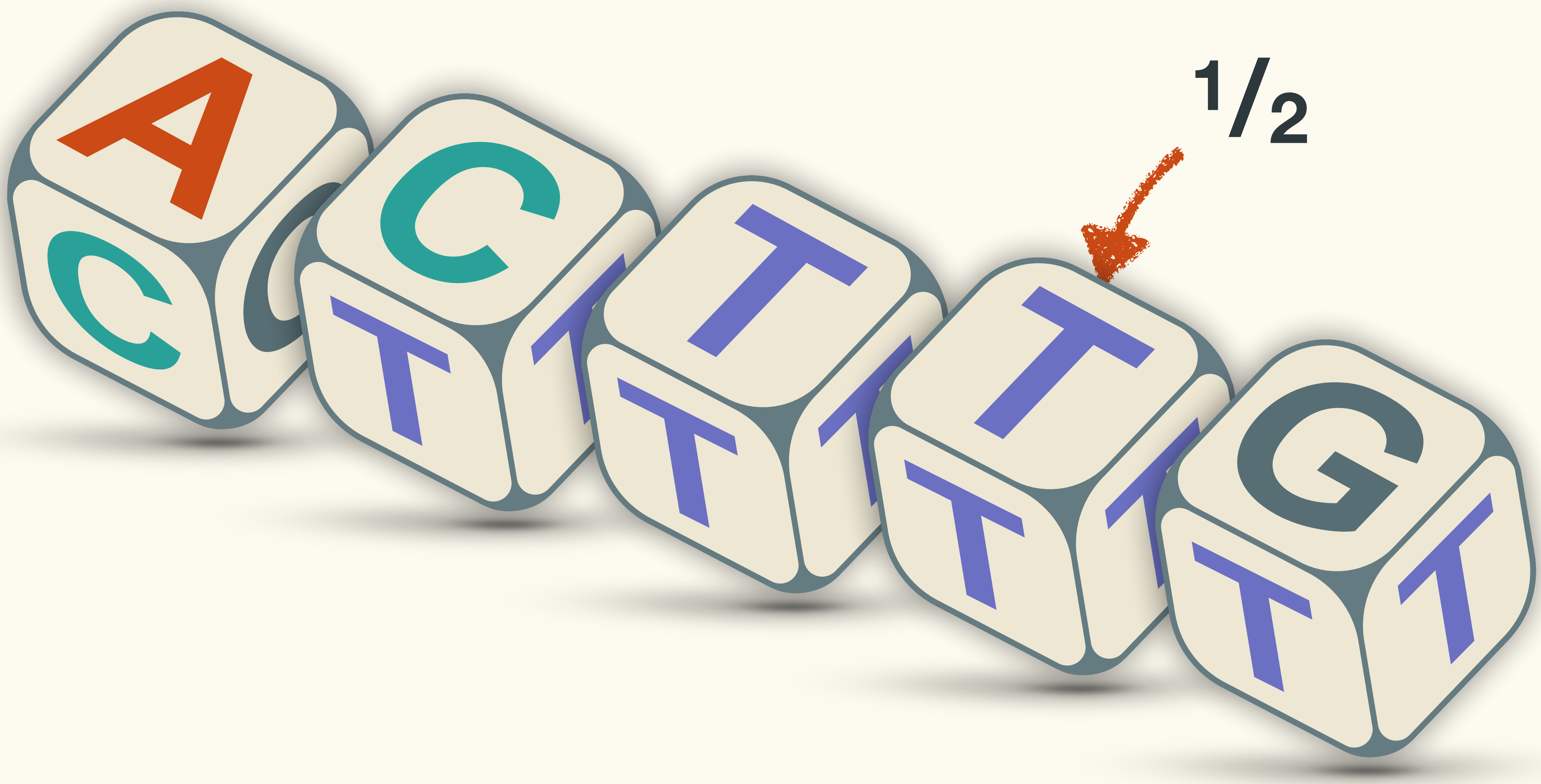
Probability



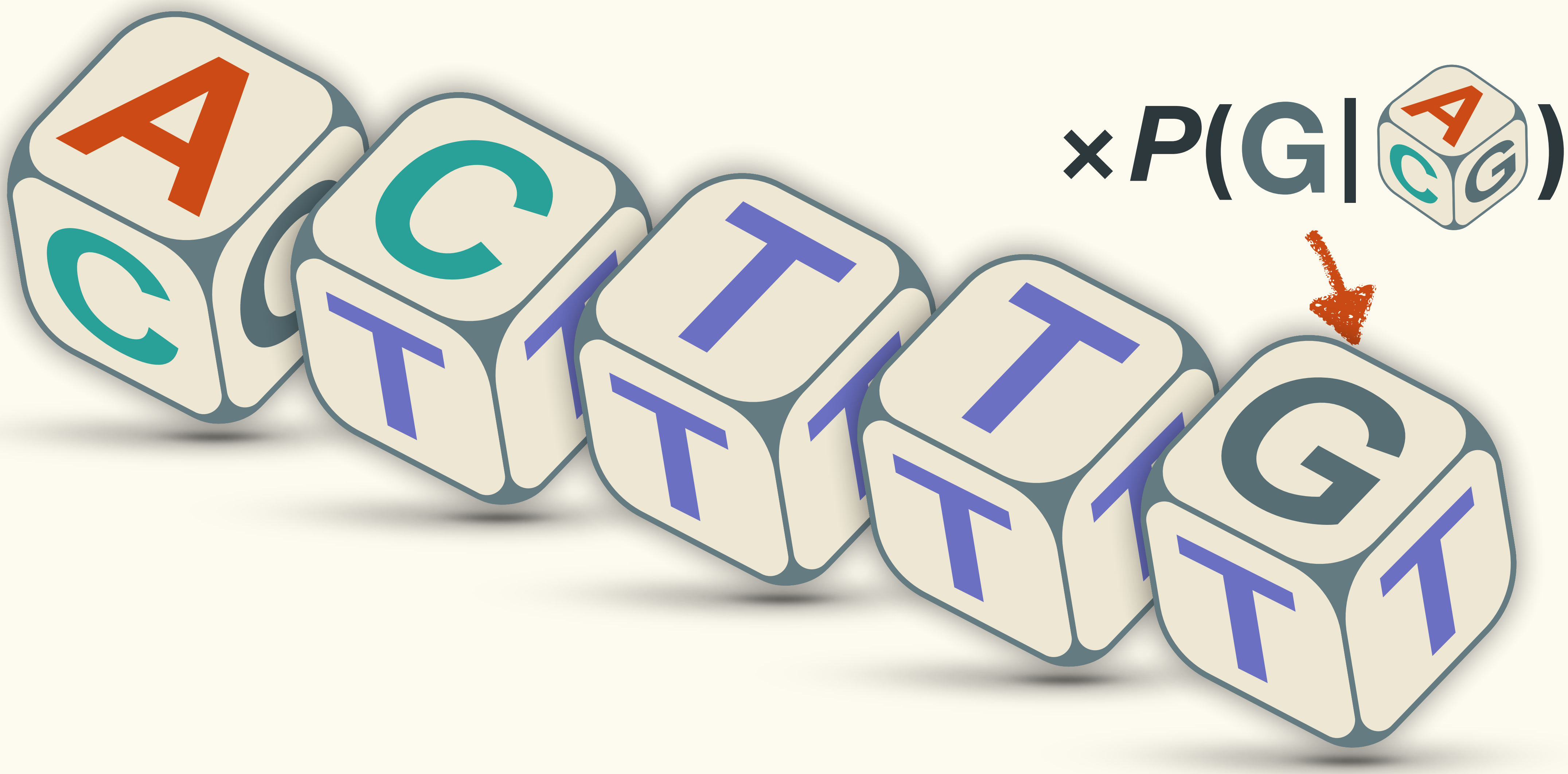
Probability



Probability



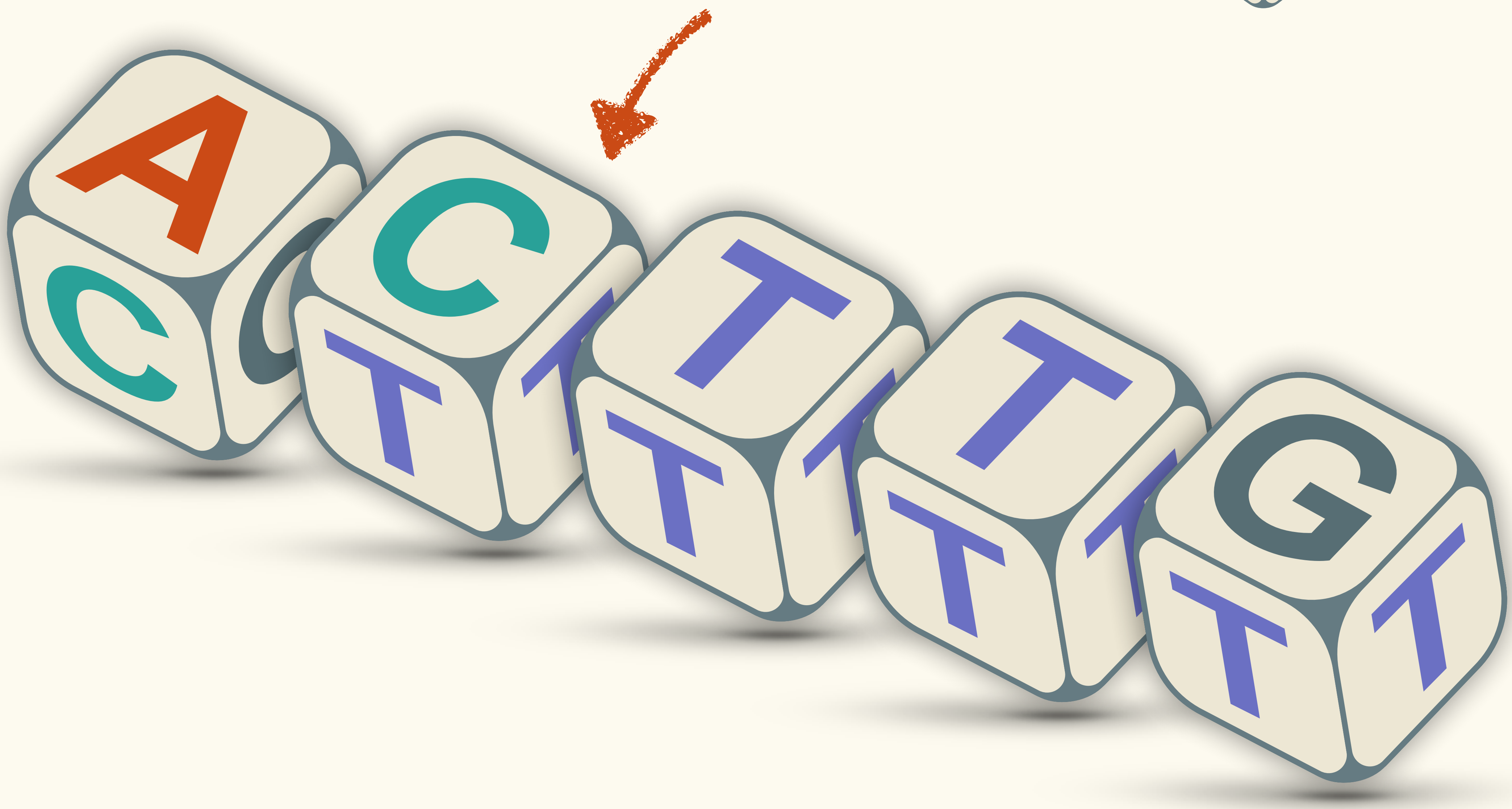
Probability



Probability

Probability

$P($A & C & T & T & G $|$  $) = ?$



Probability

P(A & C & T & T & G | ) = $(^1/_6)^3(^1/_2)^2$

Probability

$P(A \,\&\, C \,\&\, T \,\&\, T \,\&\, G \mid$ 

$) = ({}^{1}/_{4})^{5}$

$P(A \,\&\, C \,\&\, T \,\&\, T \,\&\, G \mid$ 

$) = ({}^{1}/_{6})^{3}({}^{1}/_{2})^{2}$

Probability

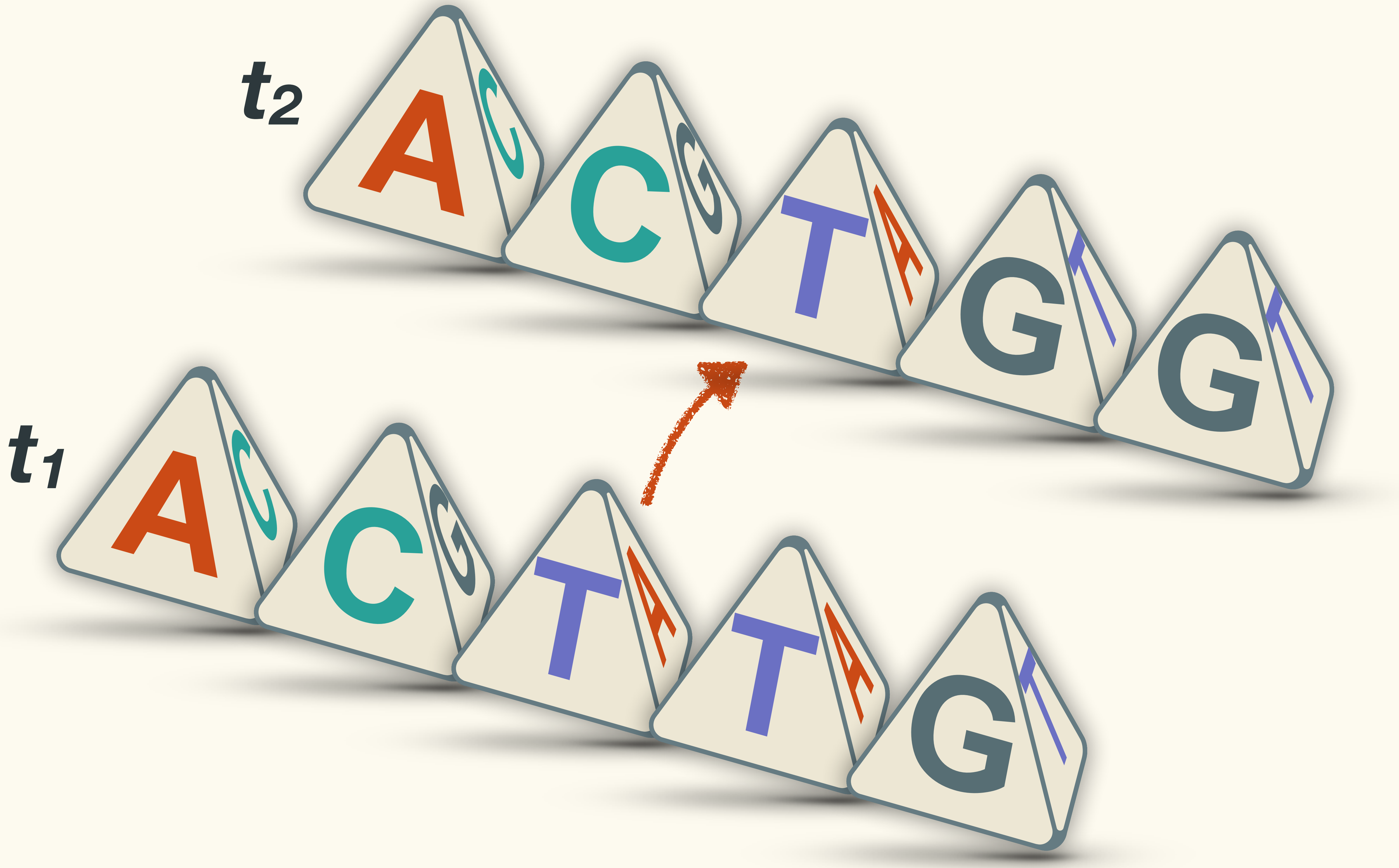
$P(A \,\&\, C \,\&\, T \,\&\, T \,\&\, G \mid$ 

$) = 0.0010$

$P(A \,\&\, C \,\&\, T \,\&\, T \,\&\, G \mid$ 

$) = 0.0012$

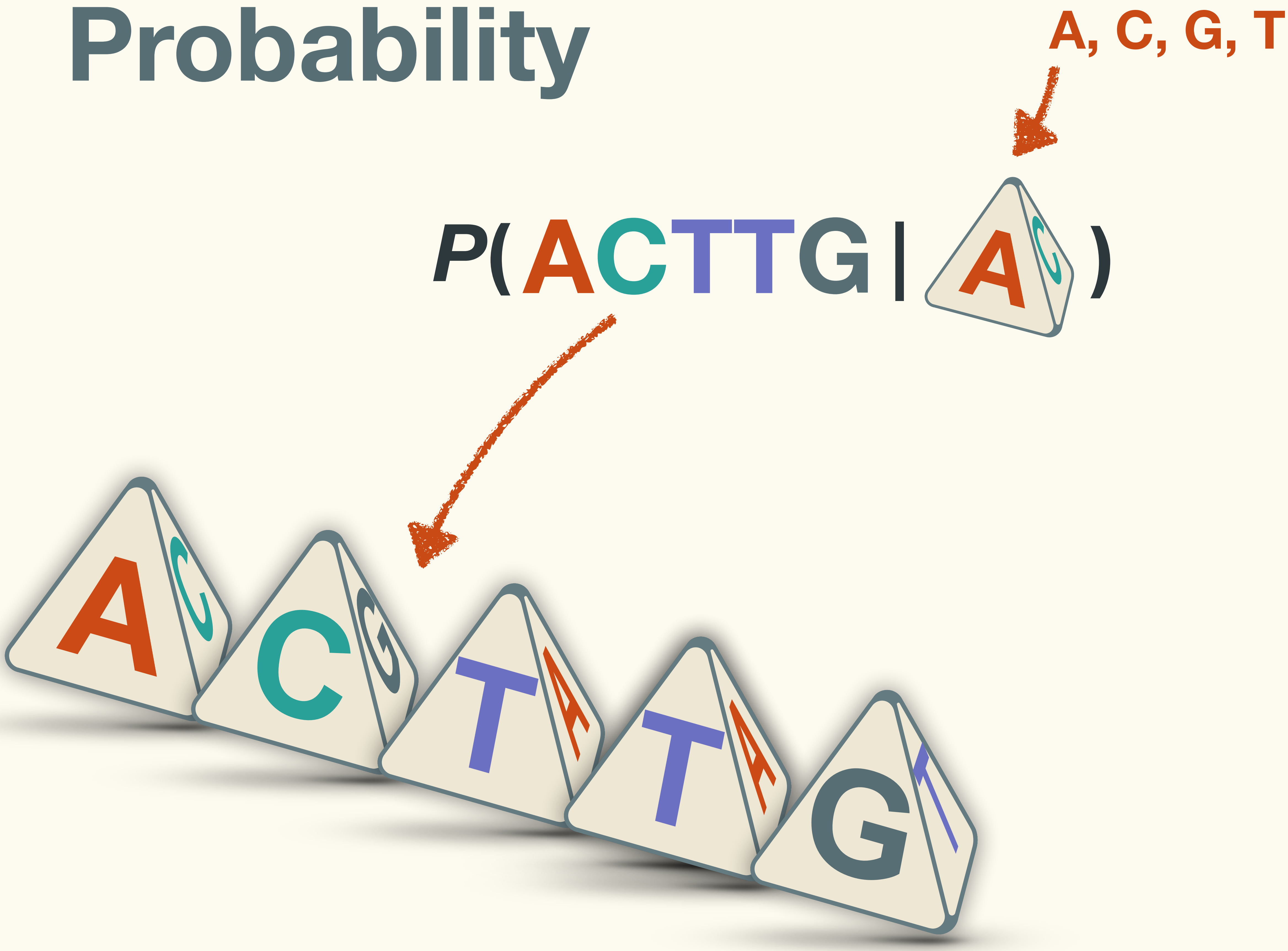
Probability



Probability

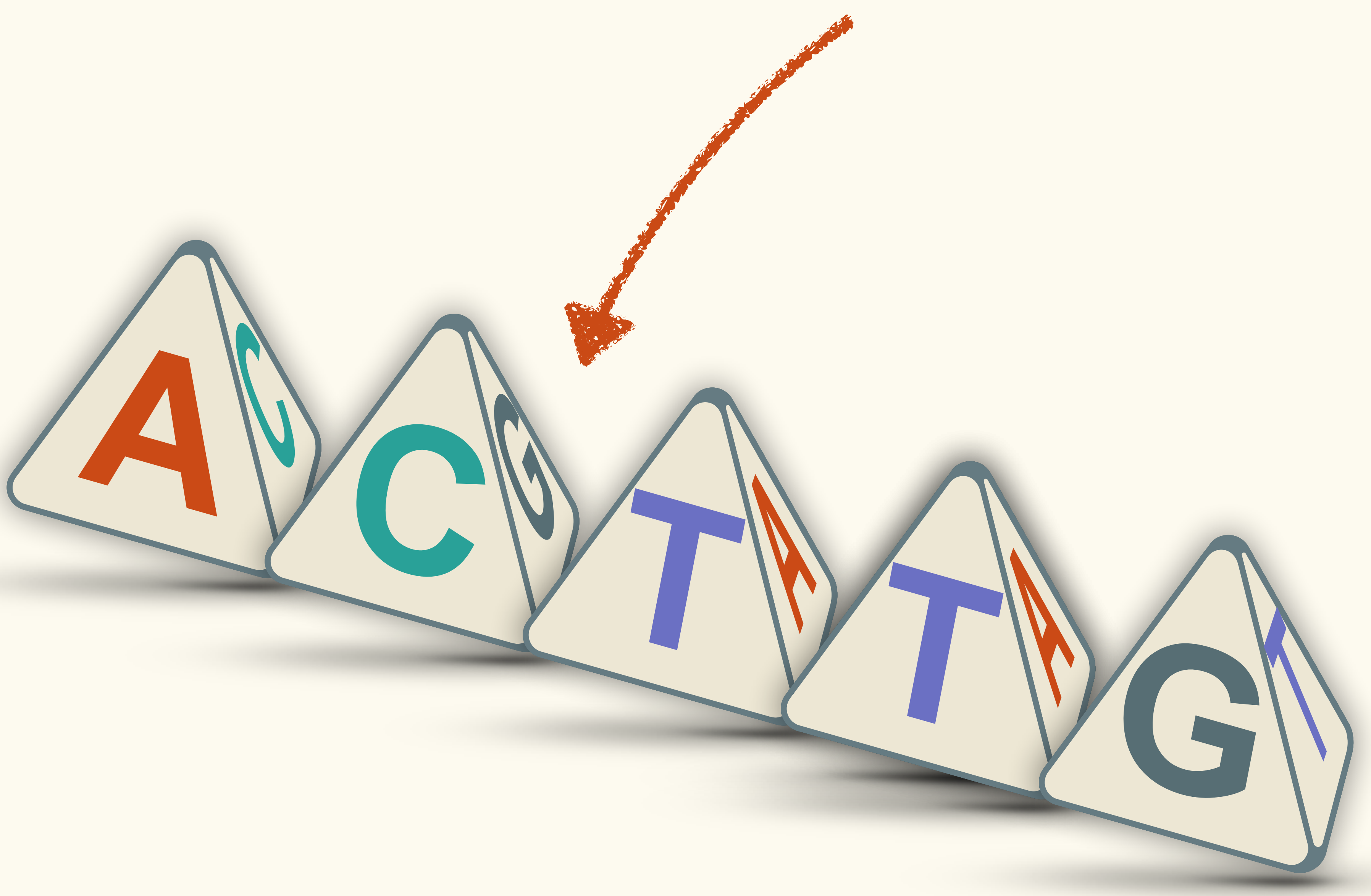
$$P\left(\begin{matrix}\text{ACTTG}\\\text{ACTGG}\end{matrix}\right) = ?$$

Probability



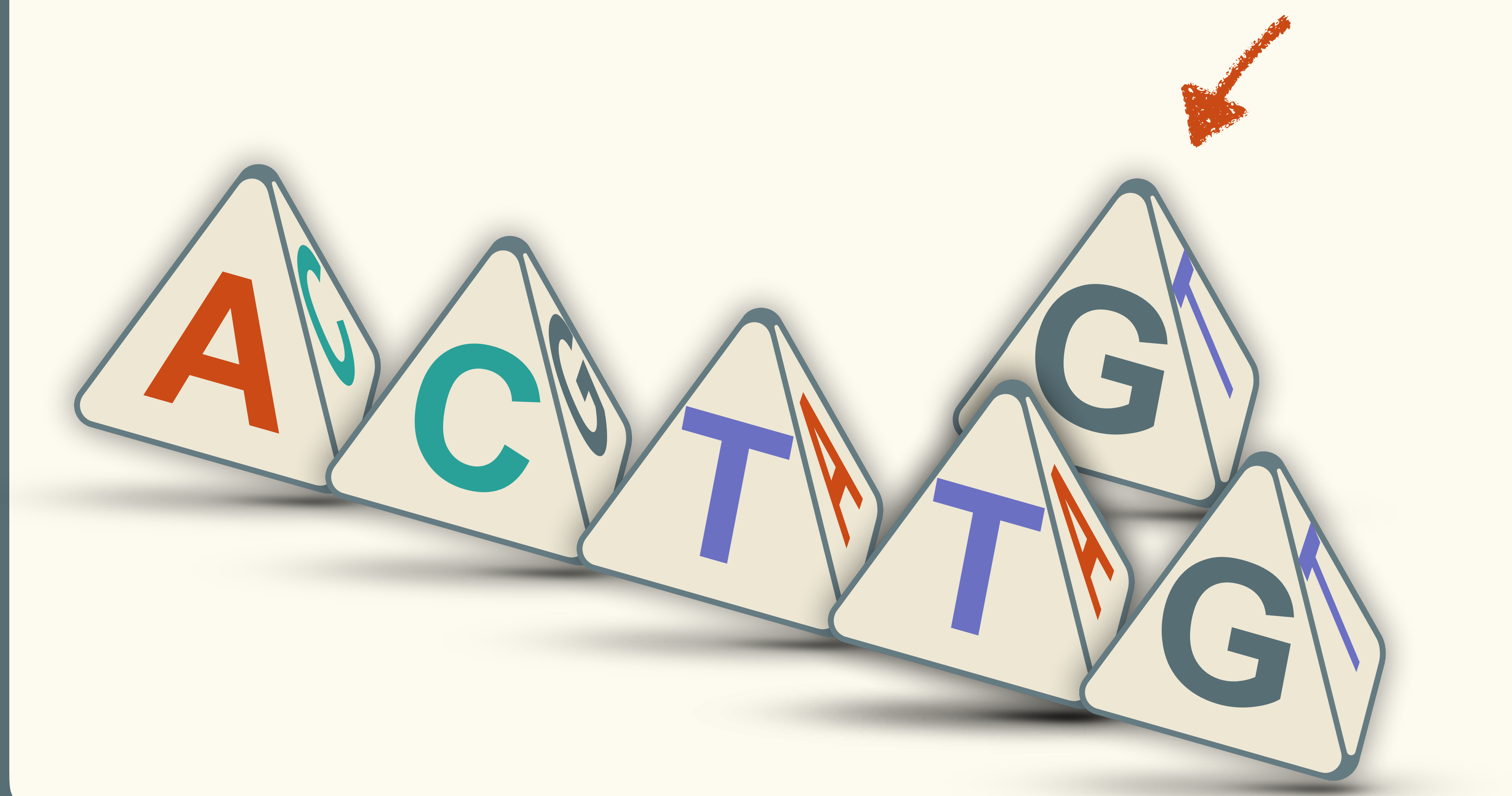
Probability

0.0010



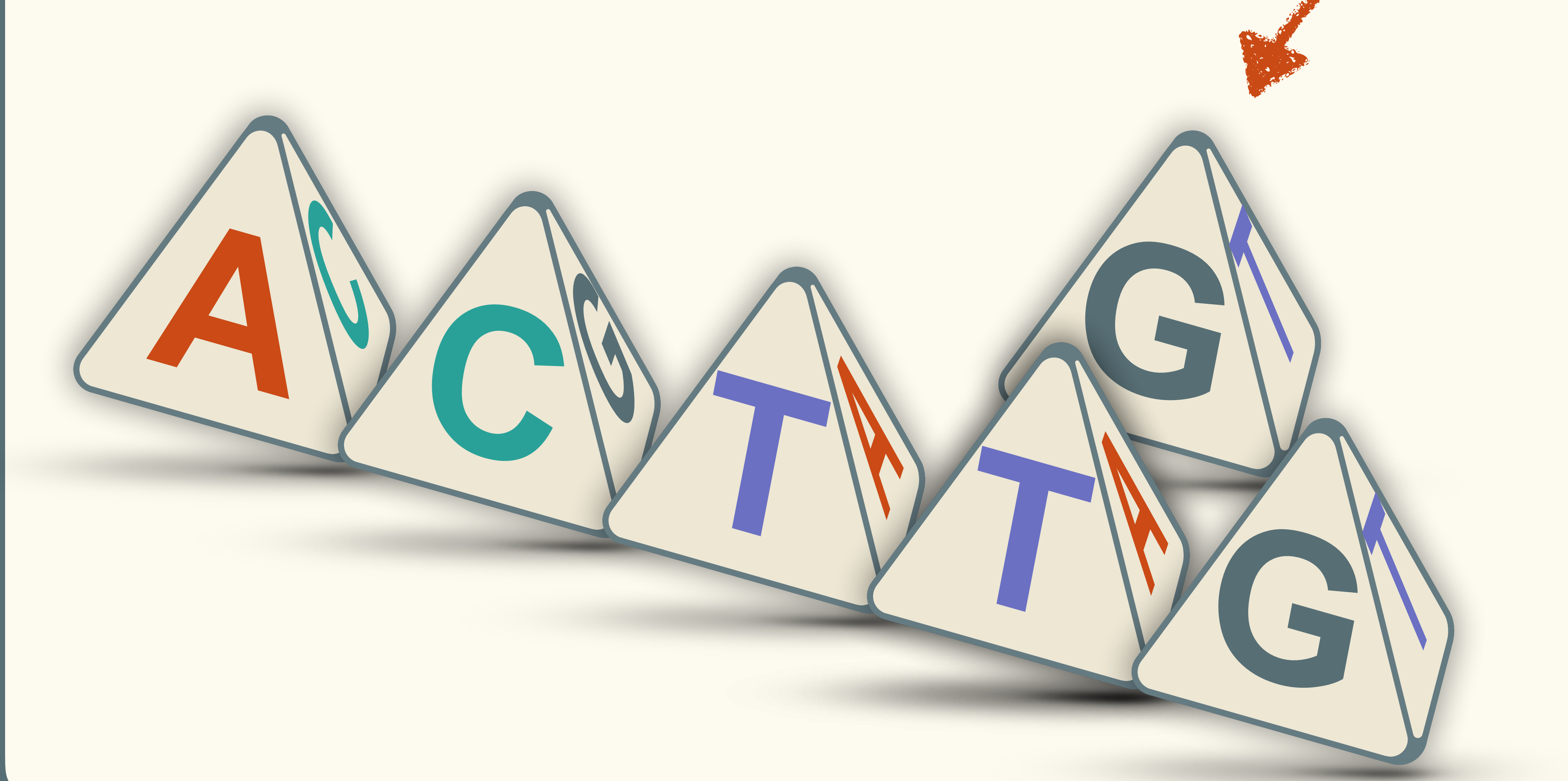
Probability

$$\times P(\mathrm{T} \rightarrow \mathrm{G})$$



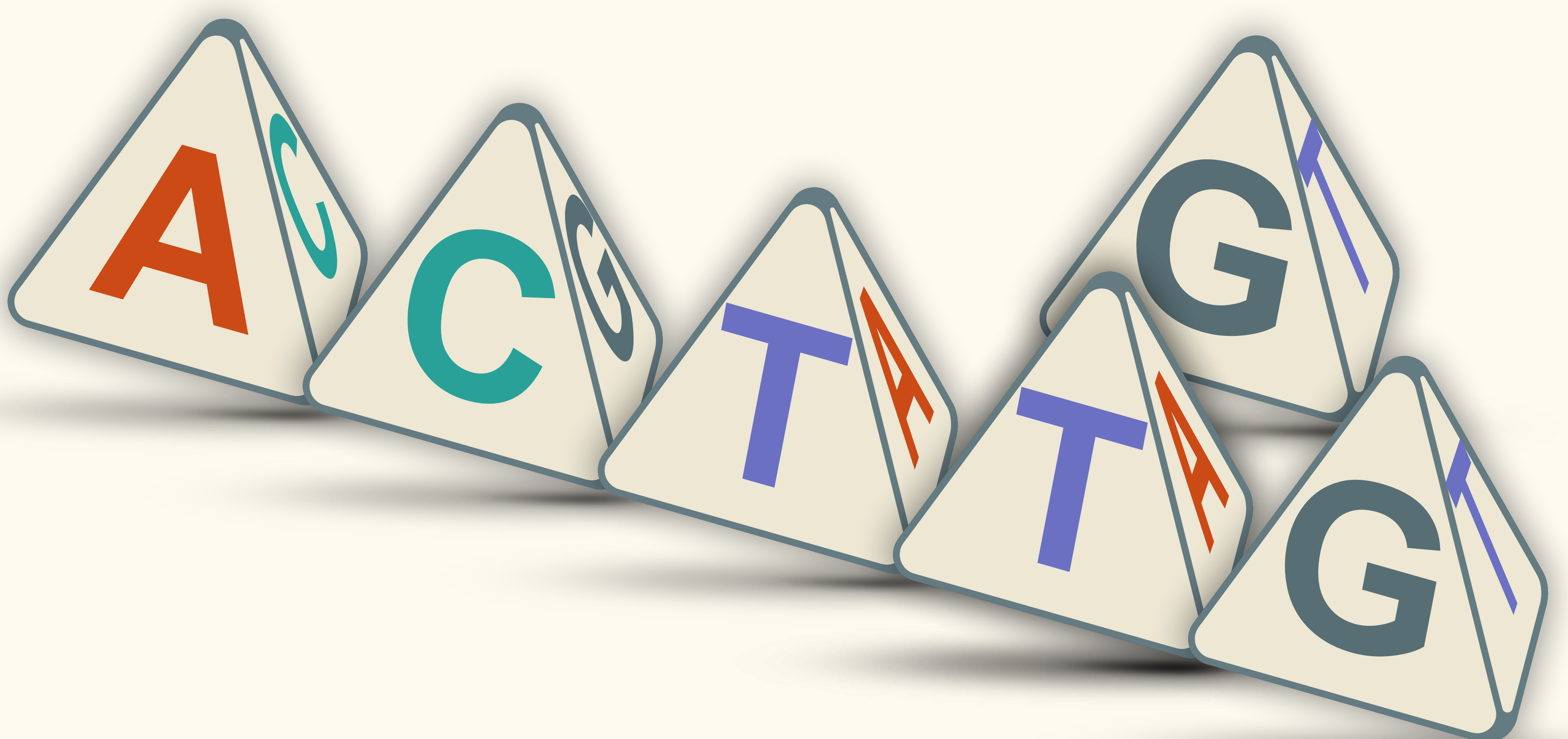
Probability

$\times P(T \to G \mid$  $)$

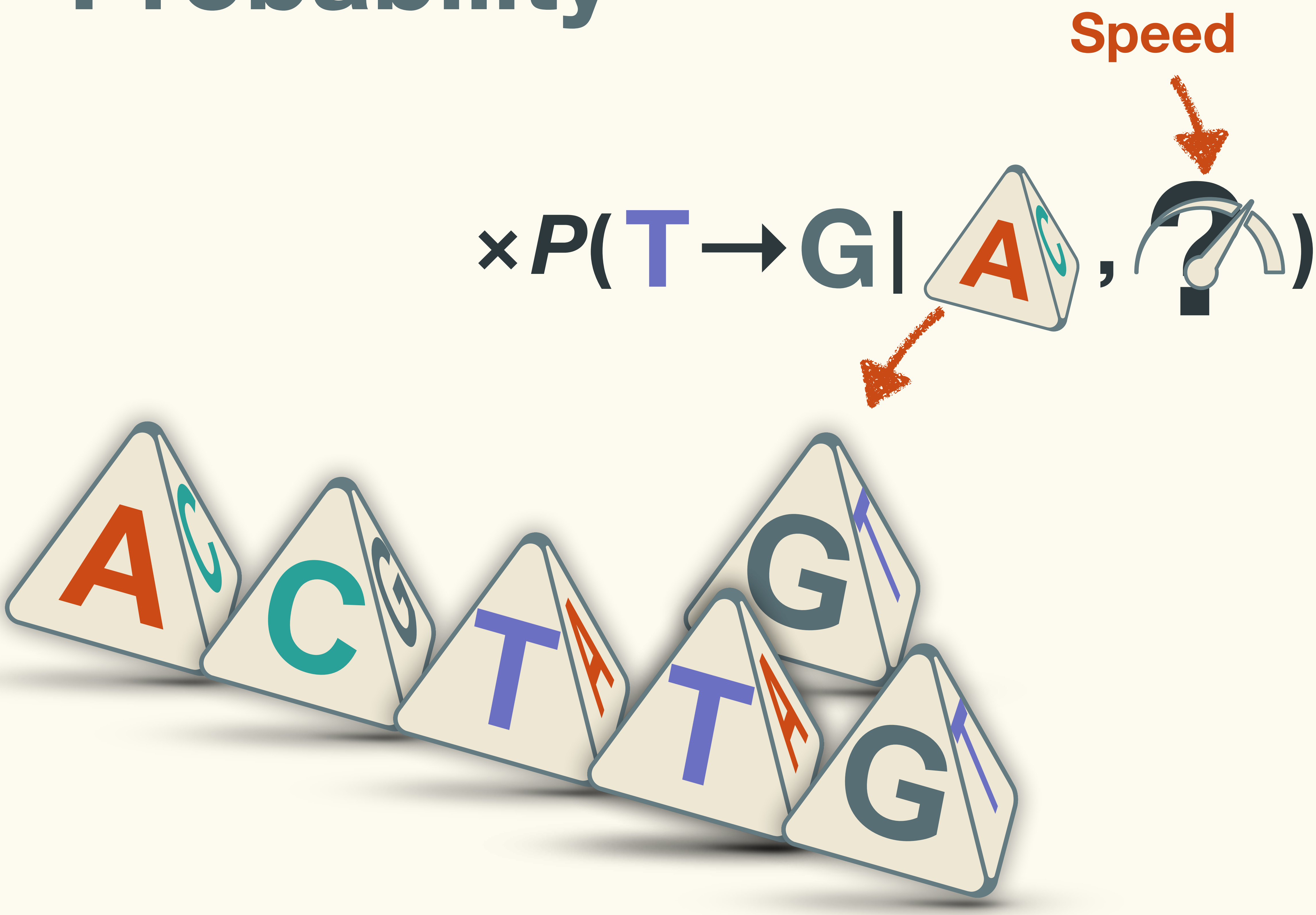


Probability

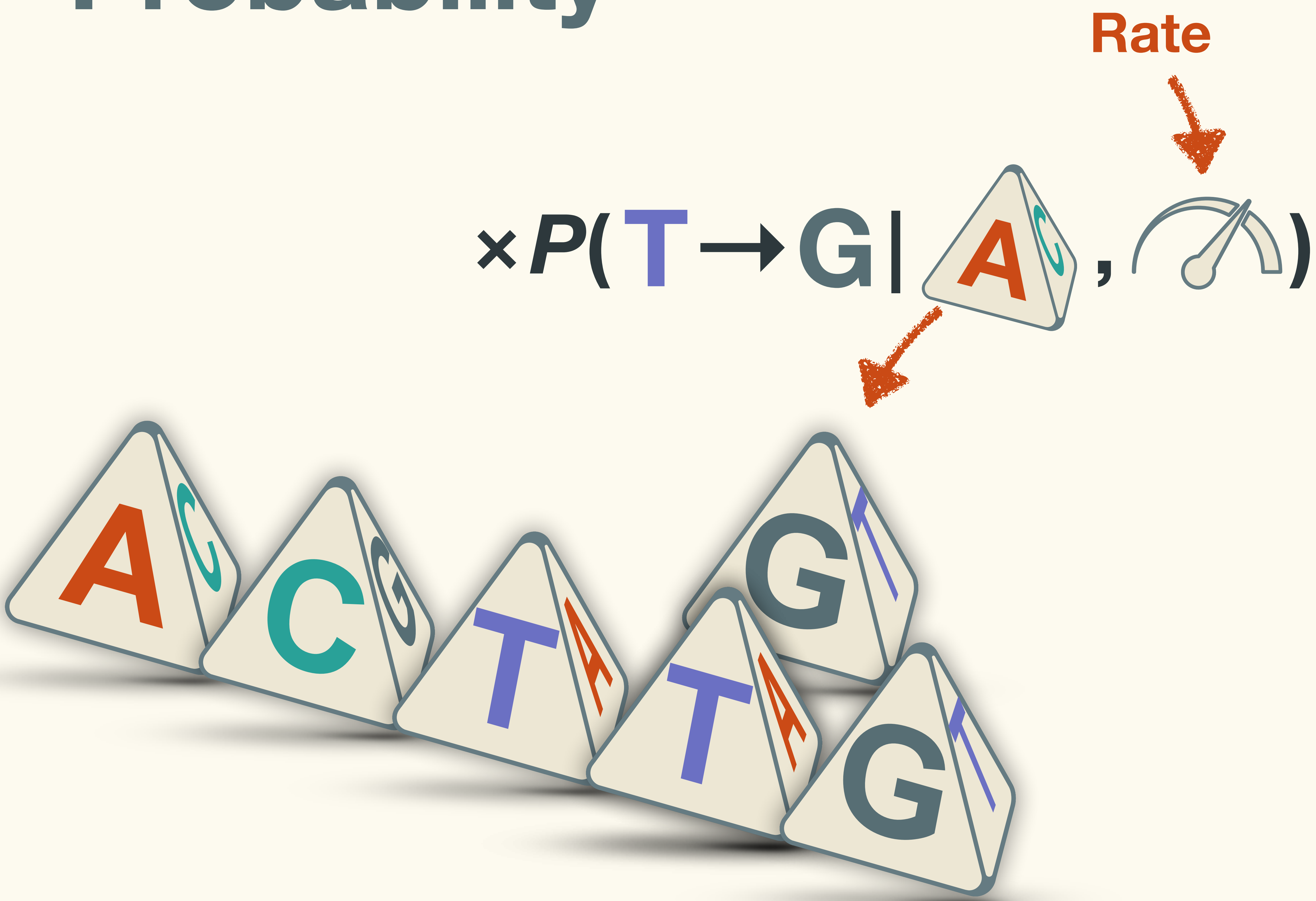
$\times P(\mathrm{T} \rightarrow \mathrm{G} \mid$  $, \; ? \;)$



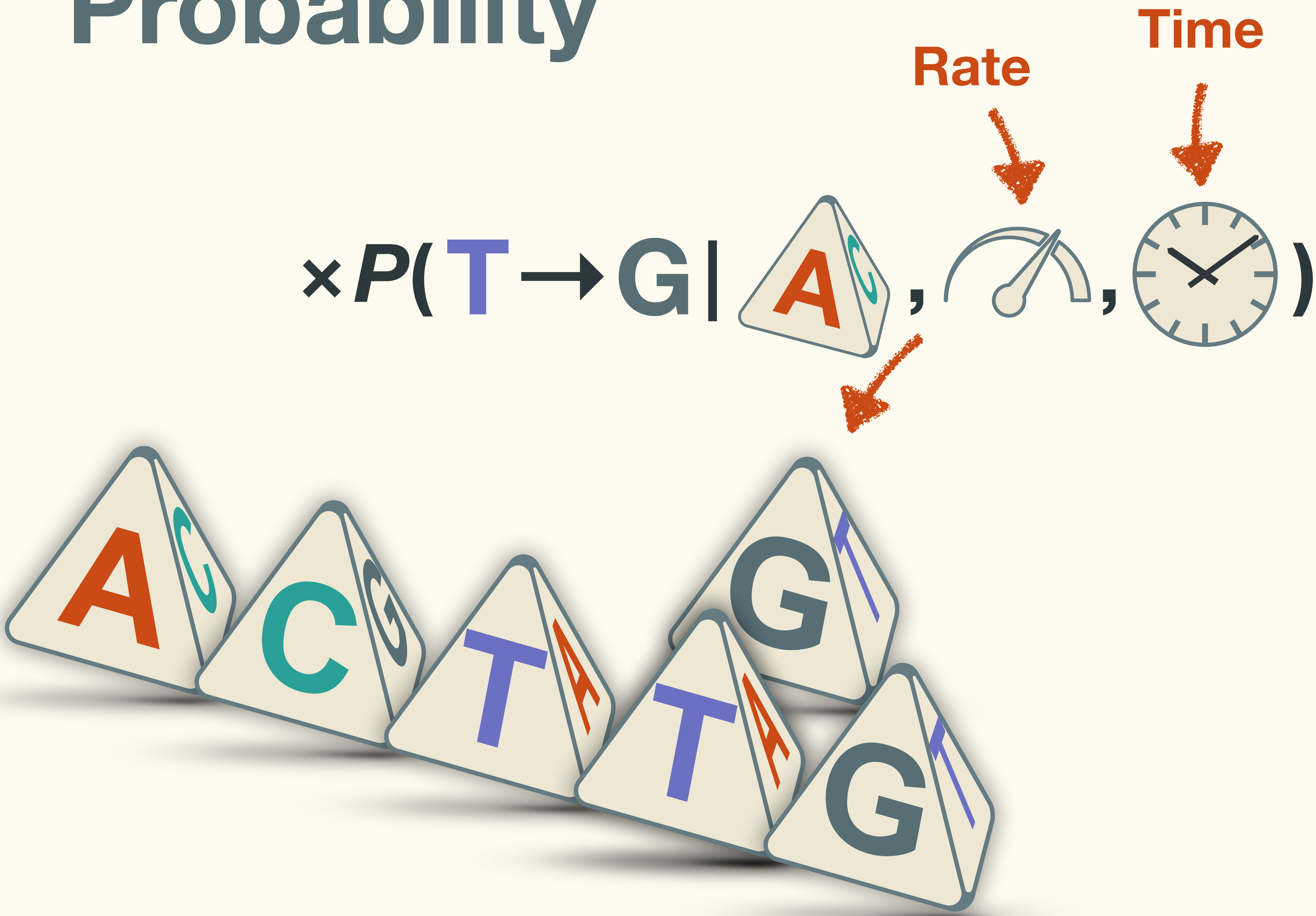
Probability



Probability



Probability



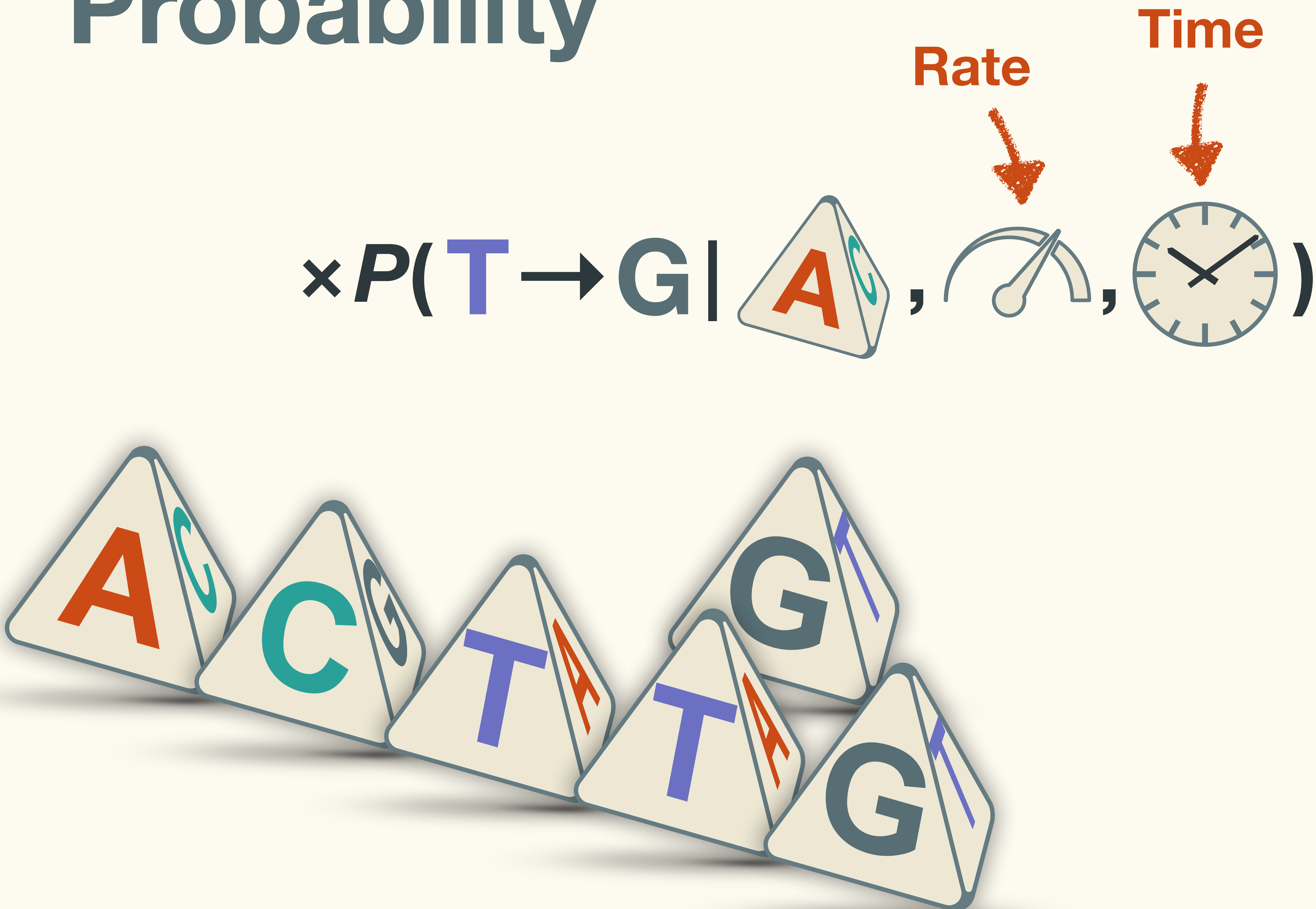
Probability

Time

Rate

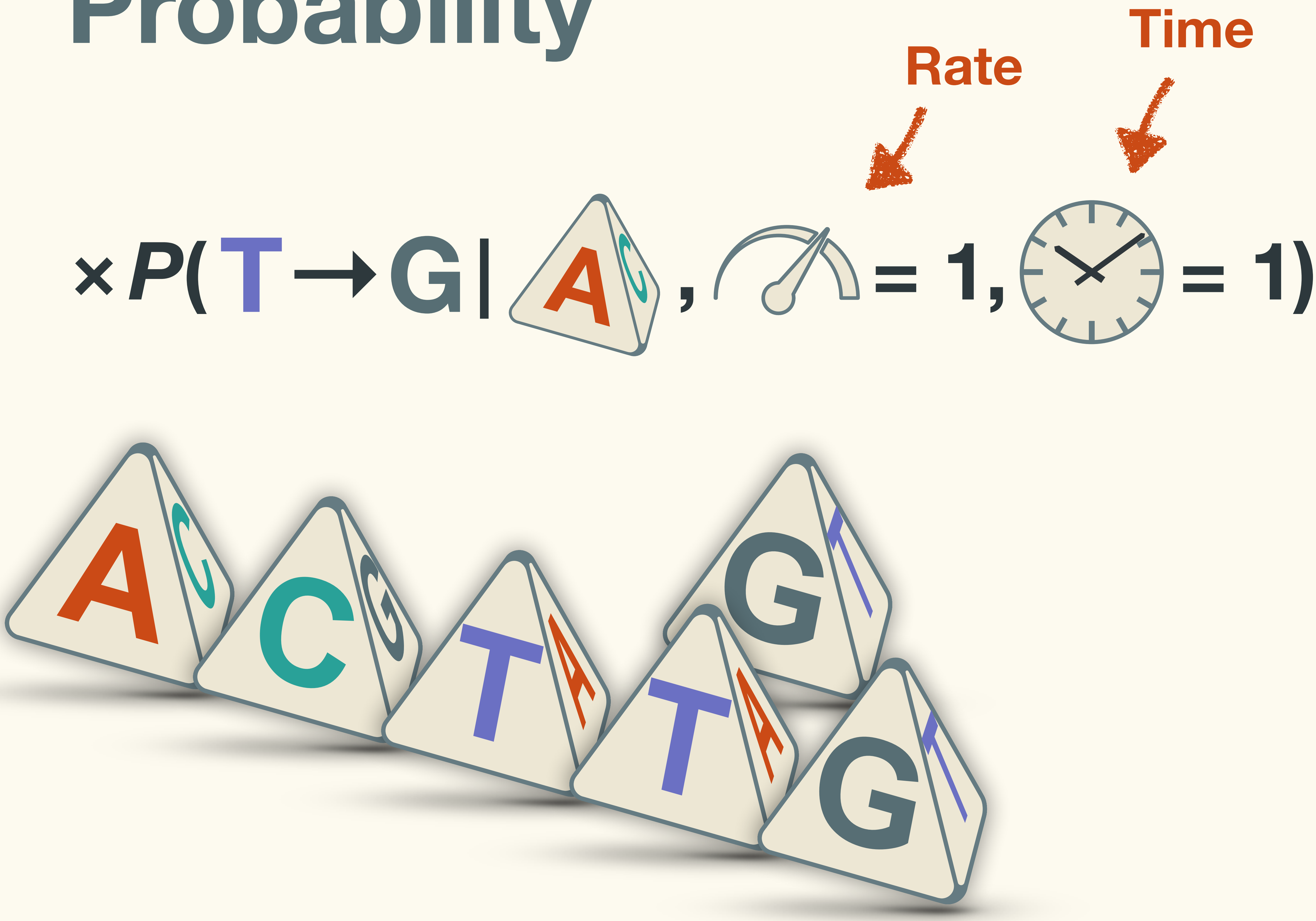
$$({}^{1}/_{4}) \times (1-e^{-4\alpha t})$$

Probability

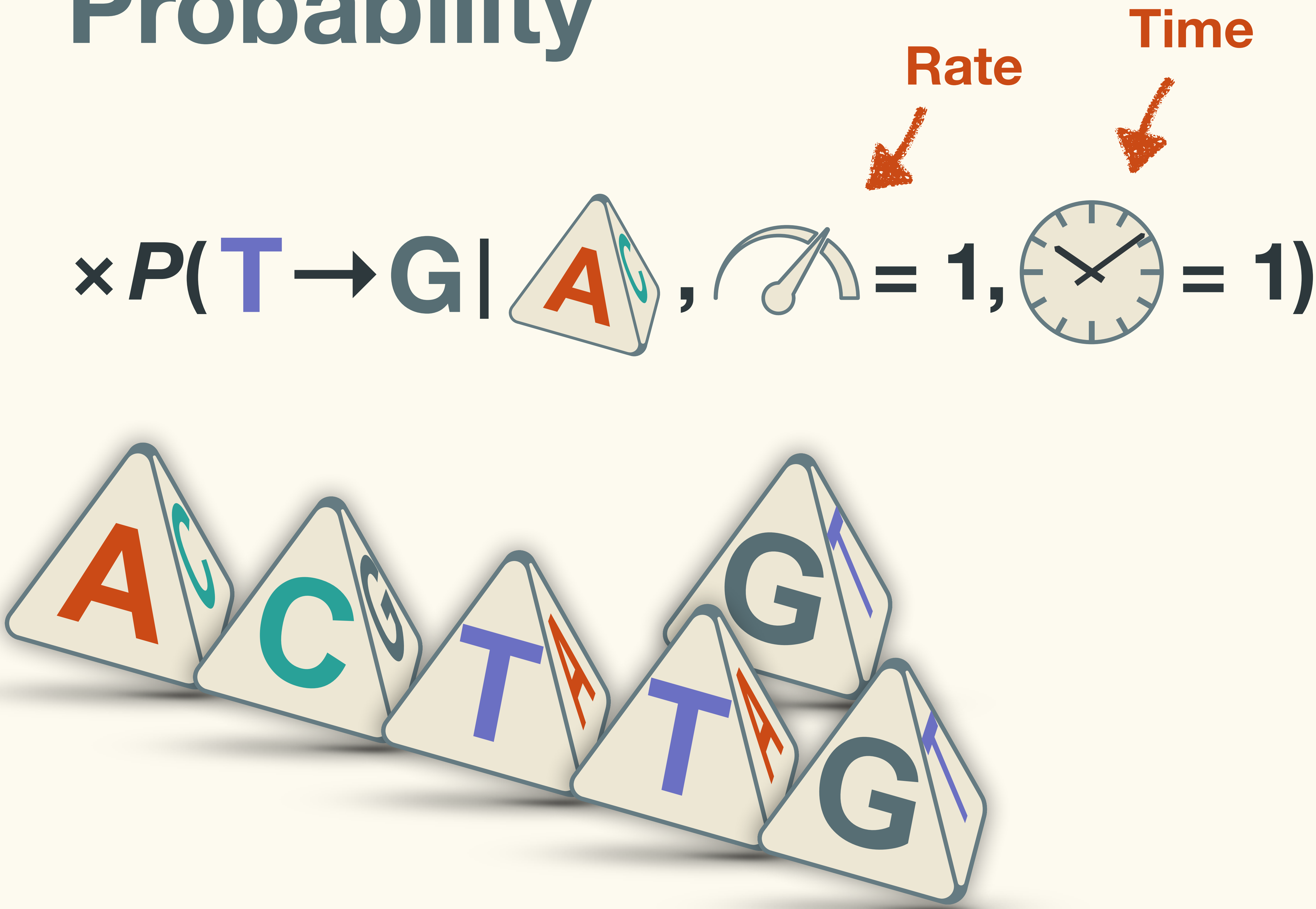


Probability

Probability

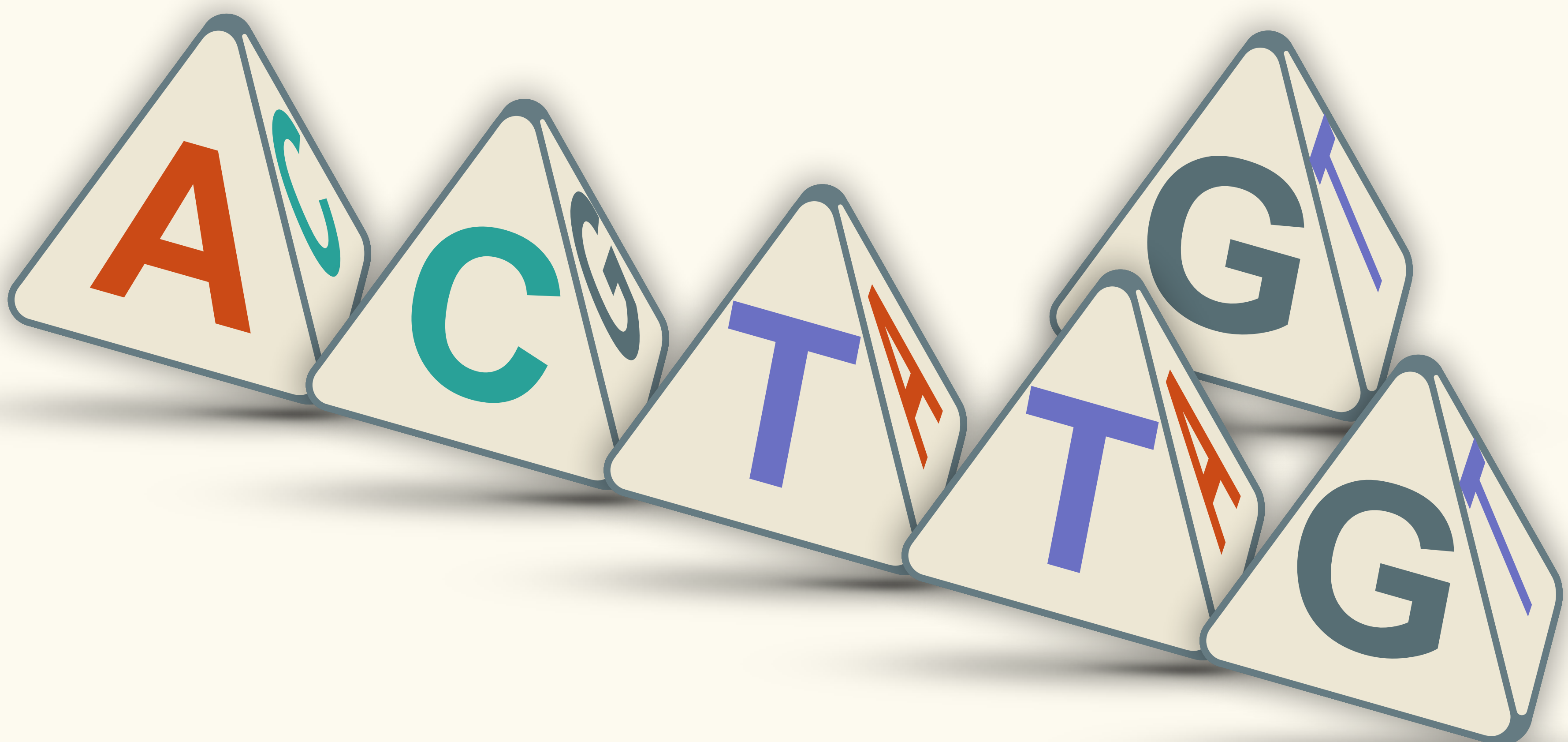


Probability



Probability

$$\times P(\mathrm{T} \rightarrow \mathrm{G} \mid \boxed{\mathrm{M}_1})$$



Probability

Time

Rate

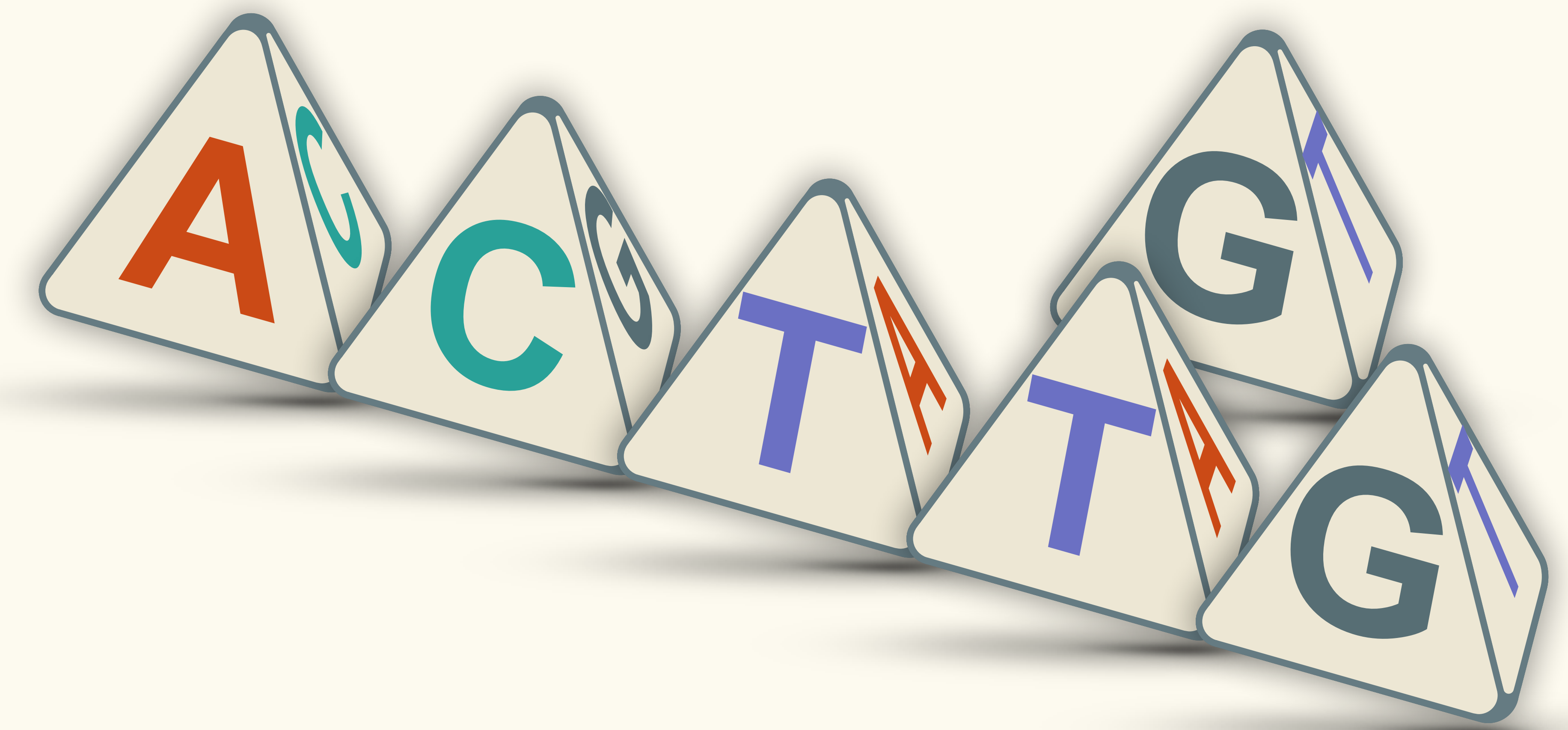
$$\left({}^{1}\!/_{4}\right) \times \left(1 - e^{-4\alpha t}\right)$$

Probability

Time

Rate

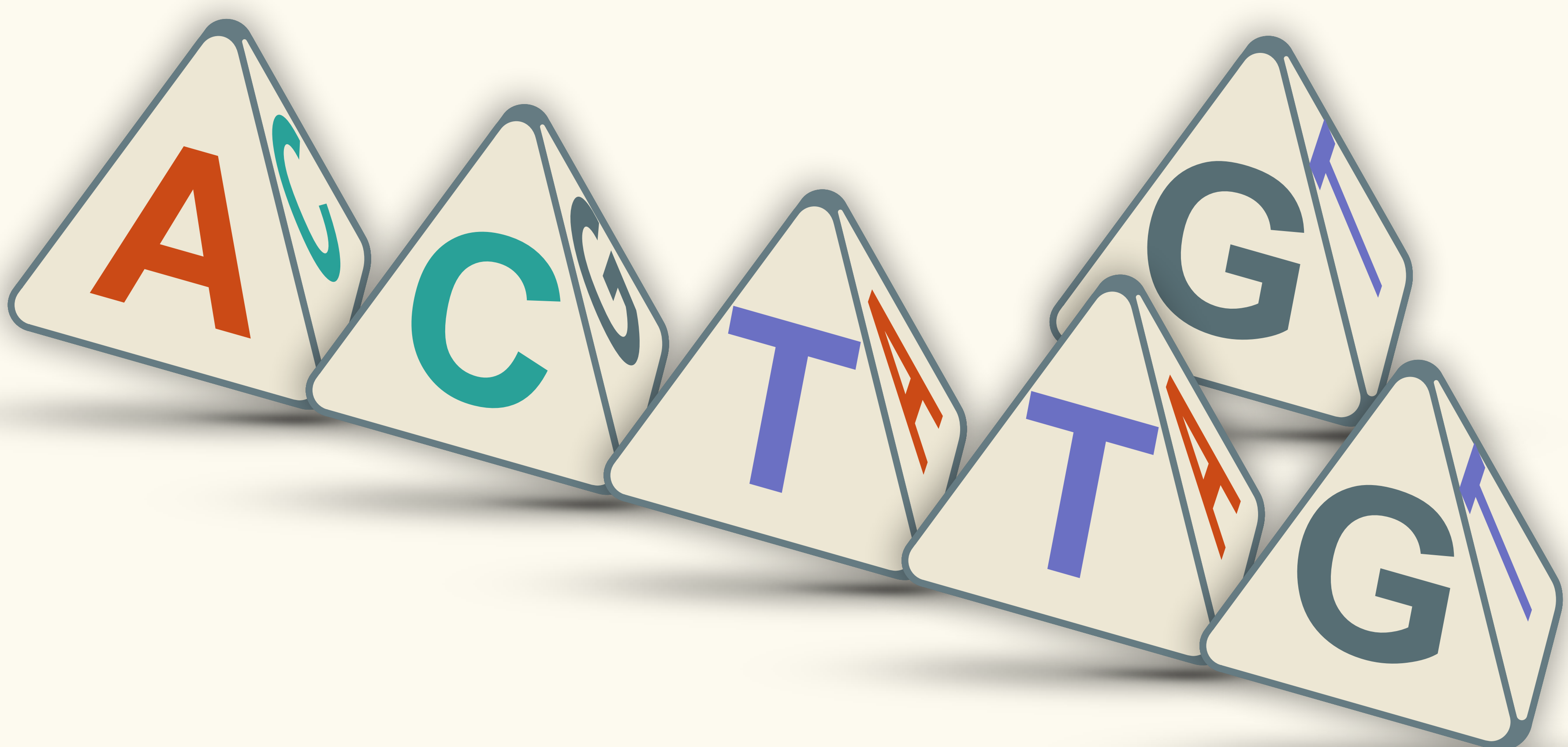
$$(^1/_4) \times (1 - e^{-4\alpha t})$$



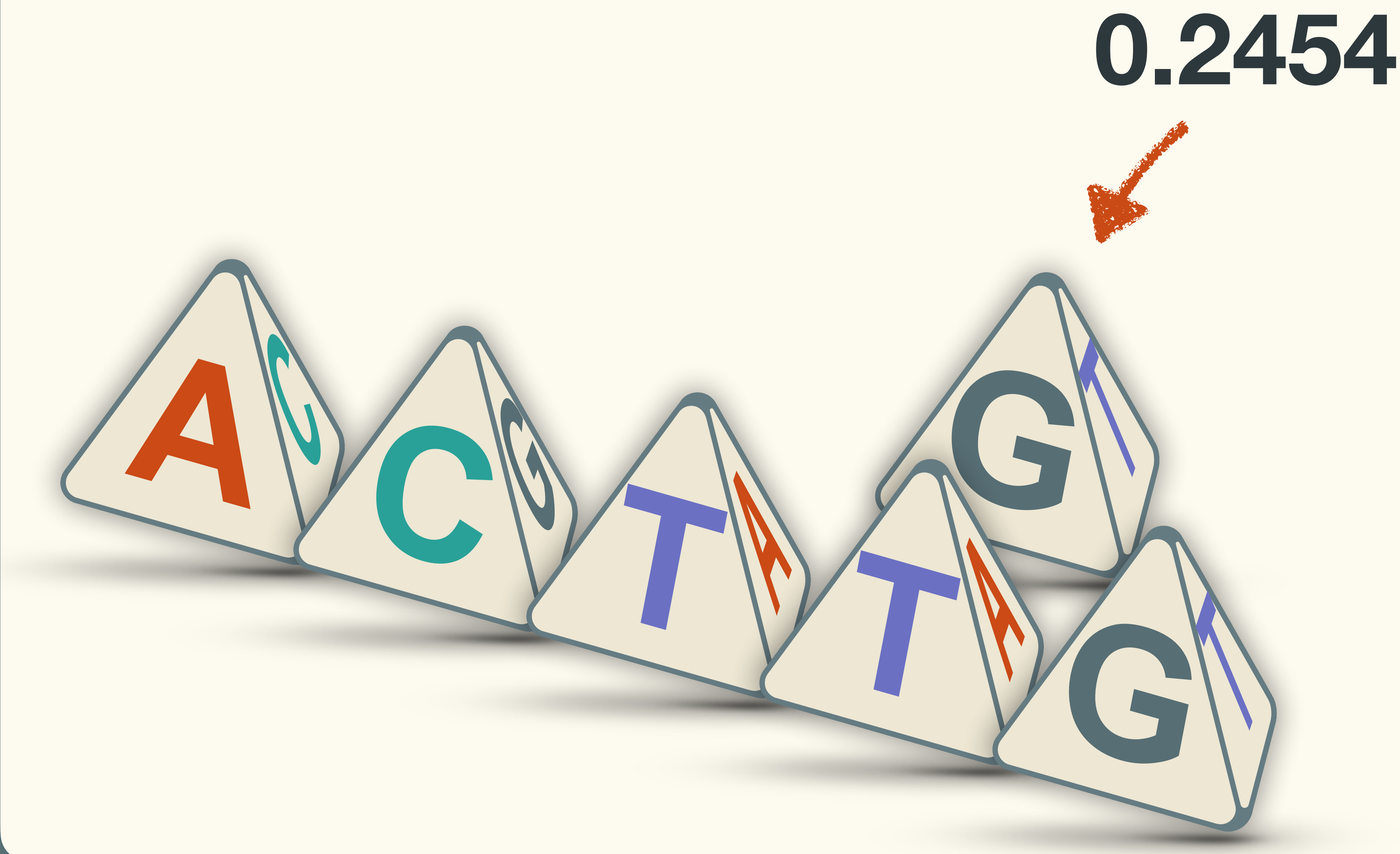
Probability

Time

$$(^1/_4) \times (1-e^{-4\ \ t})$$



Probability

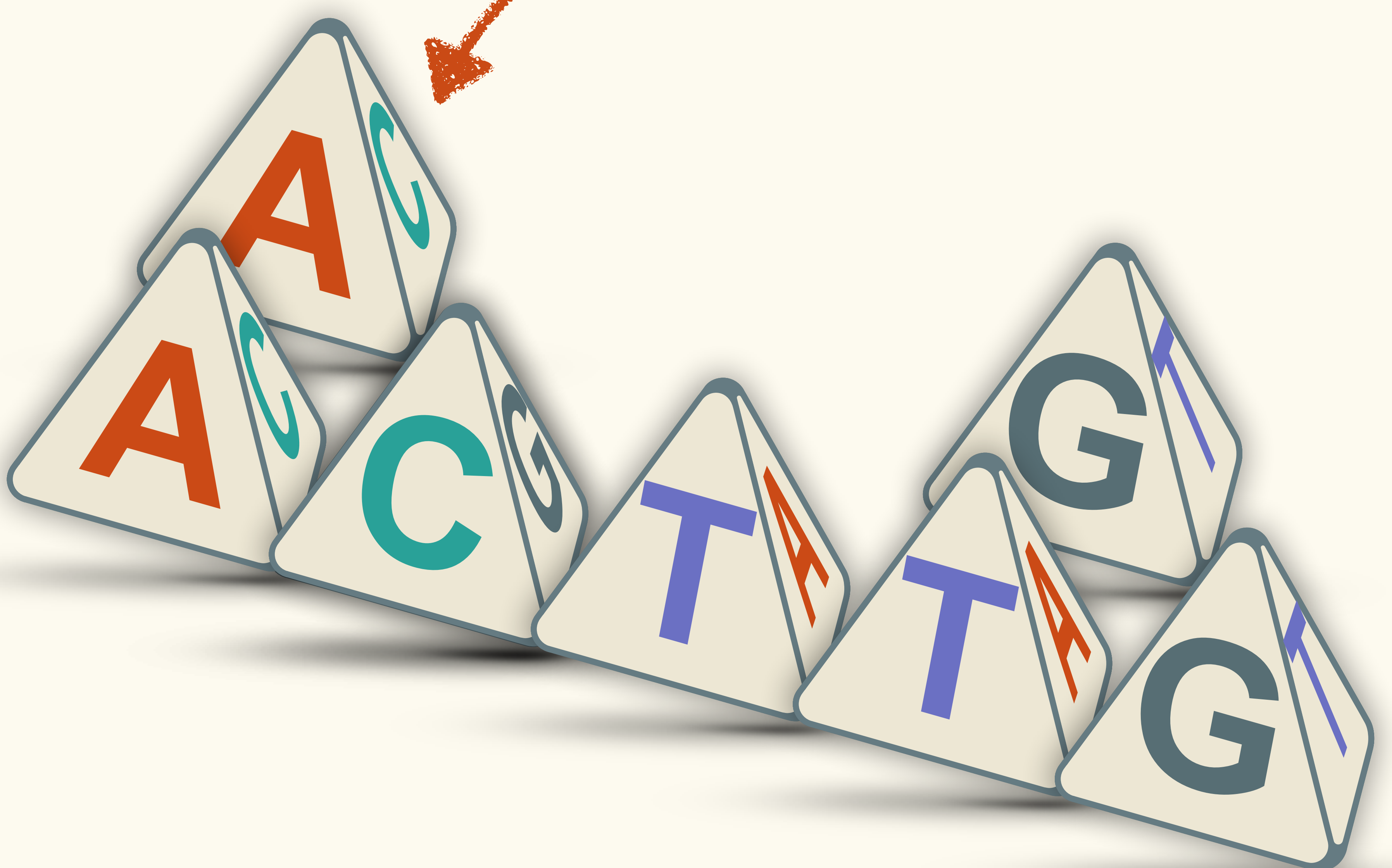


Probability

$$P\left(\begin{matrix}\text{ACTTG}\\ \text{ACTGG}\end{matrix} \,\middle|\, \boxed{M}_1\right) = ?$$

Probability

$$\times P(A \rightarrow A \mid \boxed{M_1})$$



Probability

$$\times P(C \rightarrow C \mid \boxed{M_1})$$

Probability

$$\times P(\mathrm{T} \rightarrow \mathrm{T} \mid \boxed{M_1})$$

Probability

$$\times P(G \to G \mid \boxed{M_1})$$

Probability

$$\times P(\mathrm{A} \rightarrow \mathrm{A} \mid \boxed{M_1})$$

Probability

$$(^1/_4) \times (1 - e^{-4\alpha t}) + e^{-4\alpha t}$$

Probability

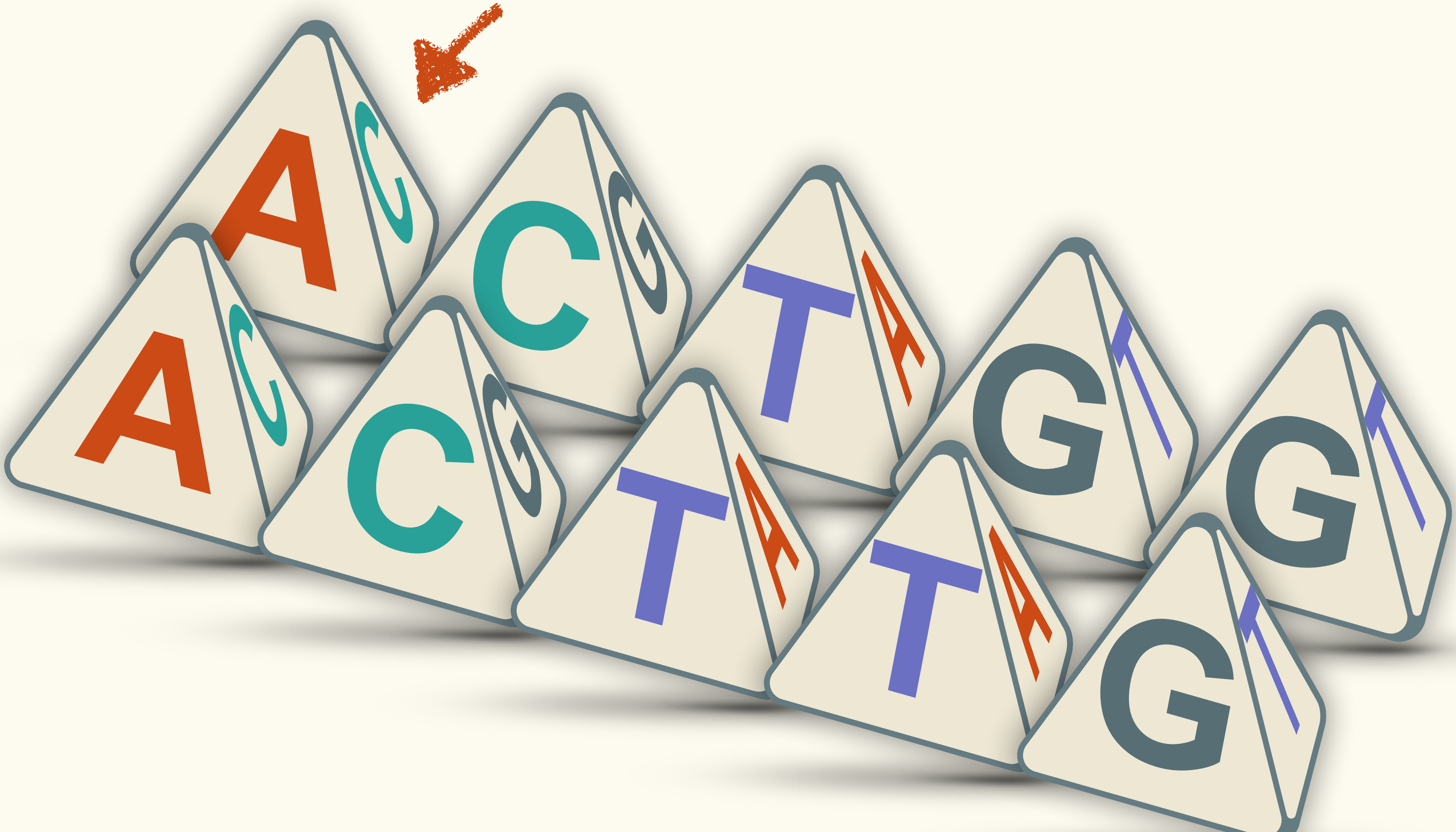
Rate

$$(^1/_4) \times (1 - e^{-4\alpha t}) + e^{-4\alpha t}$$

Probability

Time

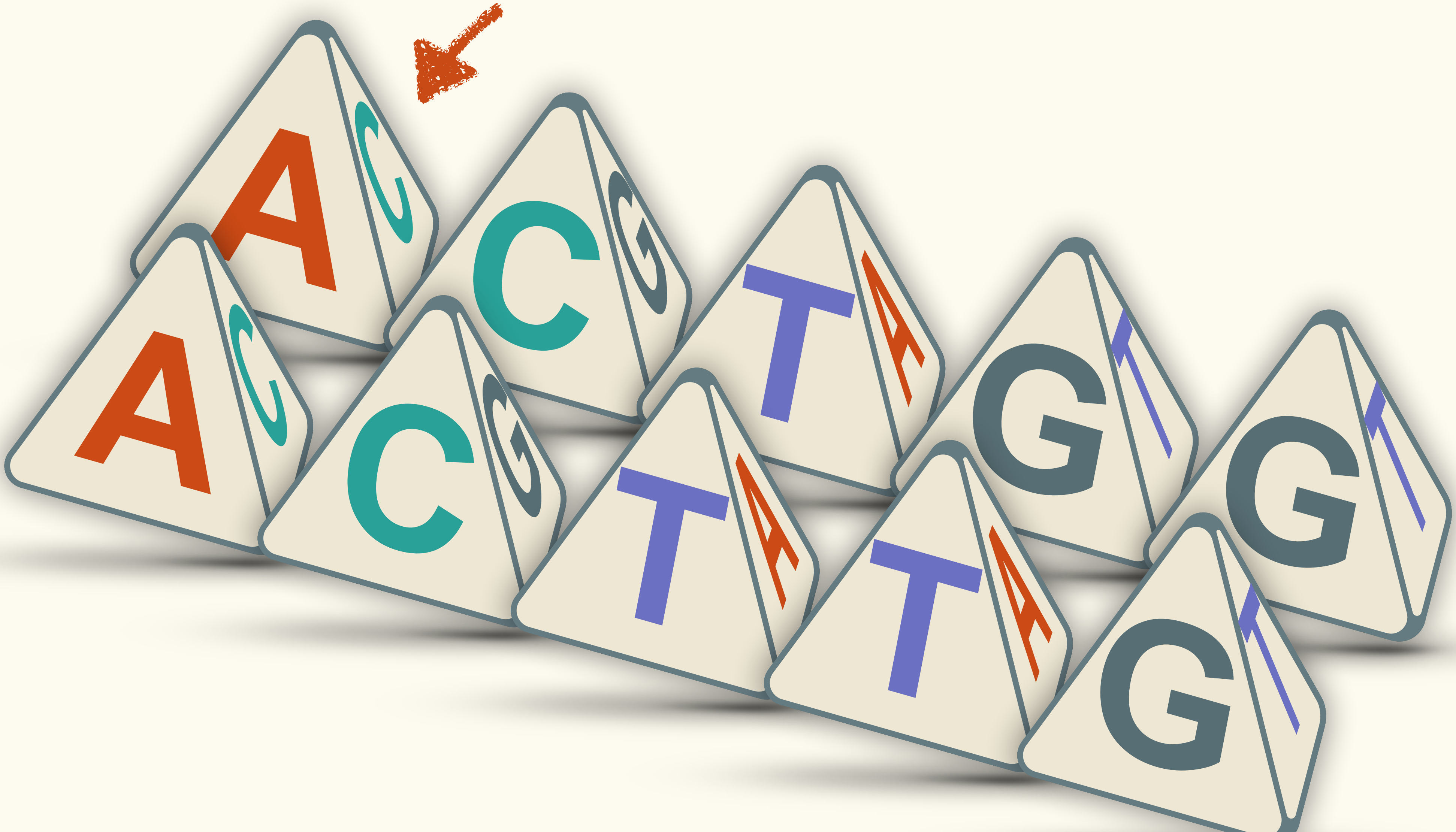
$$(^{1}/_{4}) \times (1-e^{-4\alpha t}) + e^{-4\alpha t}$$



Probability

Time

$$({}^{1}/_{4}) \times (1-e^{-4\alpha t}) + e^{-4\alpha t}$$



Probability

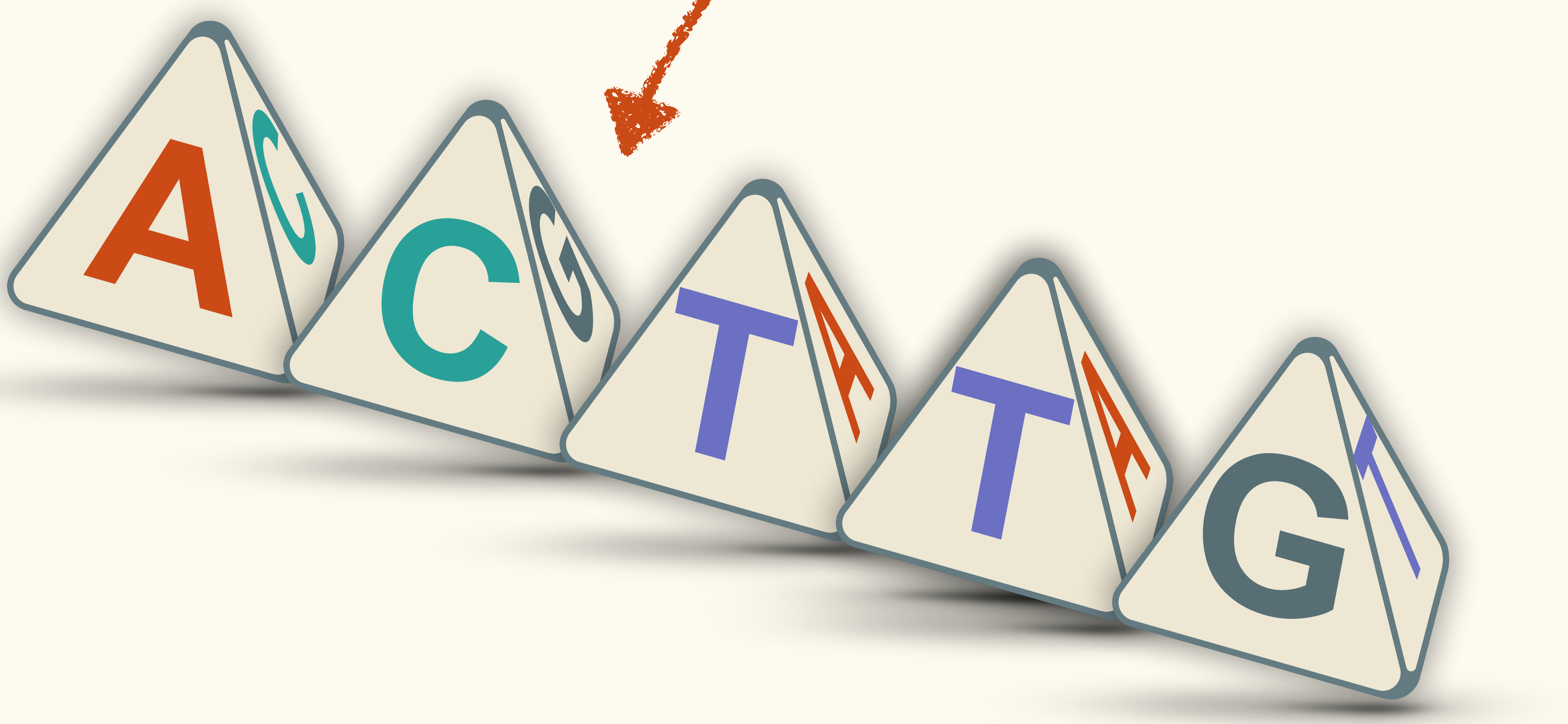
0.2637

Probability

$$P\left(\begin{matrix}\text{ACTTG}\\ \text{ACTGG}\end{matrix} \,\middle|\, \boxed{M_1}\right) = ?$$

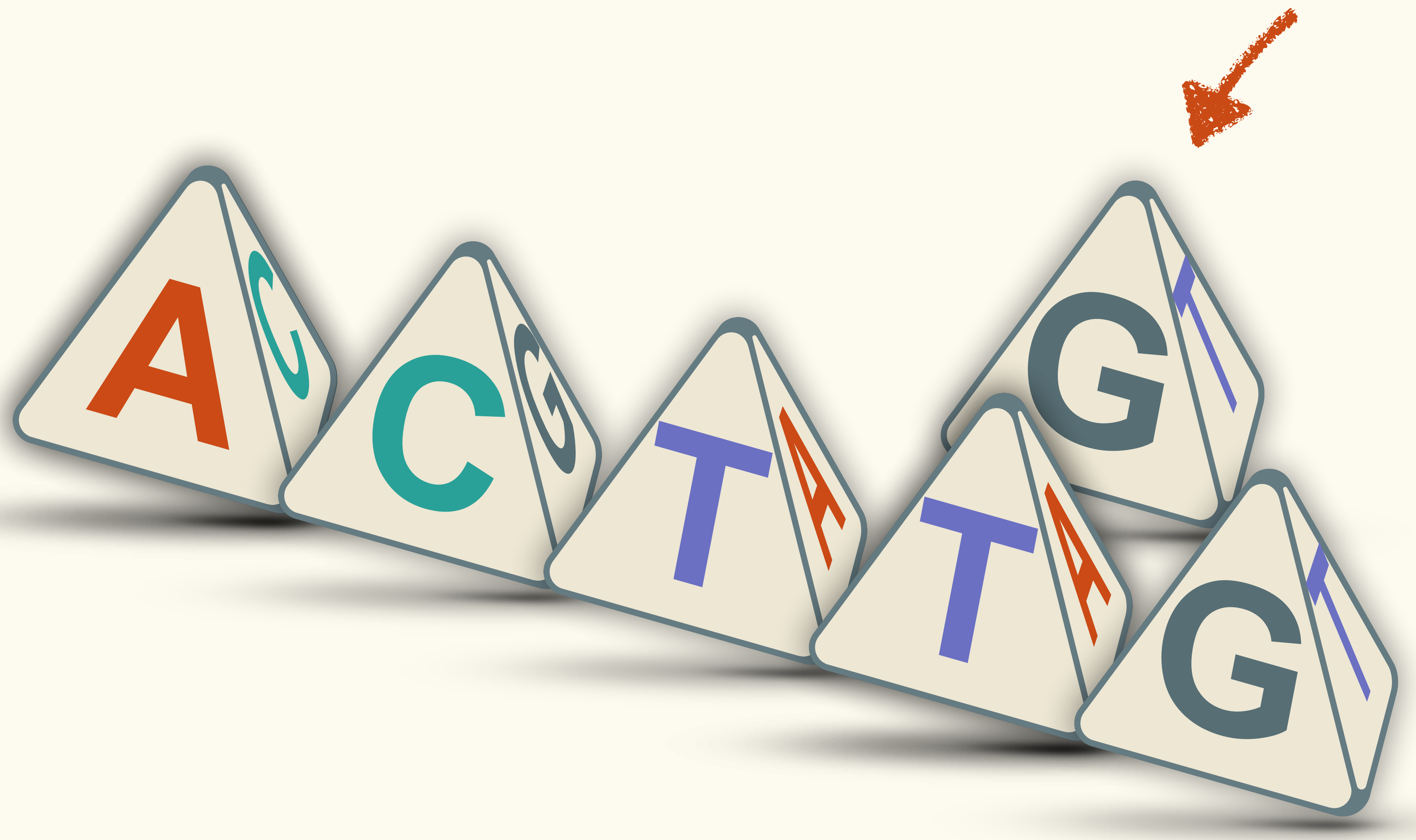
Probability

0.0010



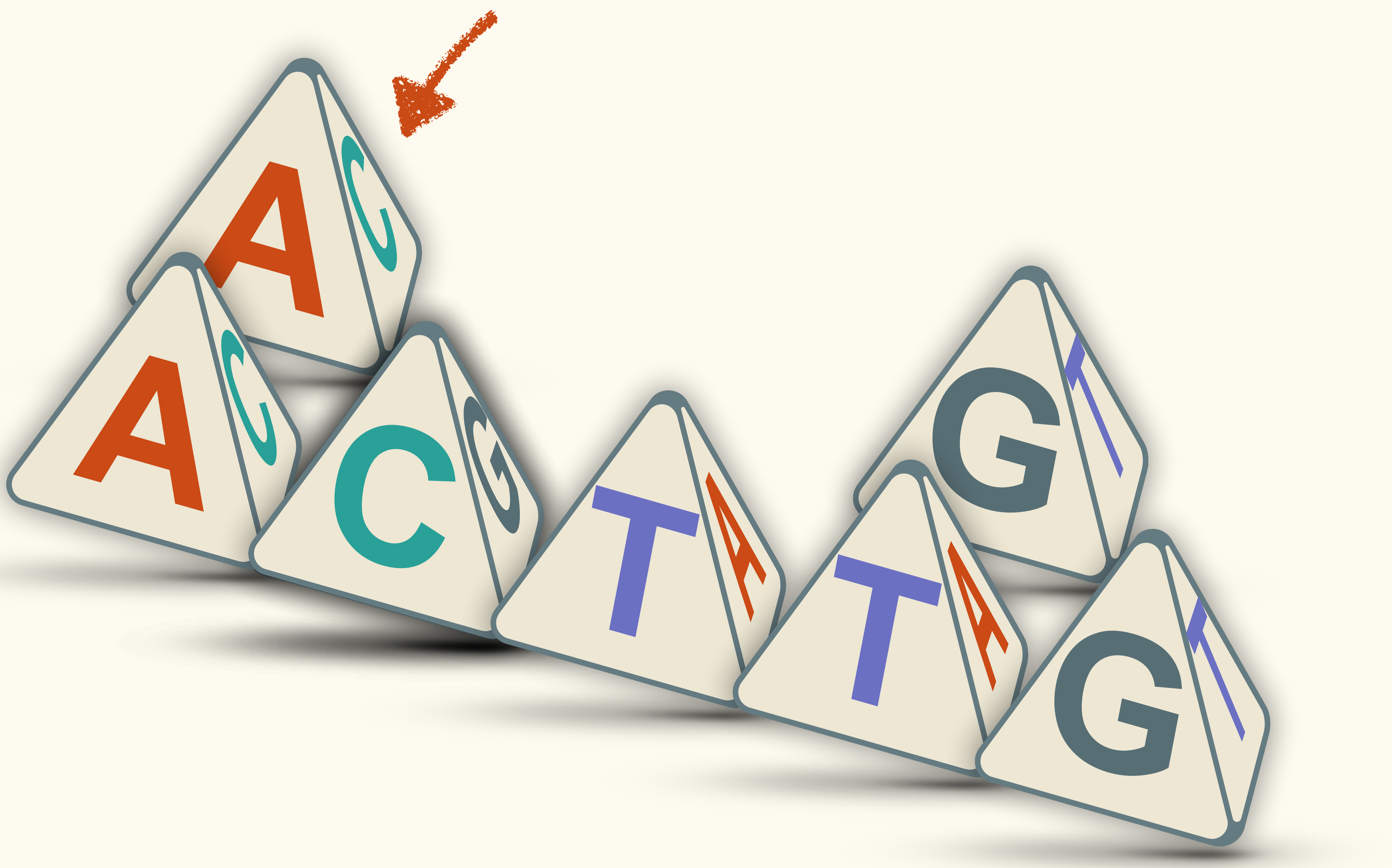
Probability

× 0.2454



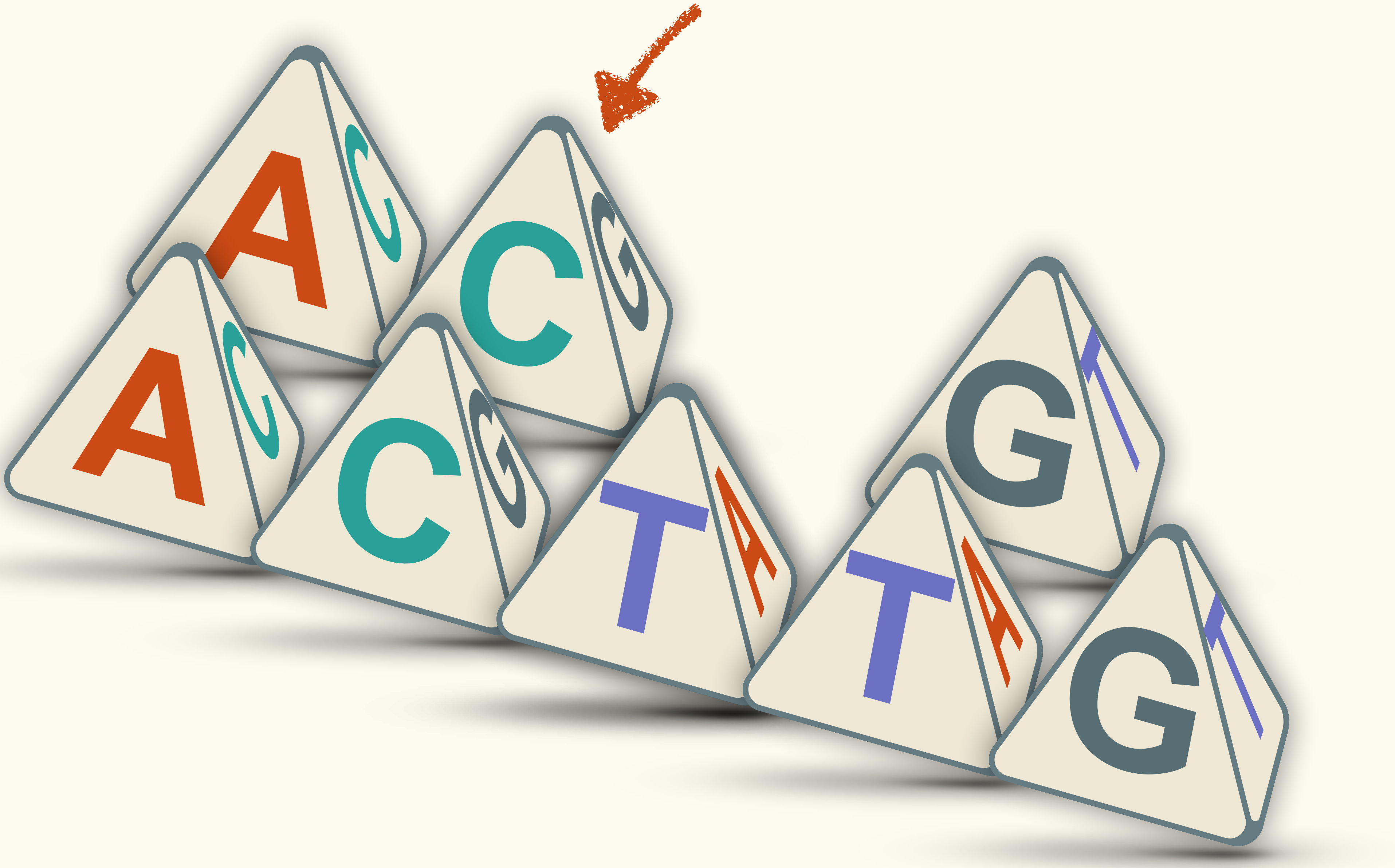
Probability

× 0.2637

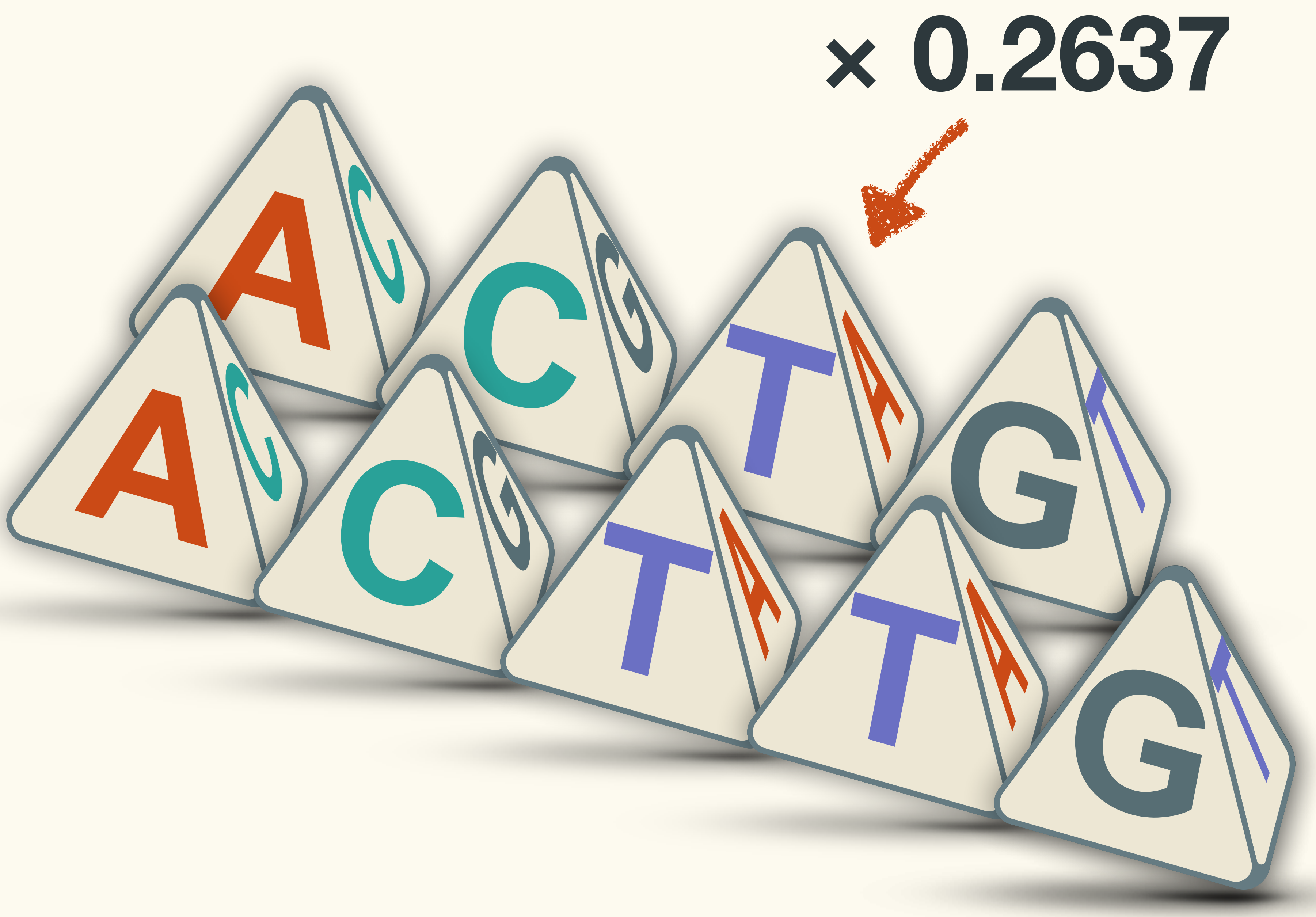


Probability

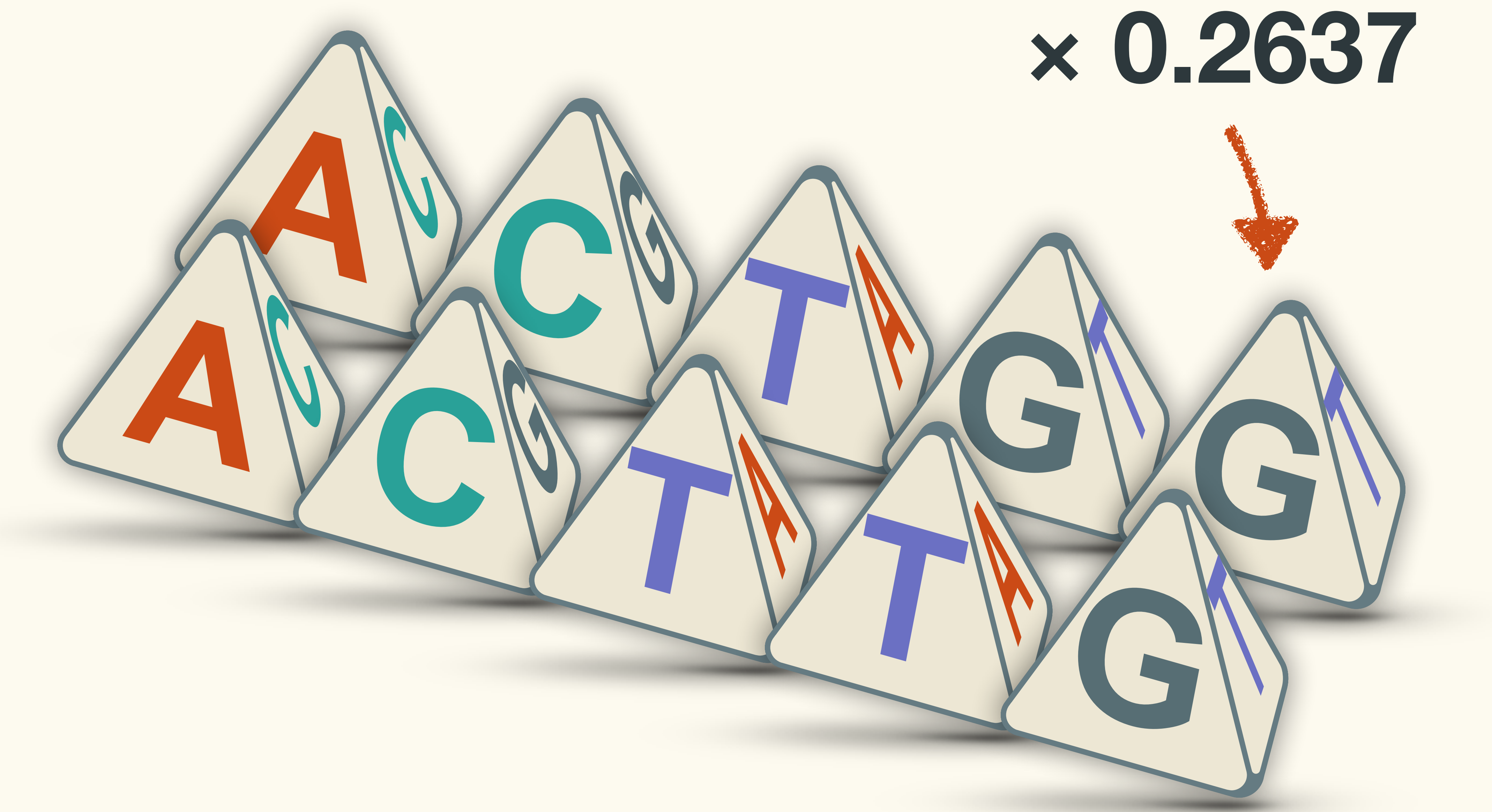
× 0.2637



Probability



Probability



Probability

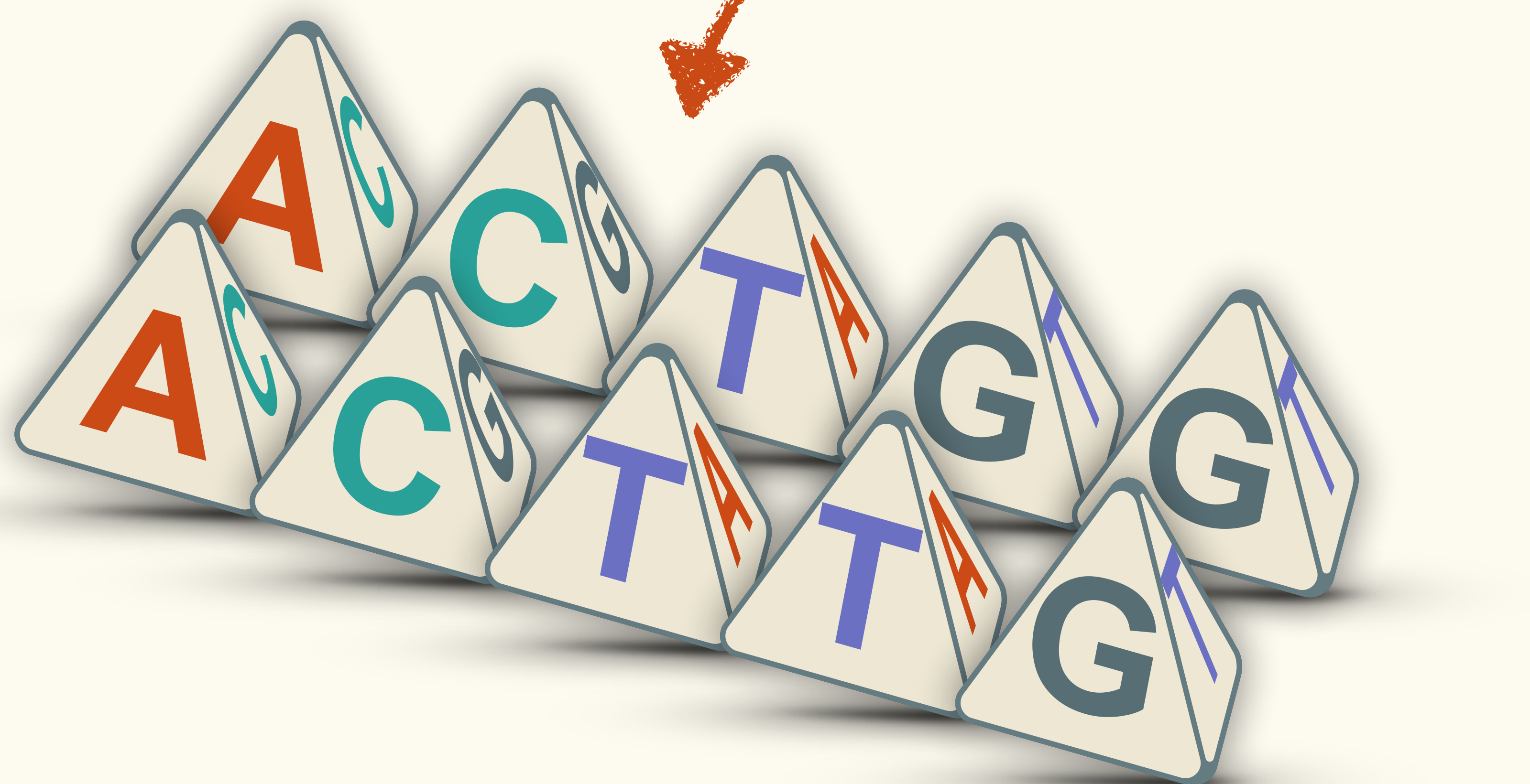
$$P\left(\begin{array}{c}\text{ACTTG}\\\text{ACTGG}\end{array} \,\middle|\, \boxed{M_1}\right) = 0.0000015$$

Likelihood

$$L\left(\boxed{M_1} \,\middle|\, \begin{matrix}\text{ACTTG}\\ \text{ACTGG}\end{matrix}\right) = 0.0000015$$

Likelihood

$$\log\left(L\left(\boxed{M_1} \,\middle|\, \begin{matrix}\text{ACTTG}\\ \text{ACTGG}\end{matrix}\right)\right) = -13.4$$

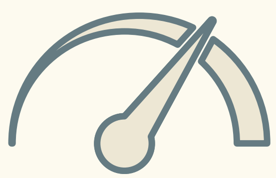


Likelihood

$$\log\left(L\left(\boxed{M}_1 \mid \begin{matrix}\text{ACTTG}\\ \text{ACTGG}\end{matrix}\right)\right) = -13.4$$

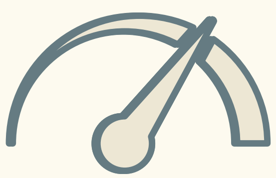
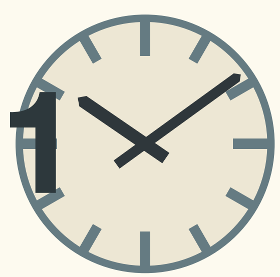
Likelihood

$$\log\left(L\left(\boxed{M}_1 \,\middle|\, \begin{matrix}\text{ACTTG}\\ \text{ACTGG}\end{matrix}\right)\right) = -13.4$$

 ,  = 1,  = 1

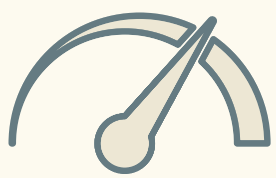
Likelihood

$$\log\left(L\left(\boxed{M_1} \,\middle|\, \begin{matrix}\text{ACTTG}\\ \text{ACTGG}\end{matrix}\right)\right) = -13.4$$

 ,  = 0,1  = 1

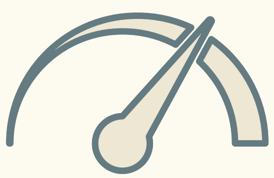
Likelihood

$$\log\left(L\left(\boxed{M_2} \,\middle|\, \begin{matrix}\text{ACTTG}\\ \text{ACTGG}\end{matrix}\right)\right) = -13.4$$

,  = 0.1,  = 1

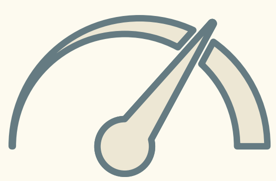
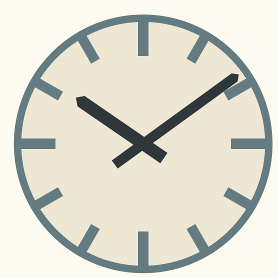
Likelihood

$$\log\left(L\left(\boxed{M}_2 \,\middle|\, \begin{matrix}\text{ACTTG}\\ \text{ACTGG}\end{matrix}\right)\right) = -10.3$$

,  = 0.1,  = 1

Likelihood

$$\log\left(L\left(\boxed{M_2} \,\middle|\, \begin{matrix}\text{ACTTG}\\ \text{ACTGG}\end{matrix}\right)\right) = -10.3$$

,  = 0.1,  = 1

Likelihood

$$\log\left(L\left(\boxed{M_1} \,\middle|\, \begin{matrix}\text{ACTTG}\\ \text{ACTGG}\end{matrix}\right)\right) = -13.4$$

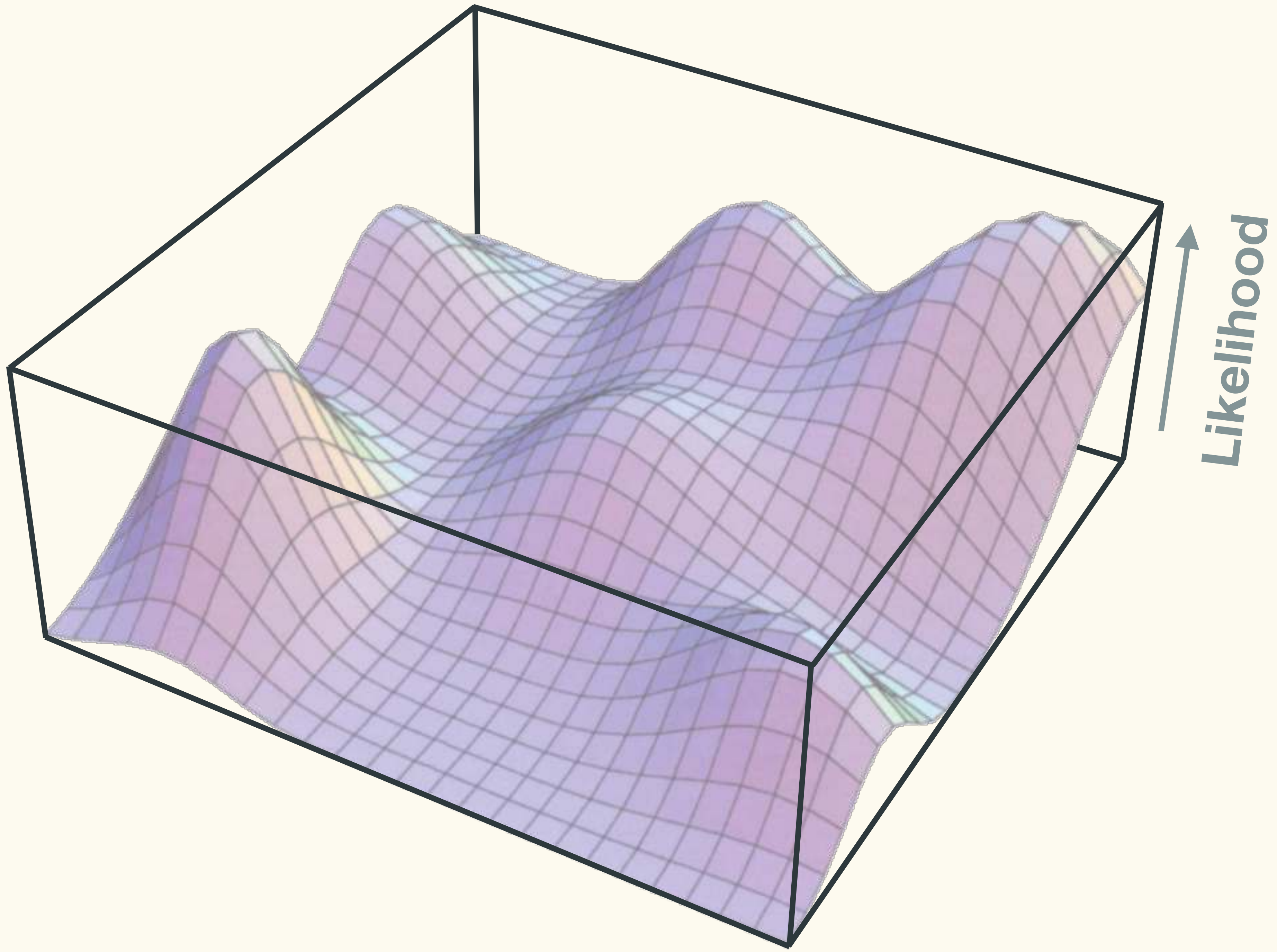
$$\log\left(L\left(\boxed{M_2} \,\middle|\, \begin{matrix}\text{ACTTG}\\ \text{ACTGG}\end{matrix}\right)\right) = -10.3$$

Estimating model parameters using likelihood

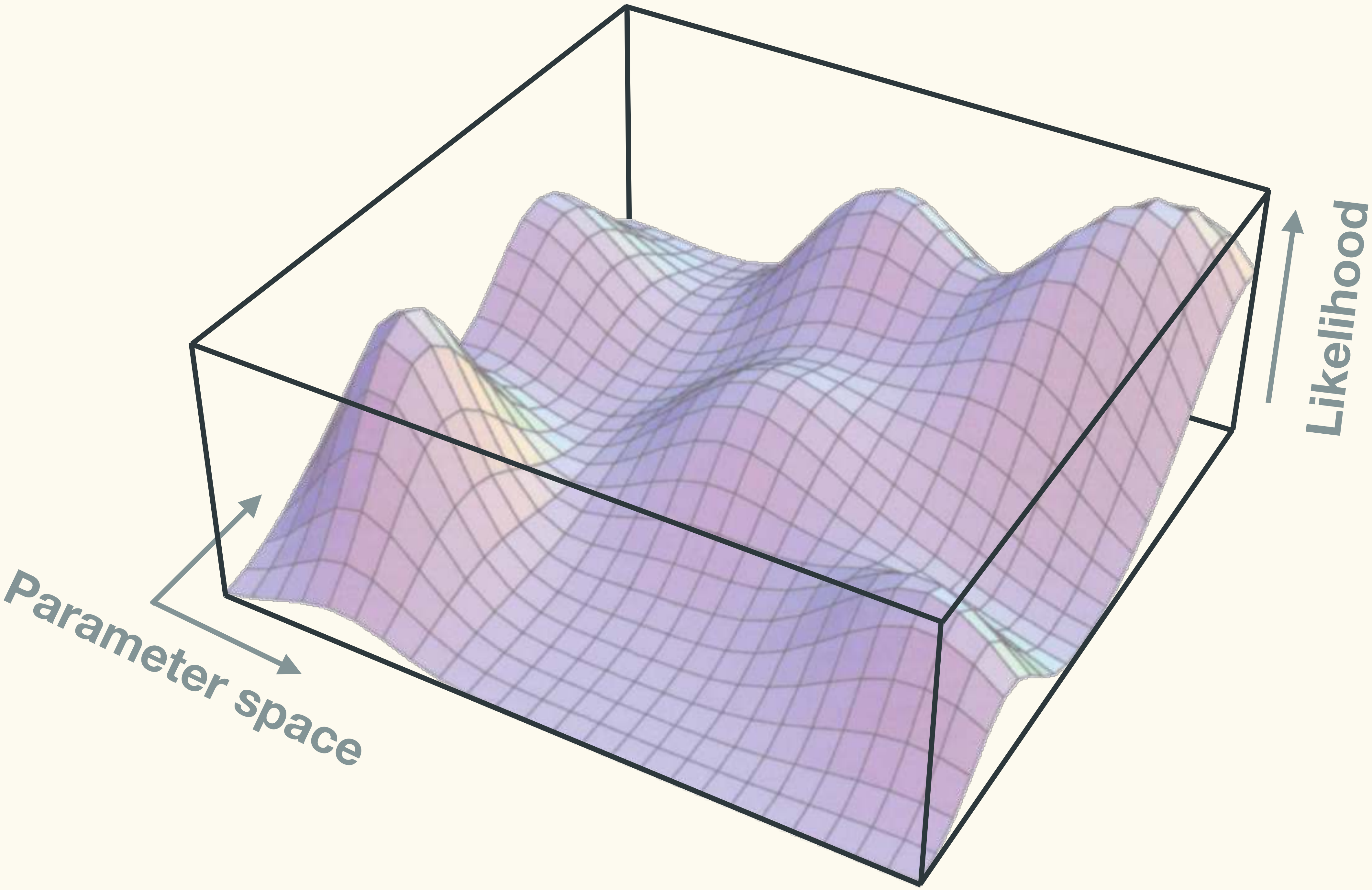
Estimating model parameters using likelihood

Maximum likelihood

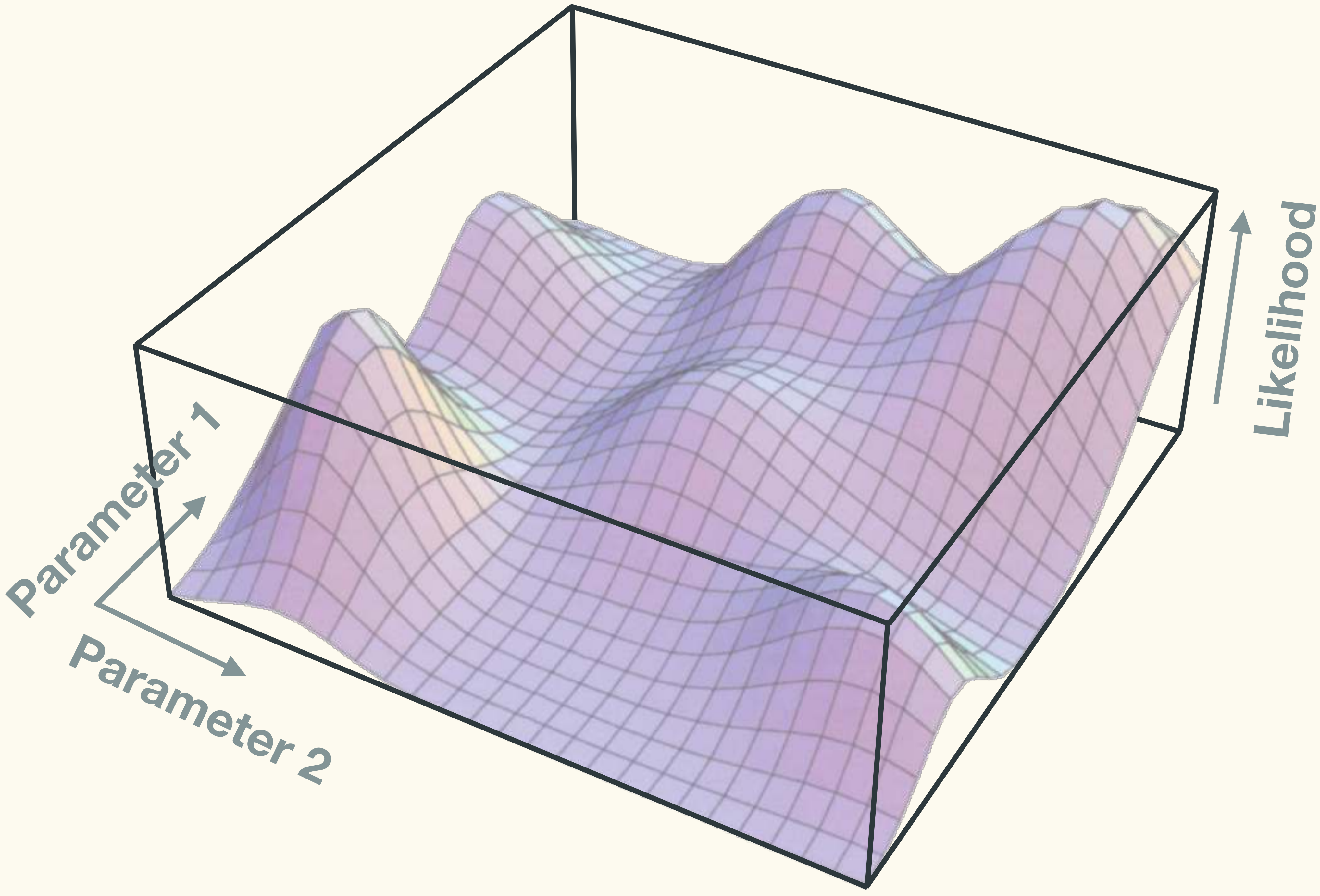
Likelihood surface



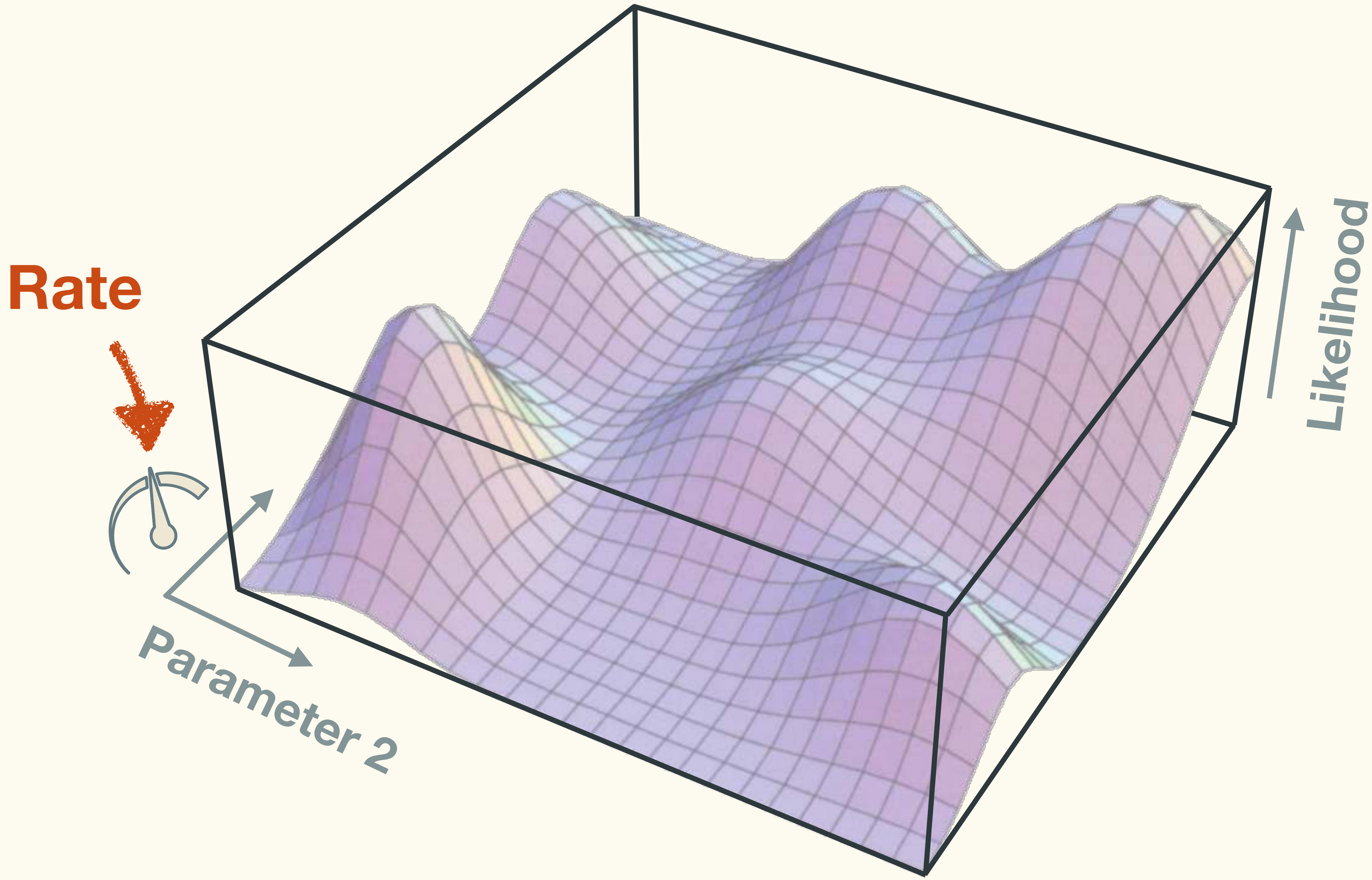
Likelihood surface



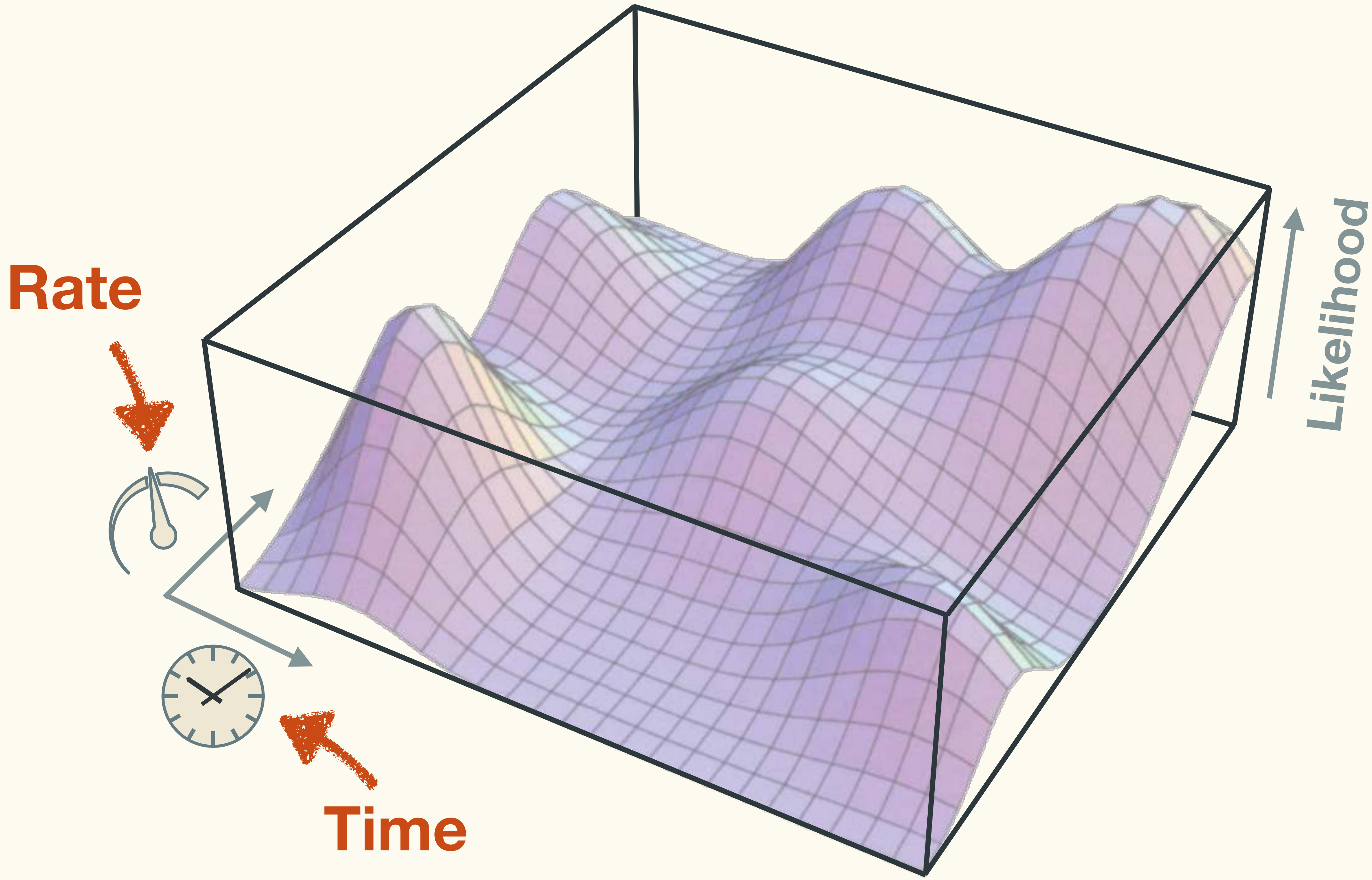
Likelihood surface



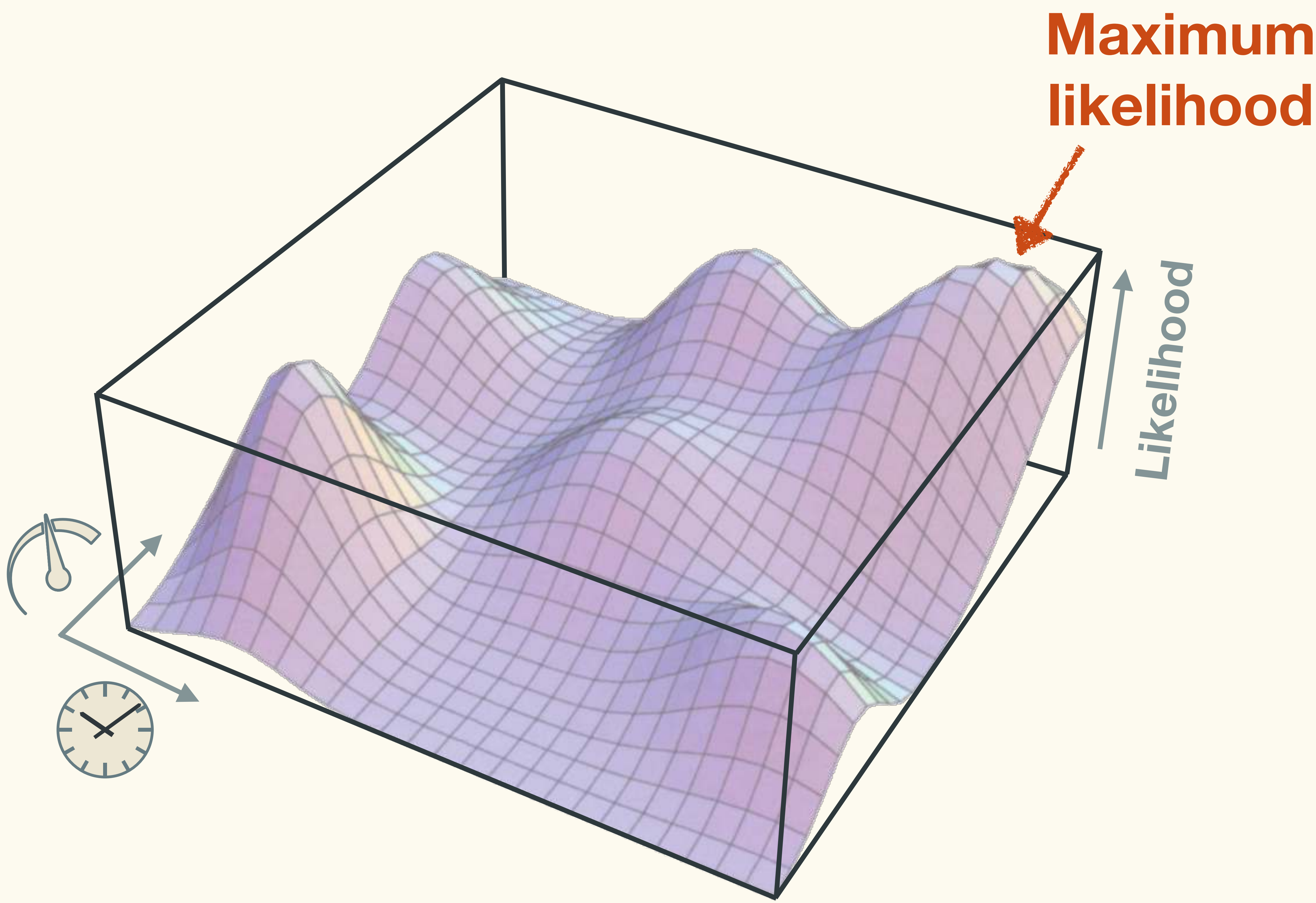
Likelihood surface



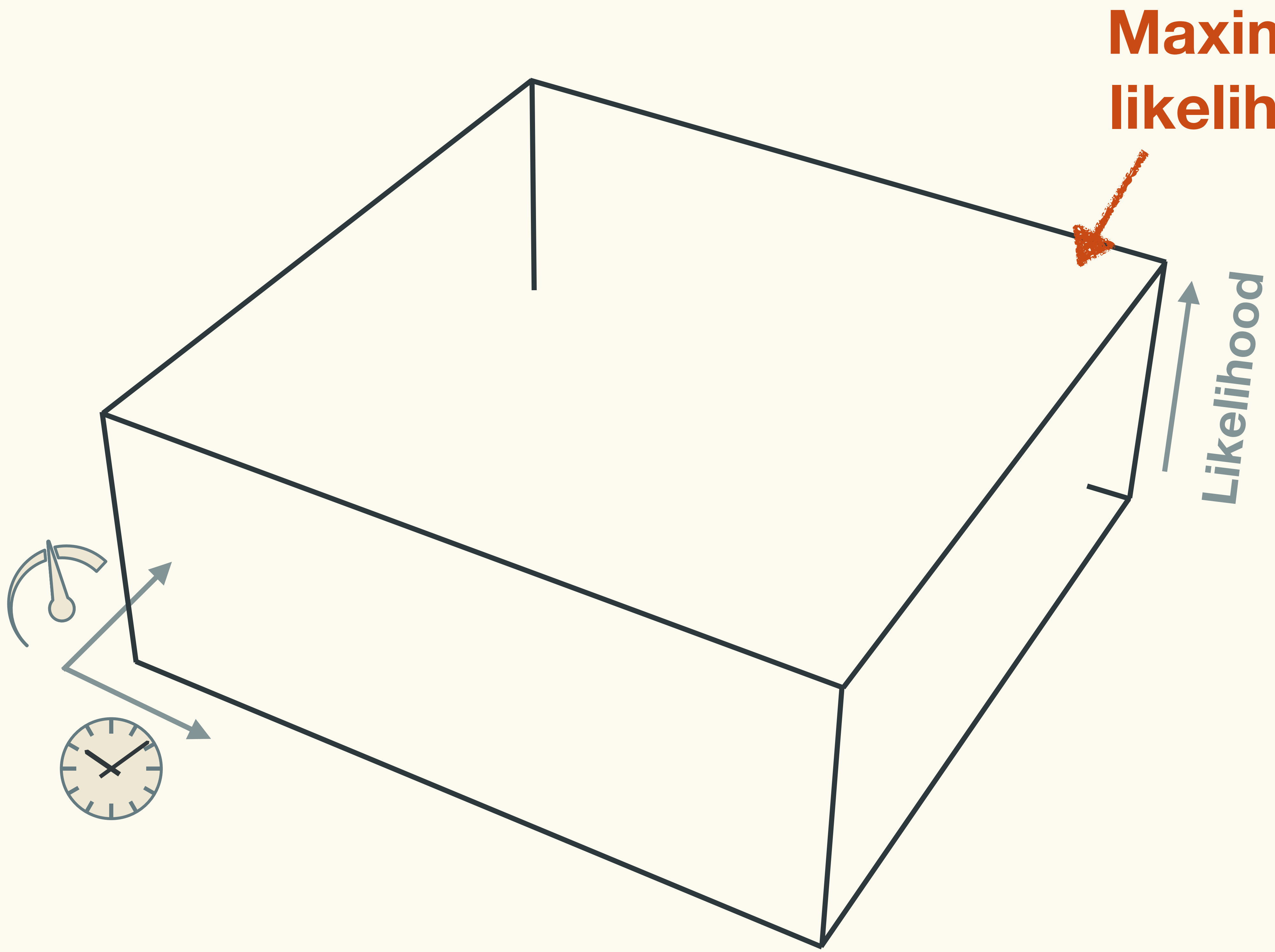
Likelihood surface



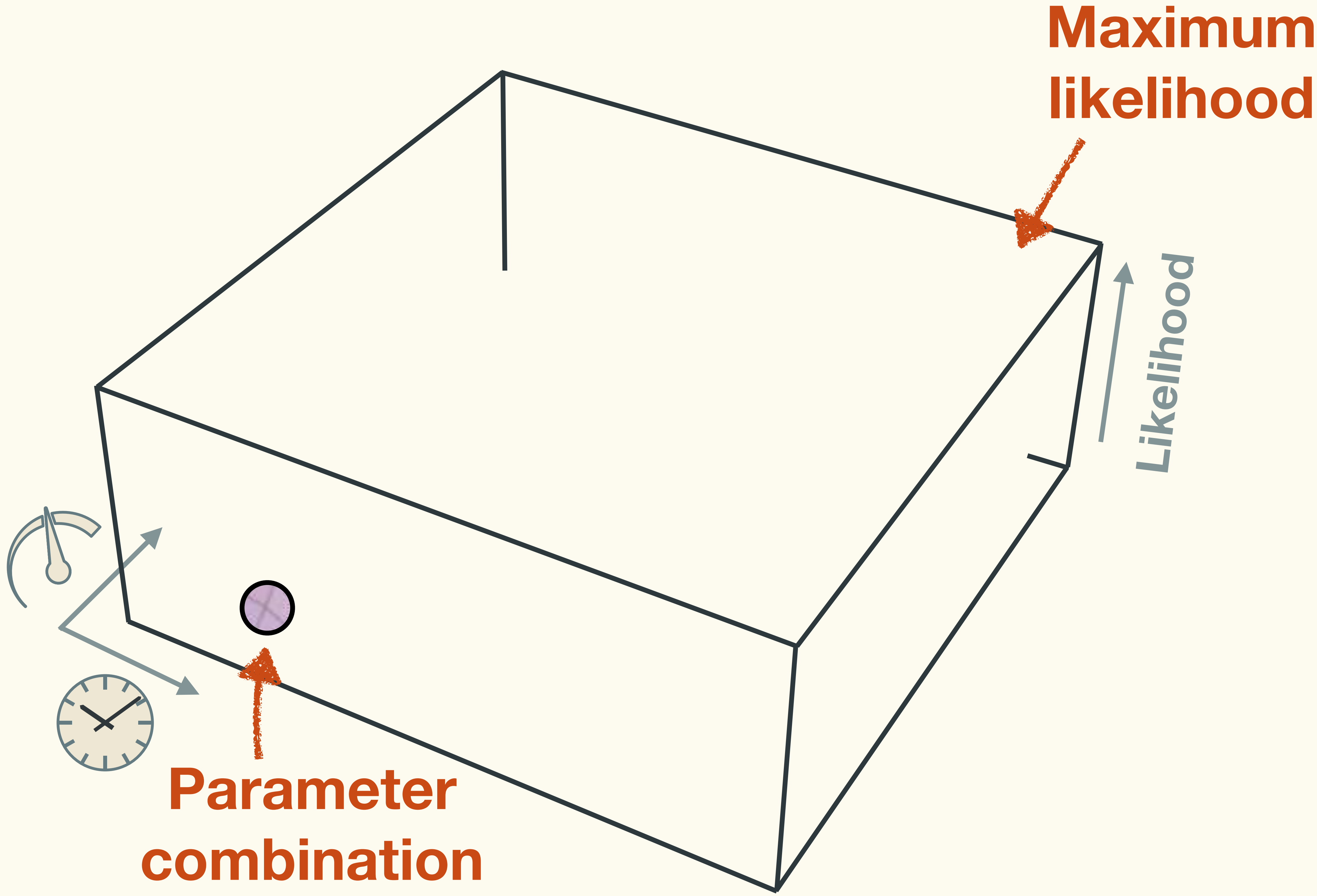
Likelihood surface



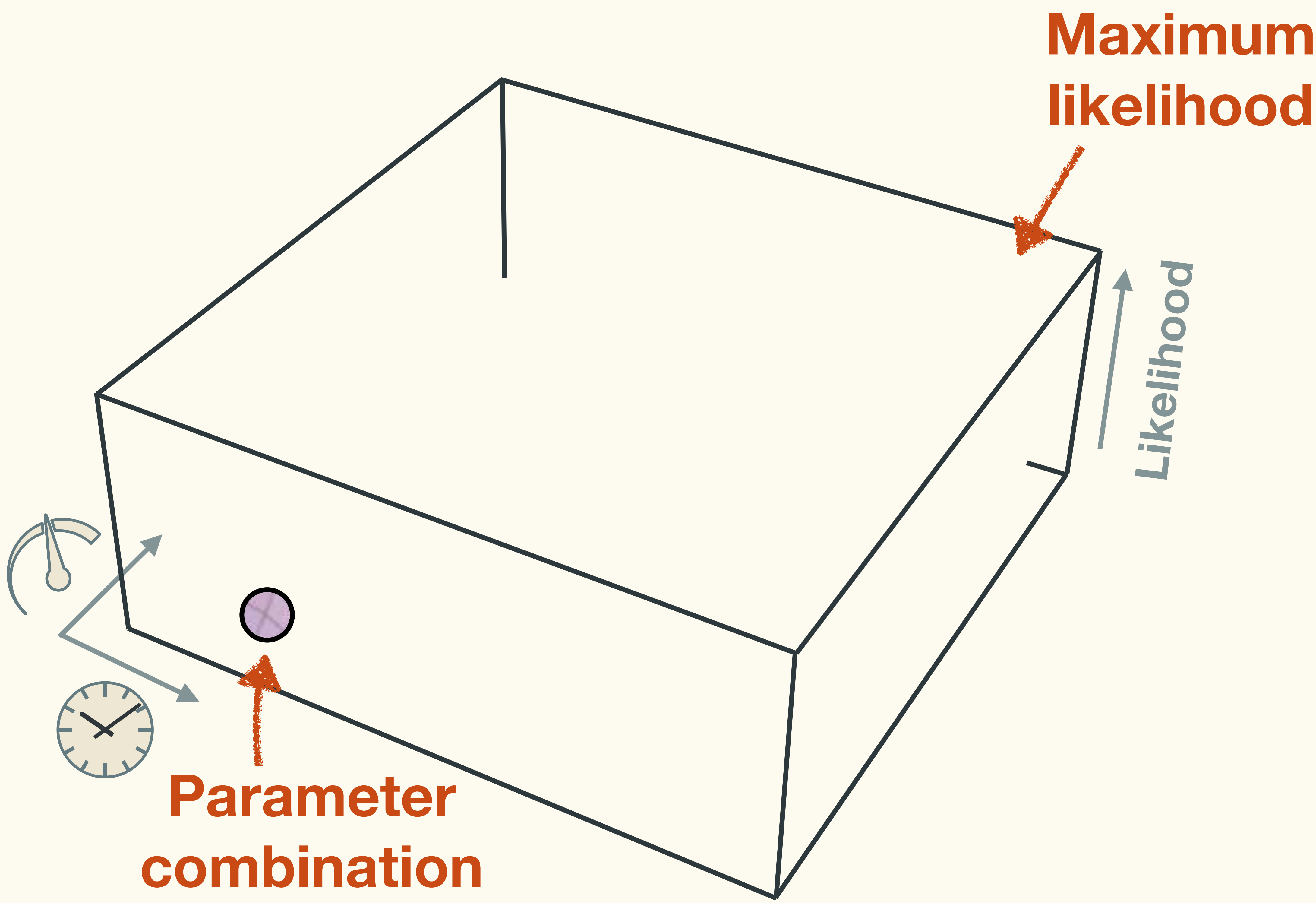
Likelihood surface



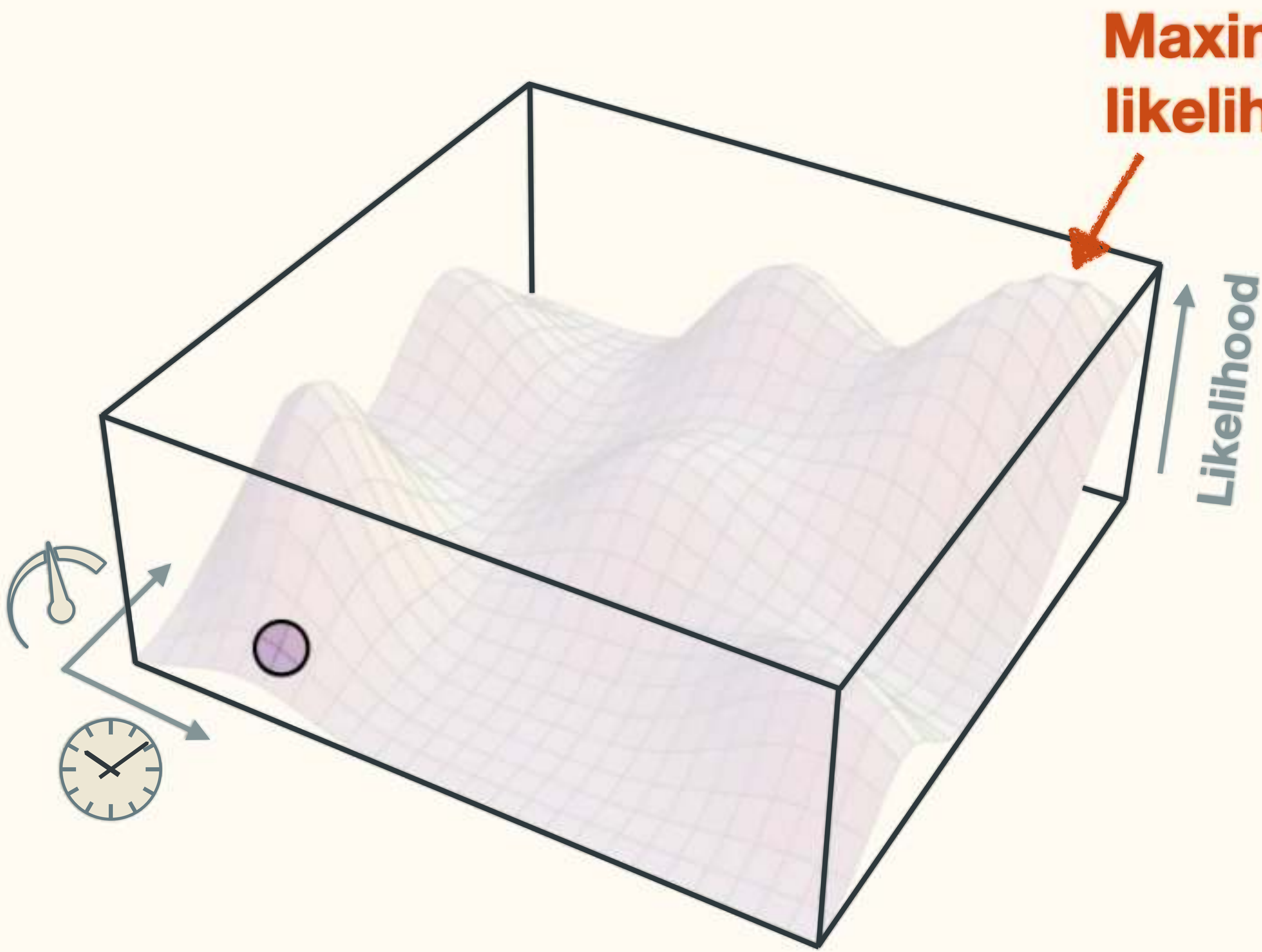
Likelihood surface



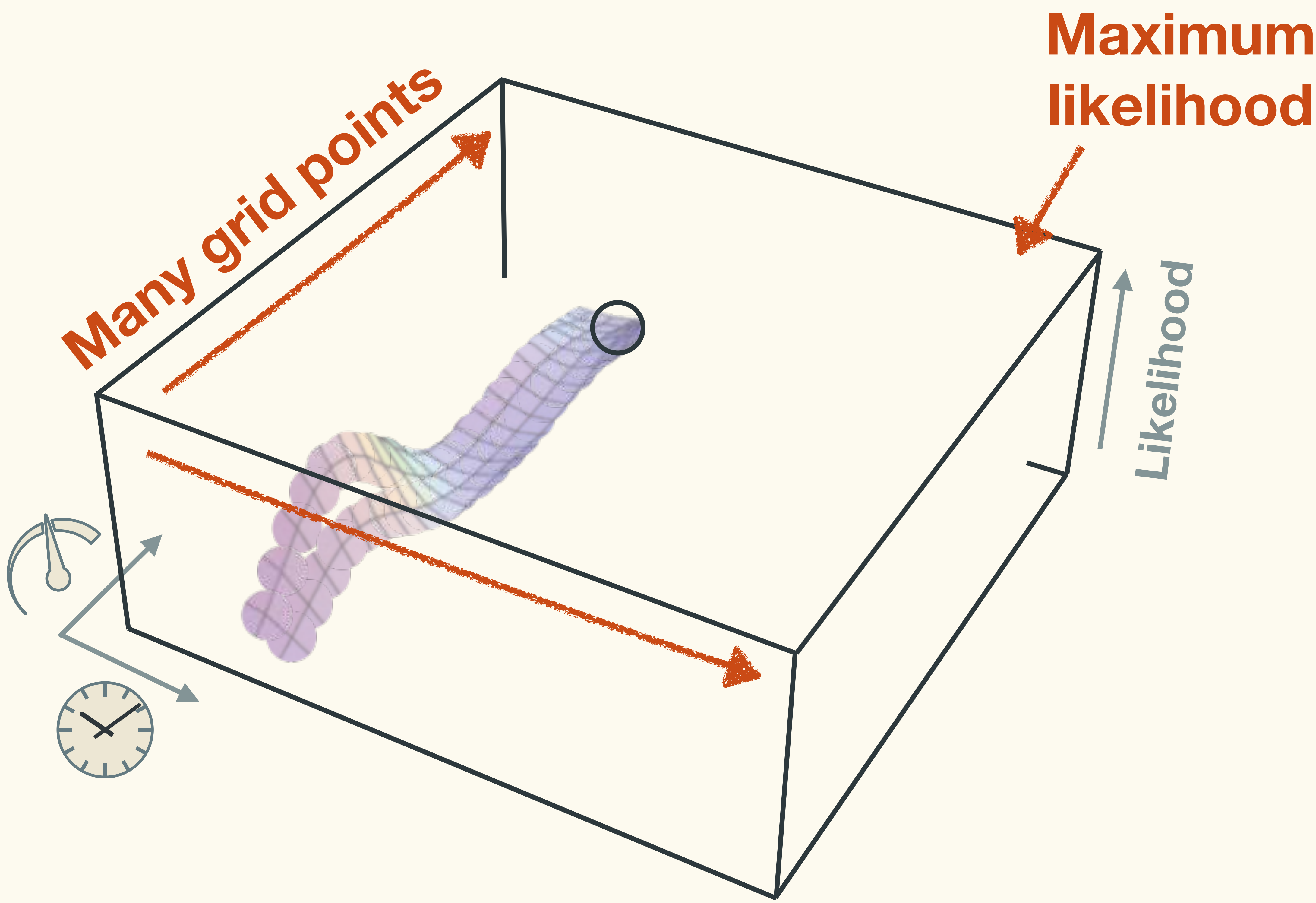
Likelihood surface



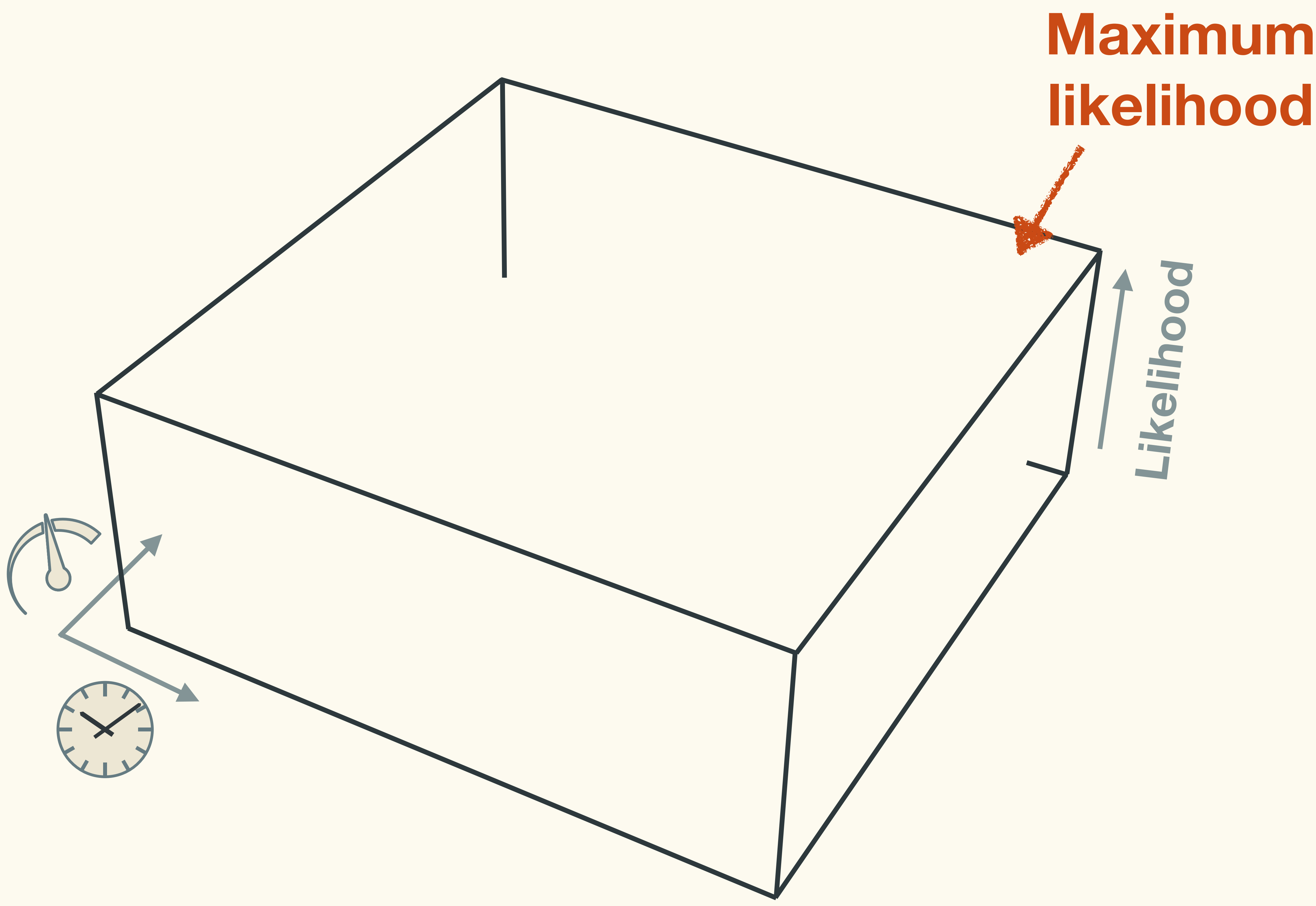
Grid search



Grid search

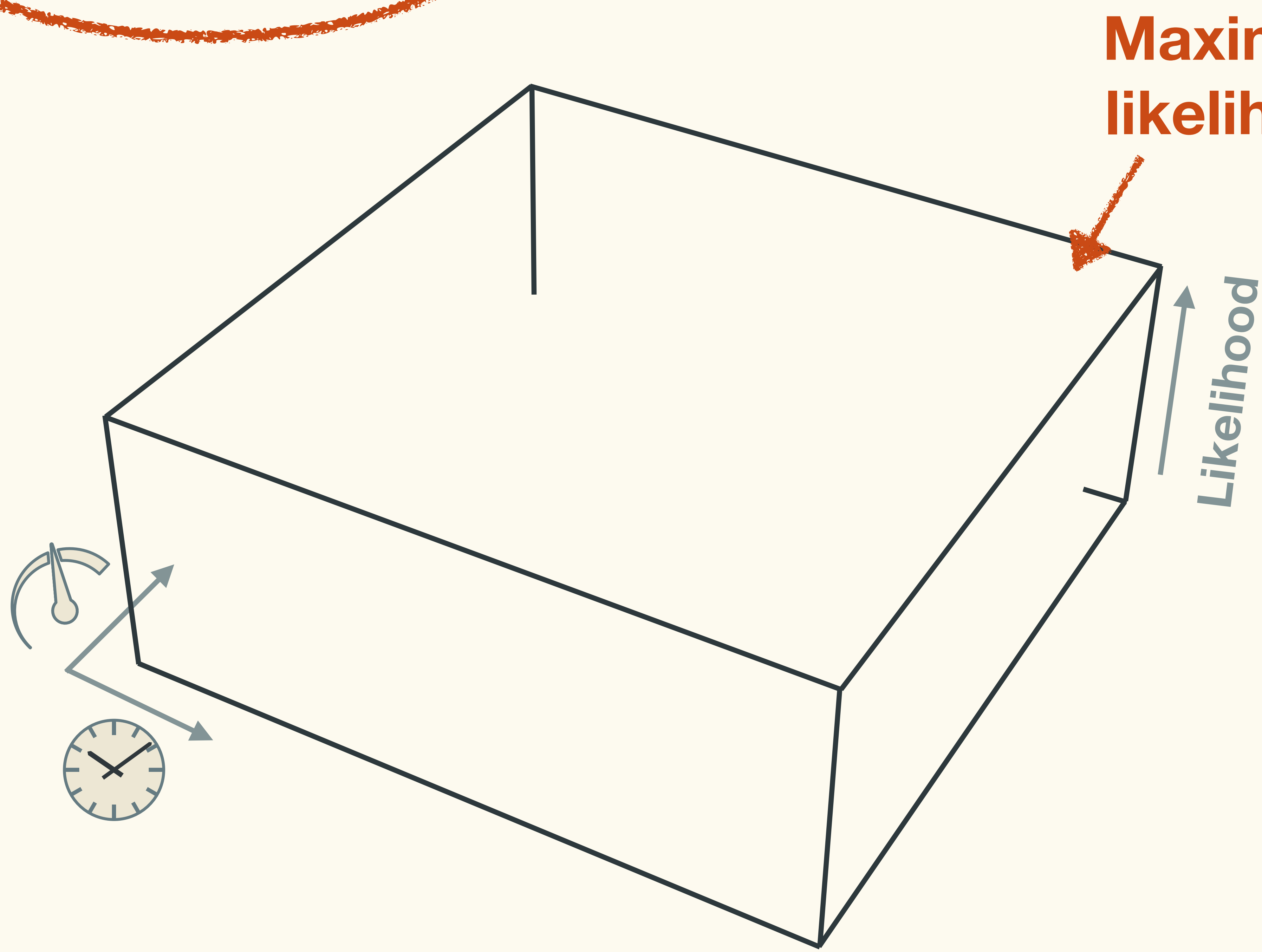


Heuristic search



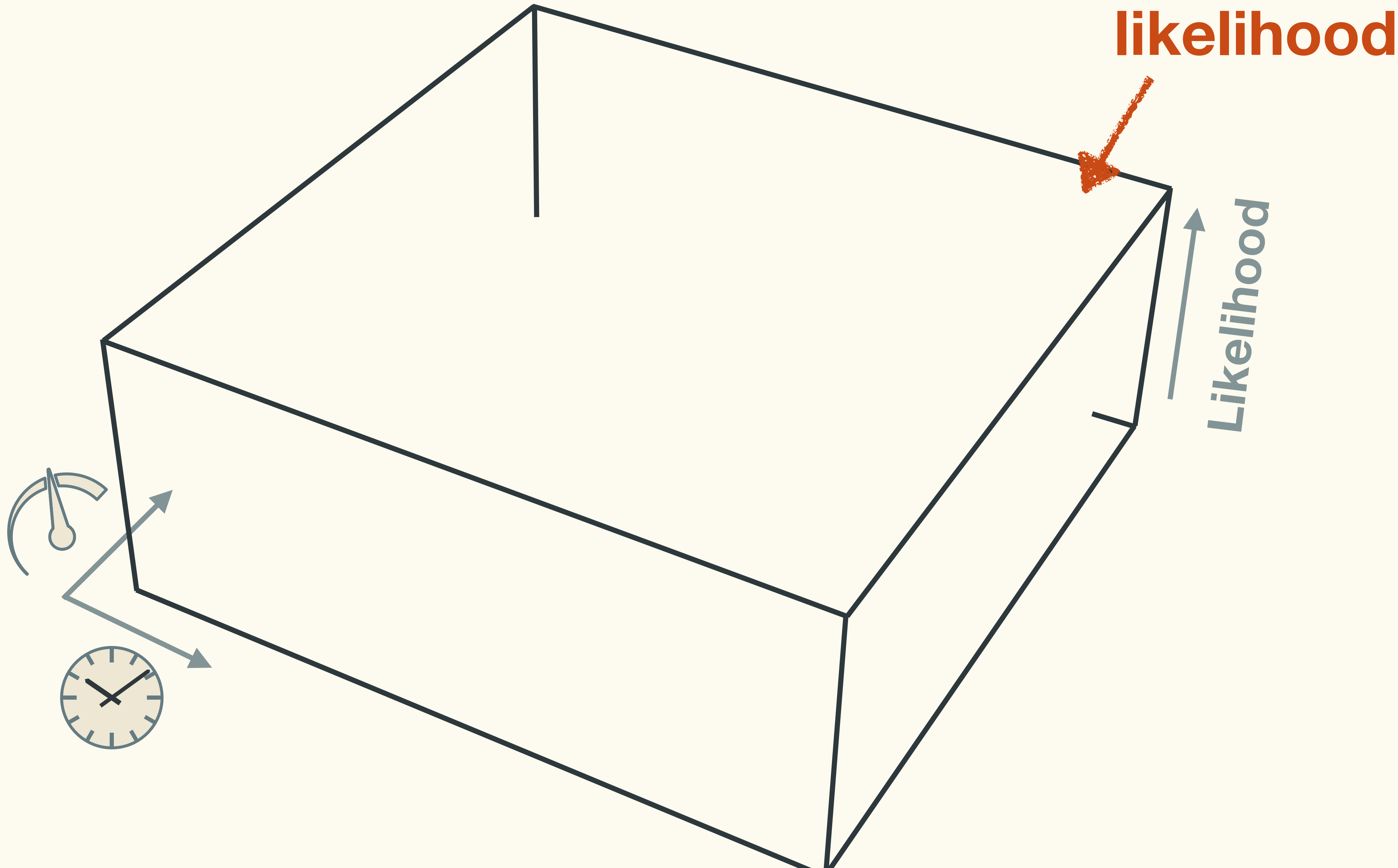
“proceeding to a solution by trial and error”

Heuristic search



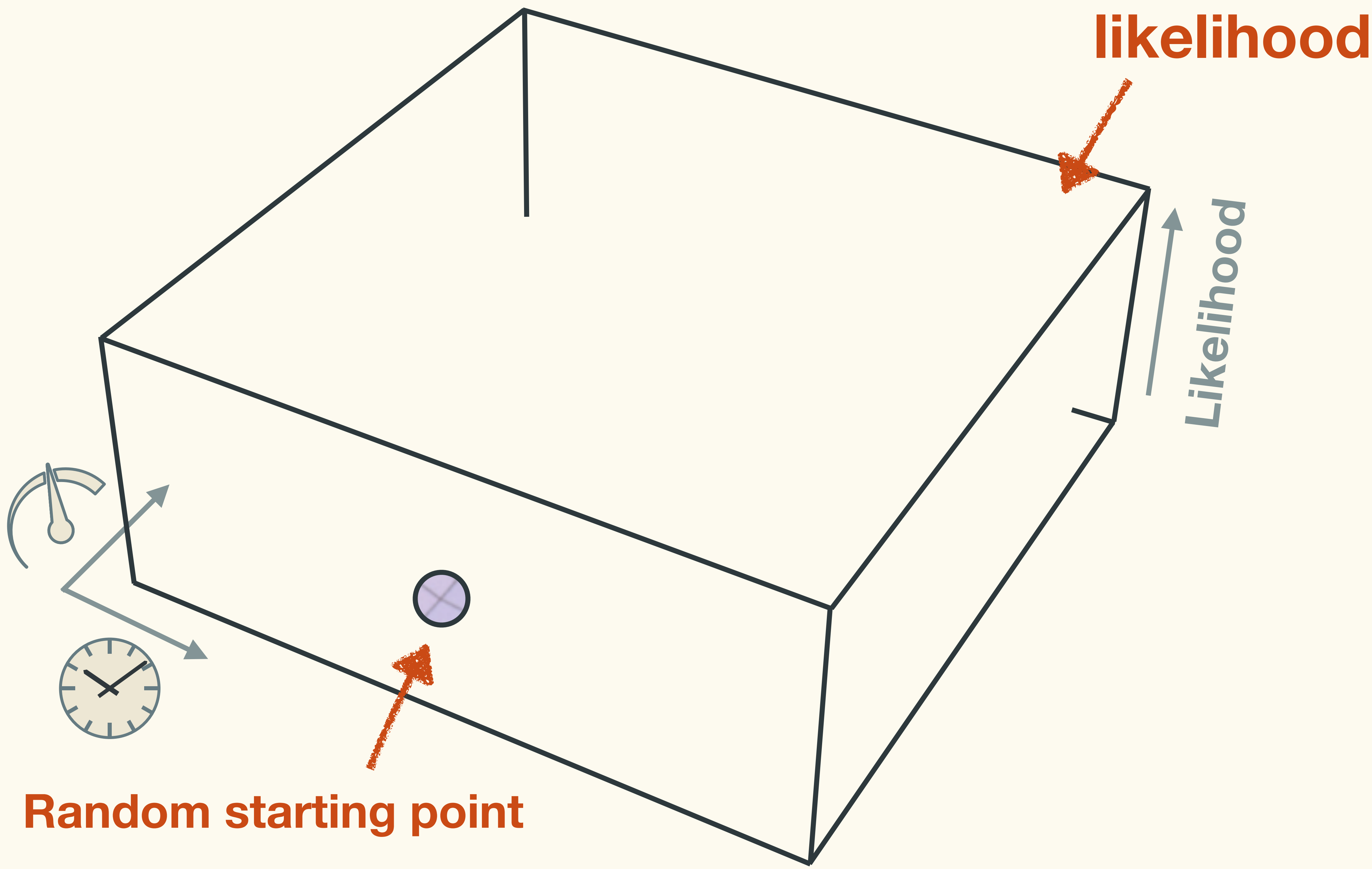
Heuristic search

Hill climbing



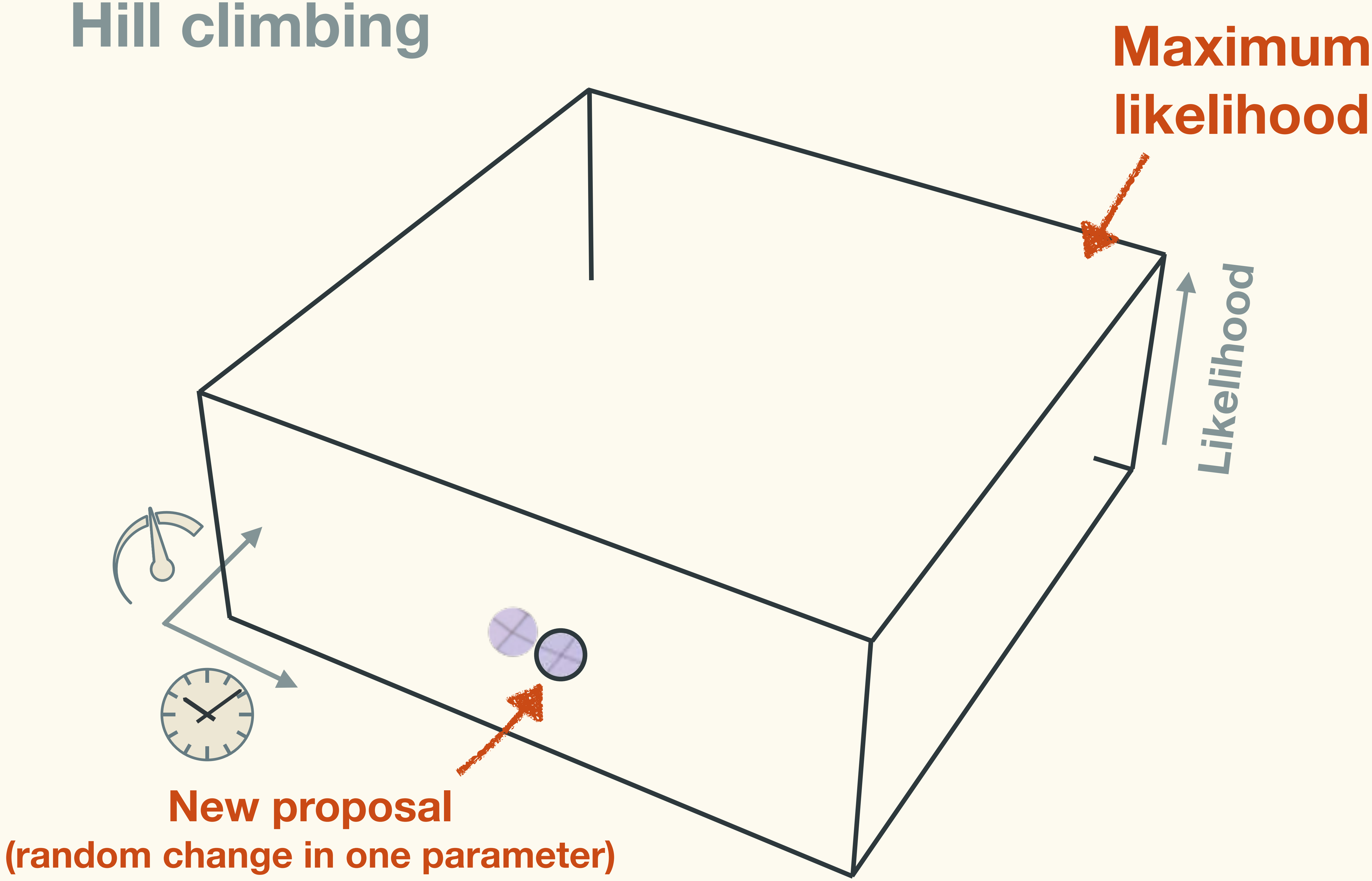
Heuristic search

Hill climbing



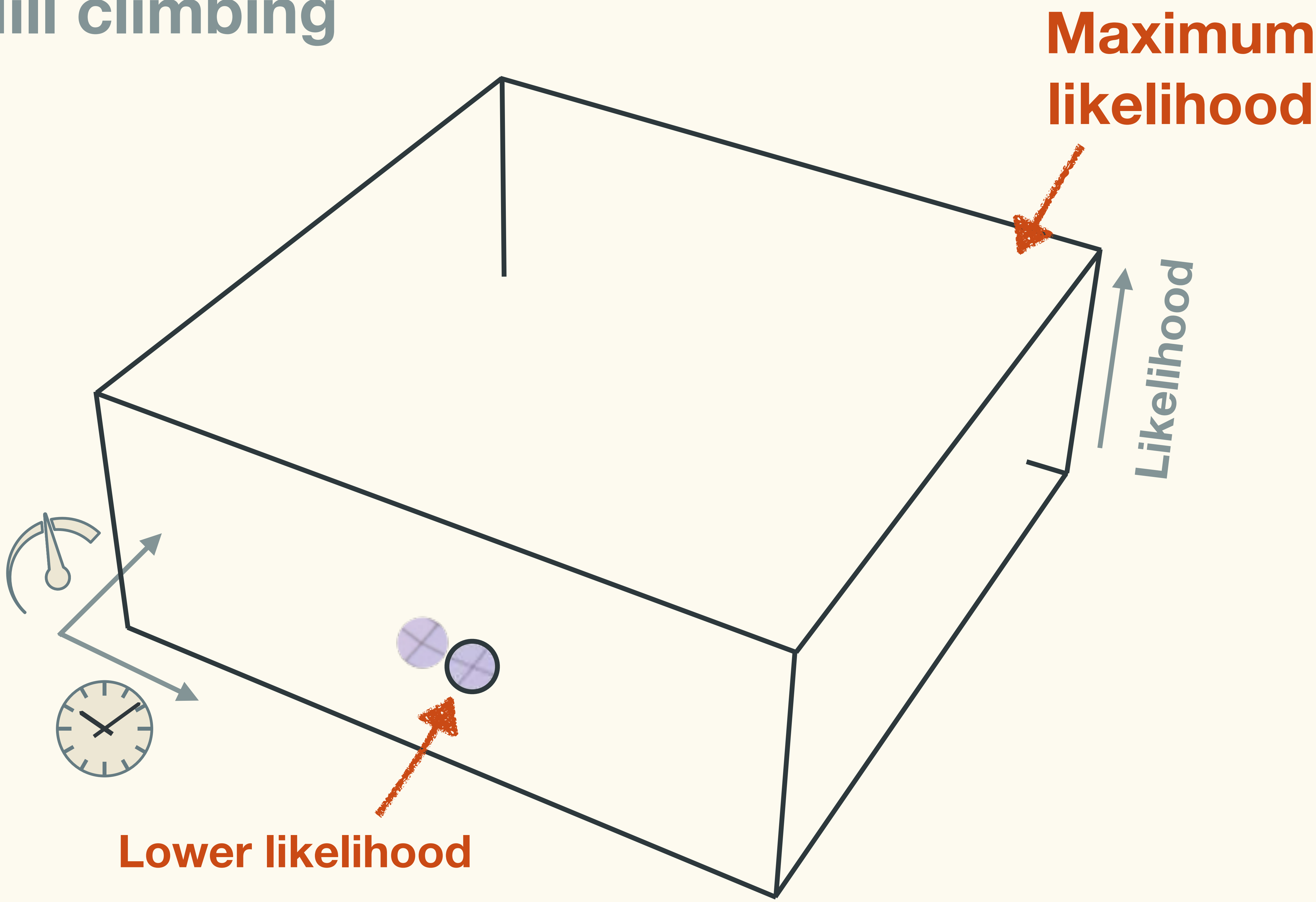
Heuristic search

Hill climbing



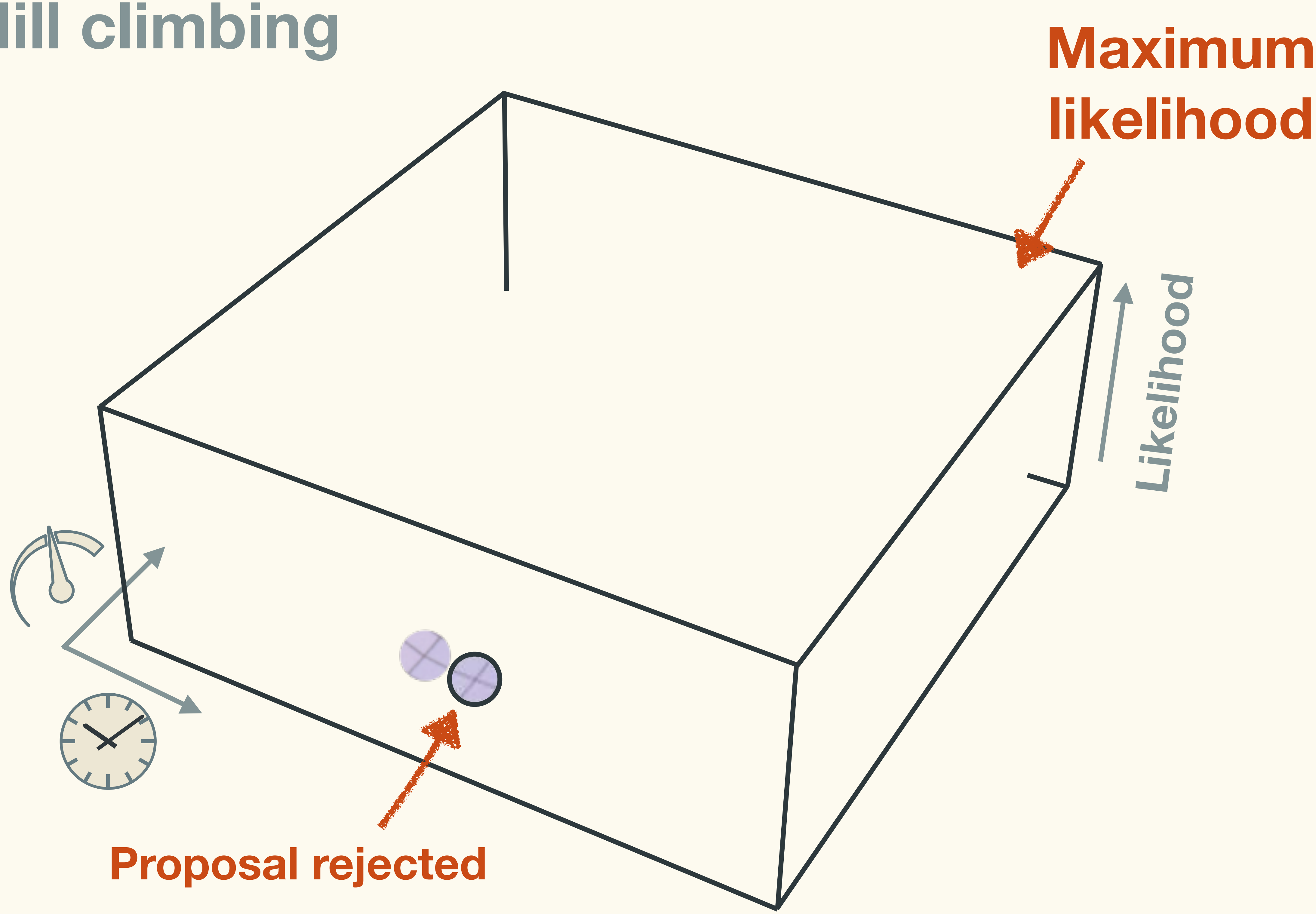
Heuristic search

Hill climbing



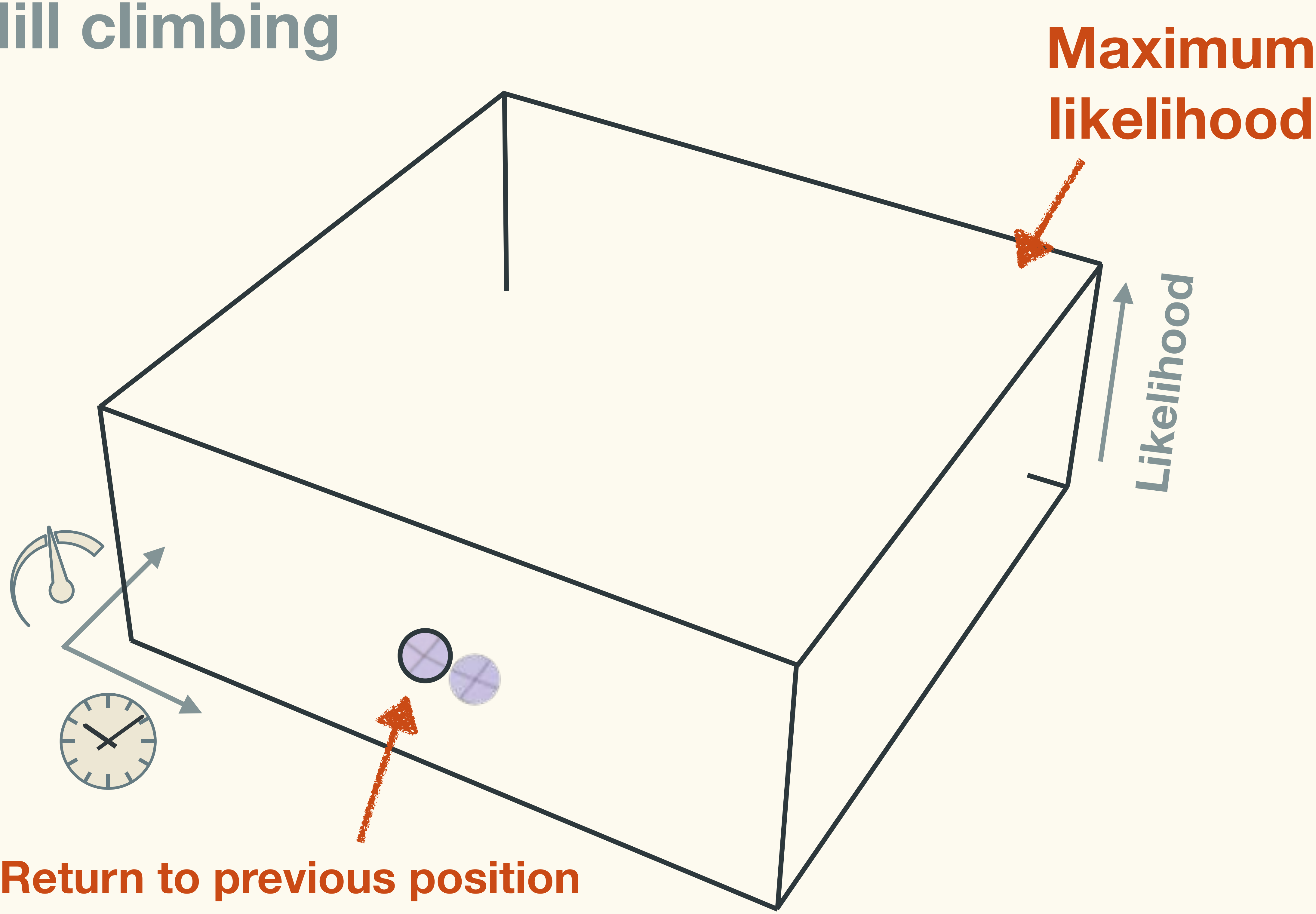
Heuristic search

Hill climbing



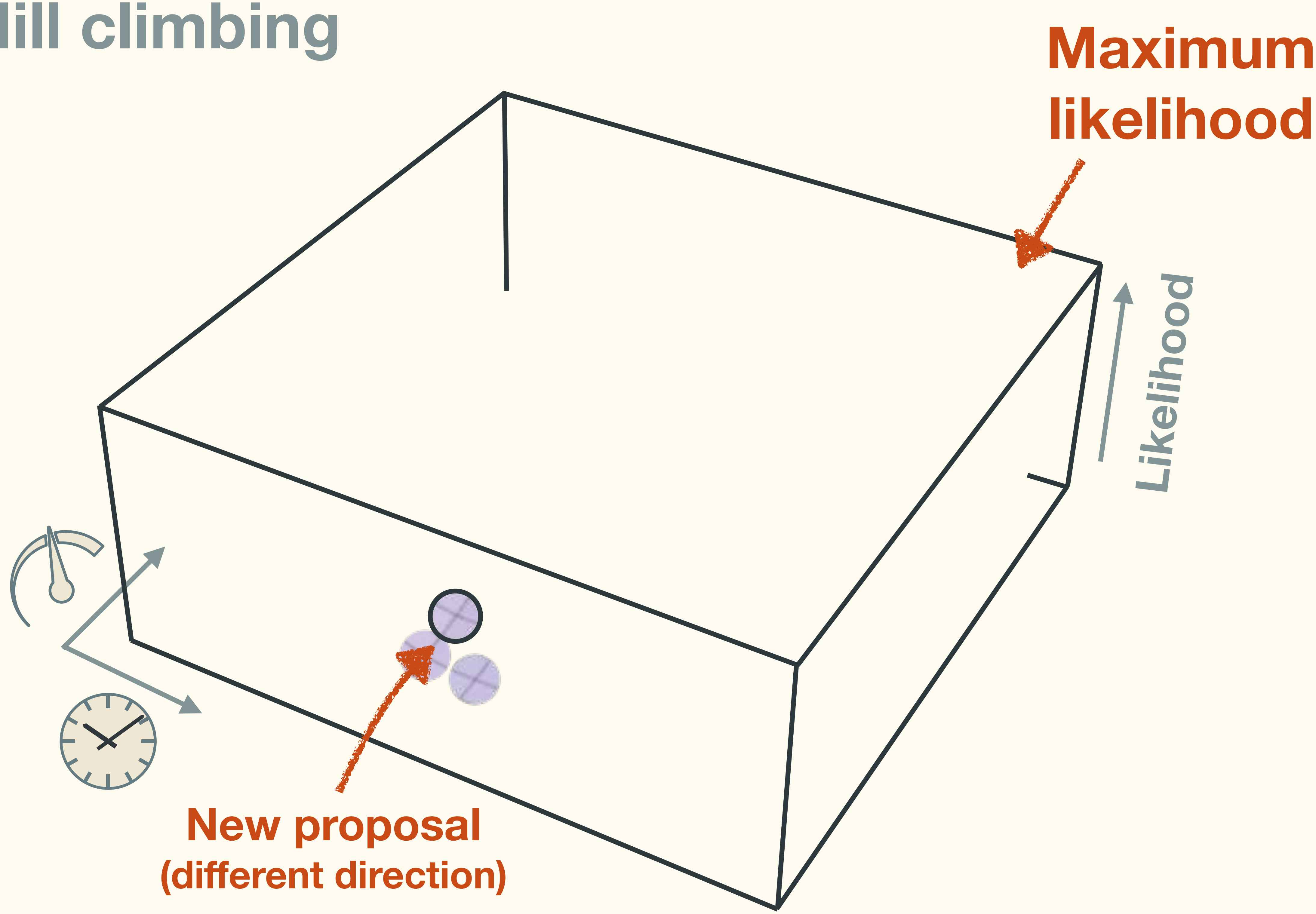
Heuristic search

Hill climbing



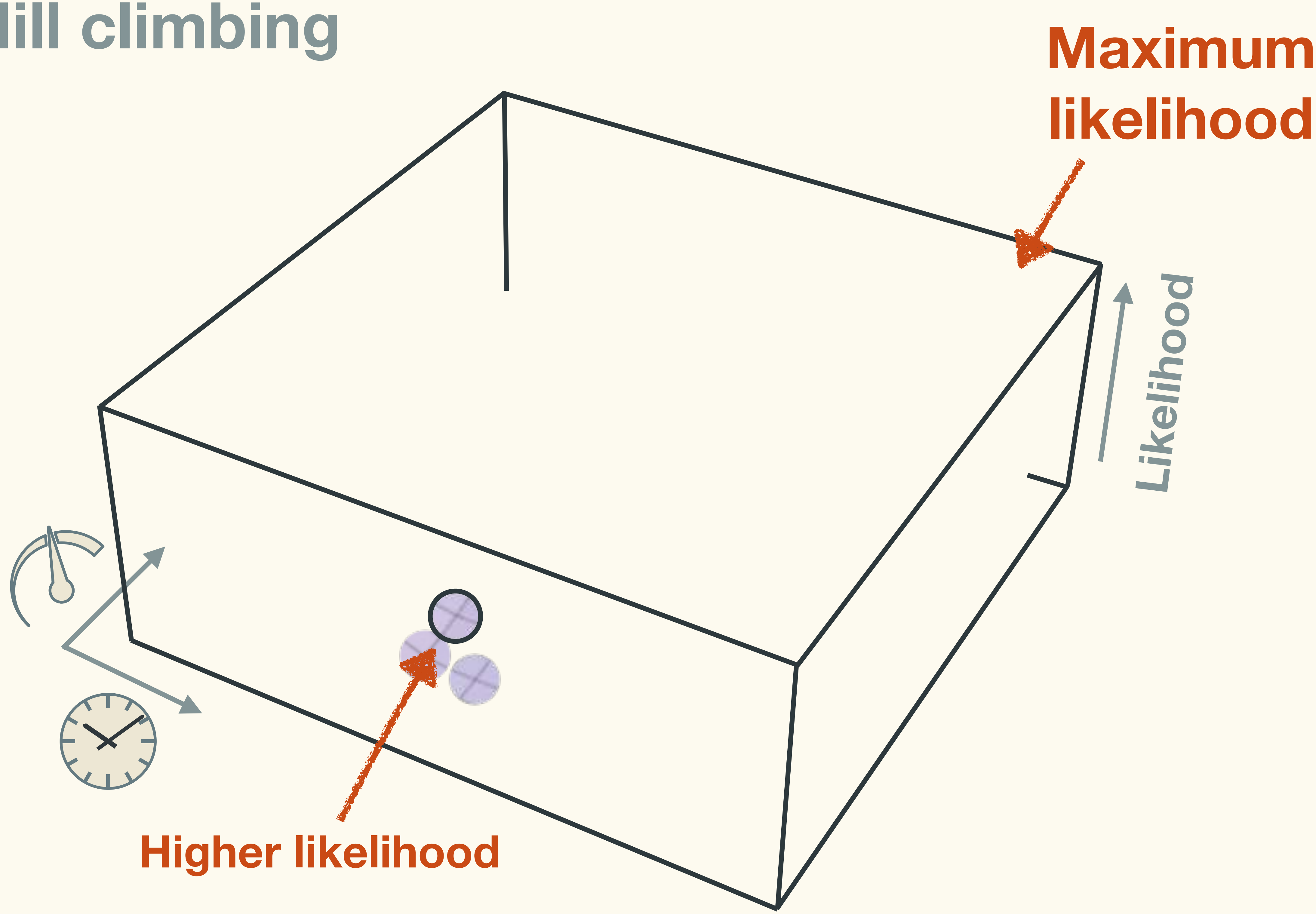
Heuristic search

Hill climbing



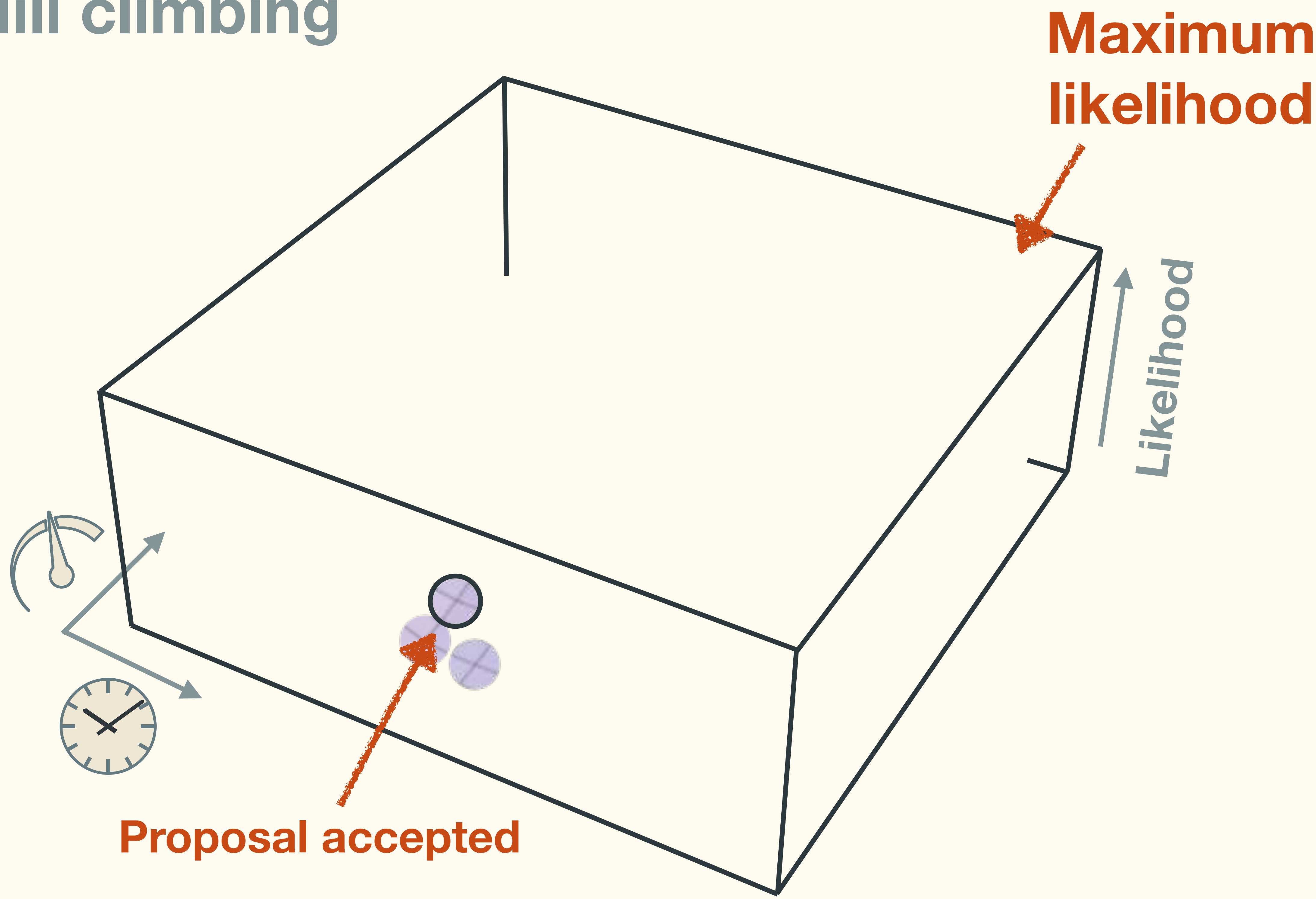
Heuristic search

Hill climbing



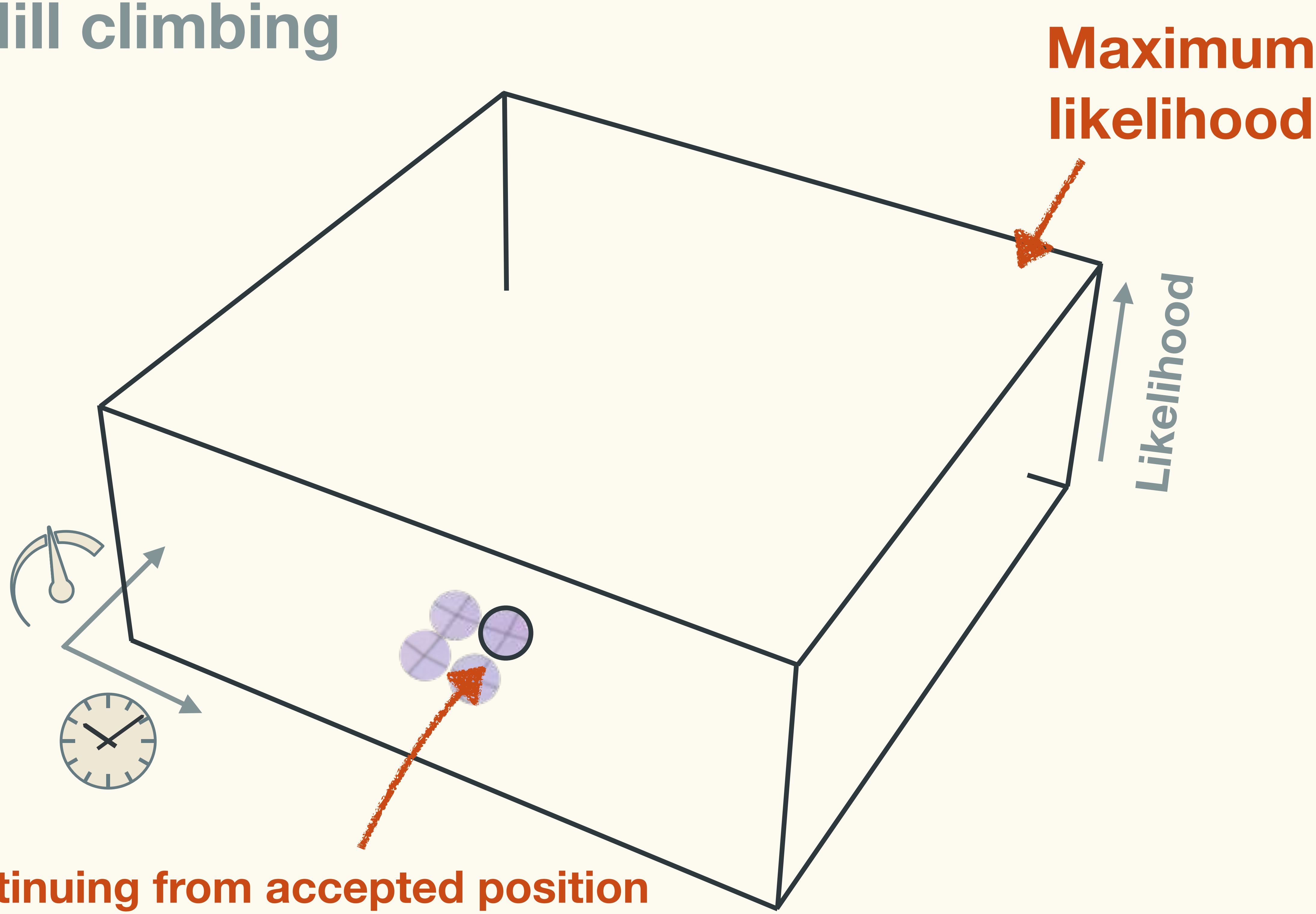
Heuristic search

Hill climbing



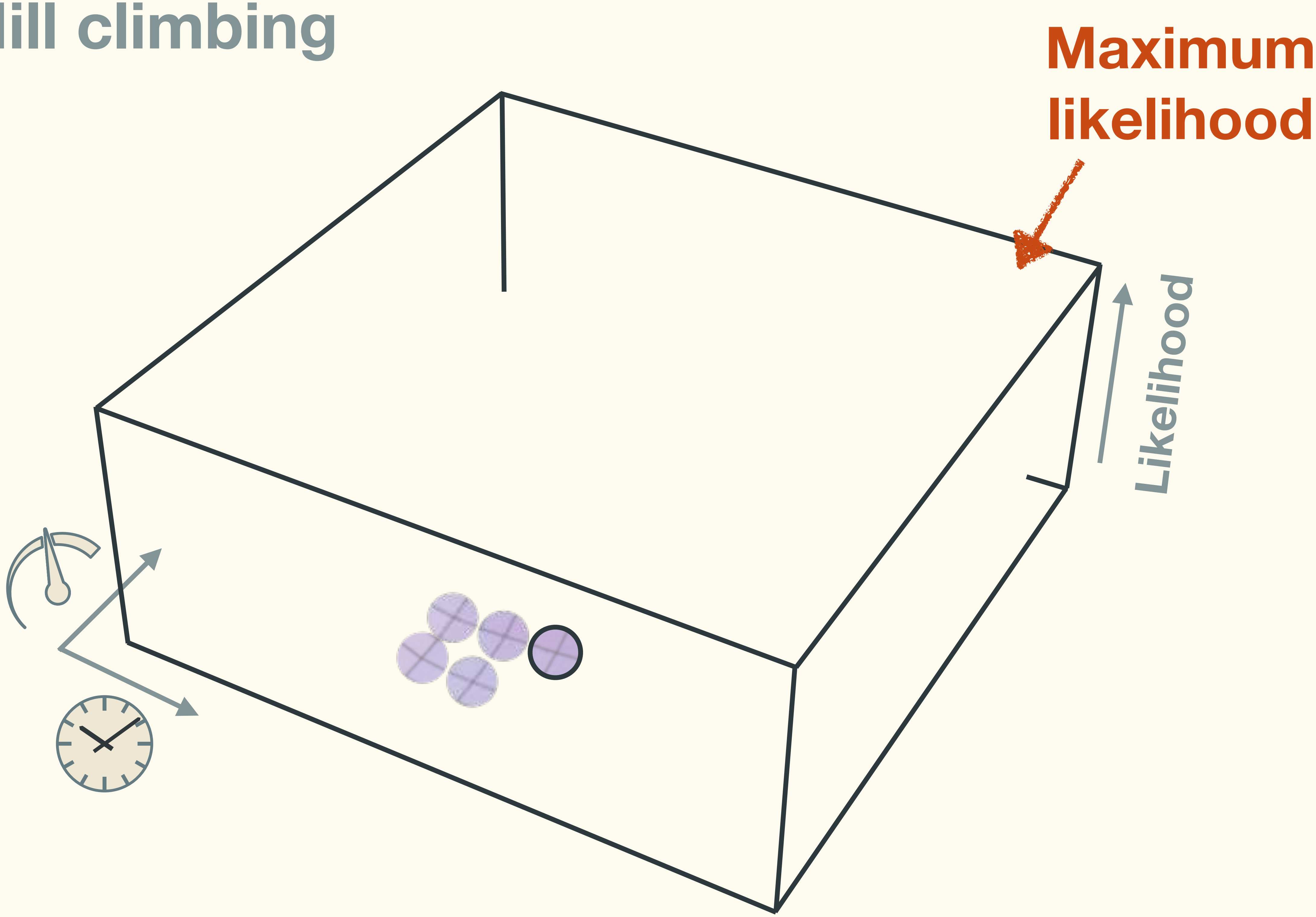
Heuristic search

Hill climbing



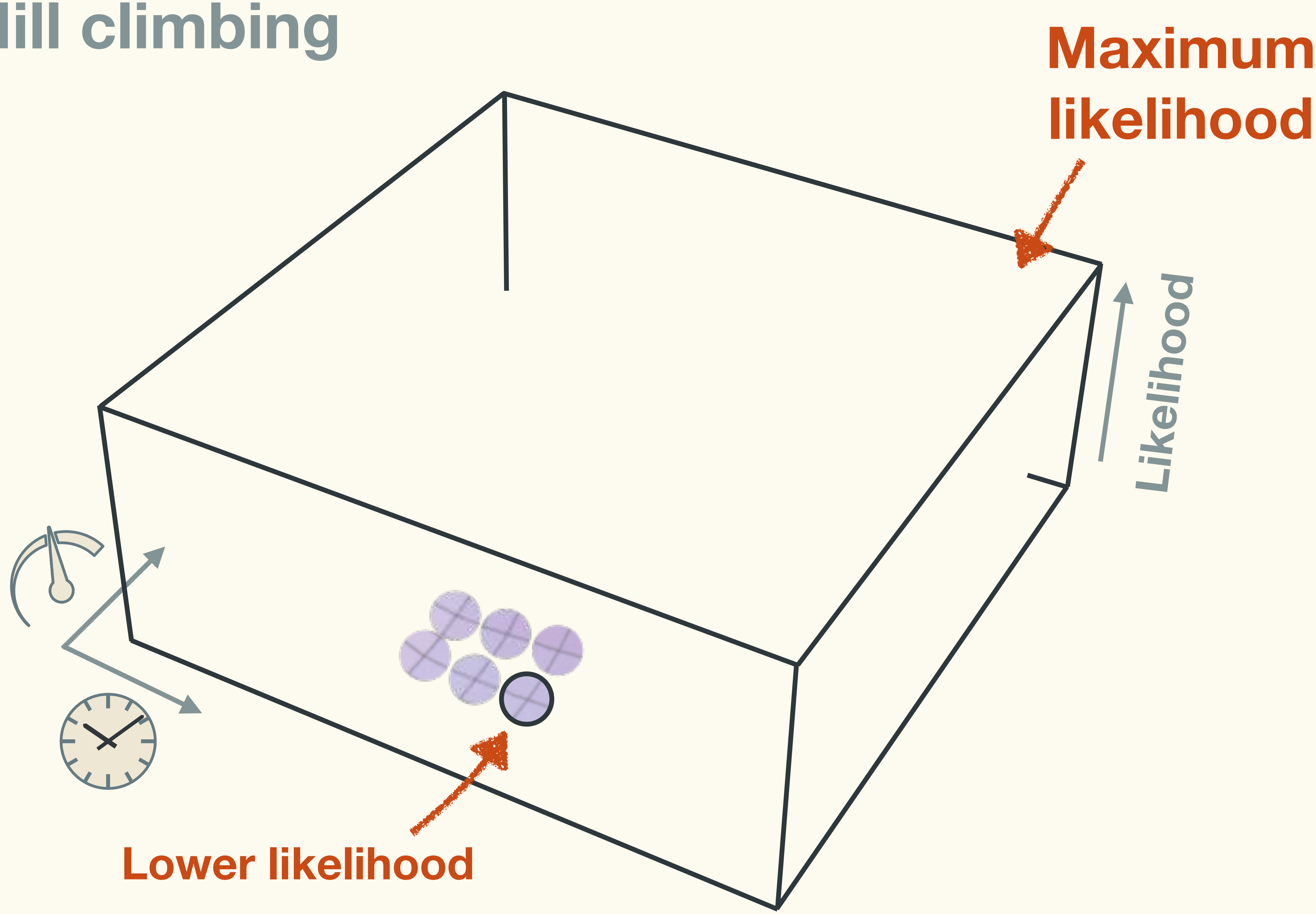
Heuristic search

Hill climbing



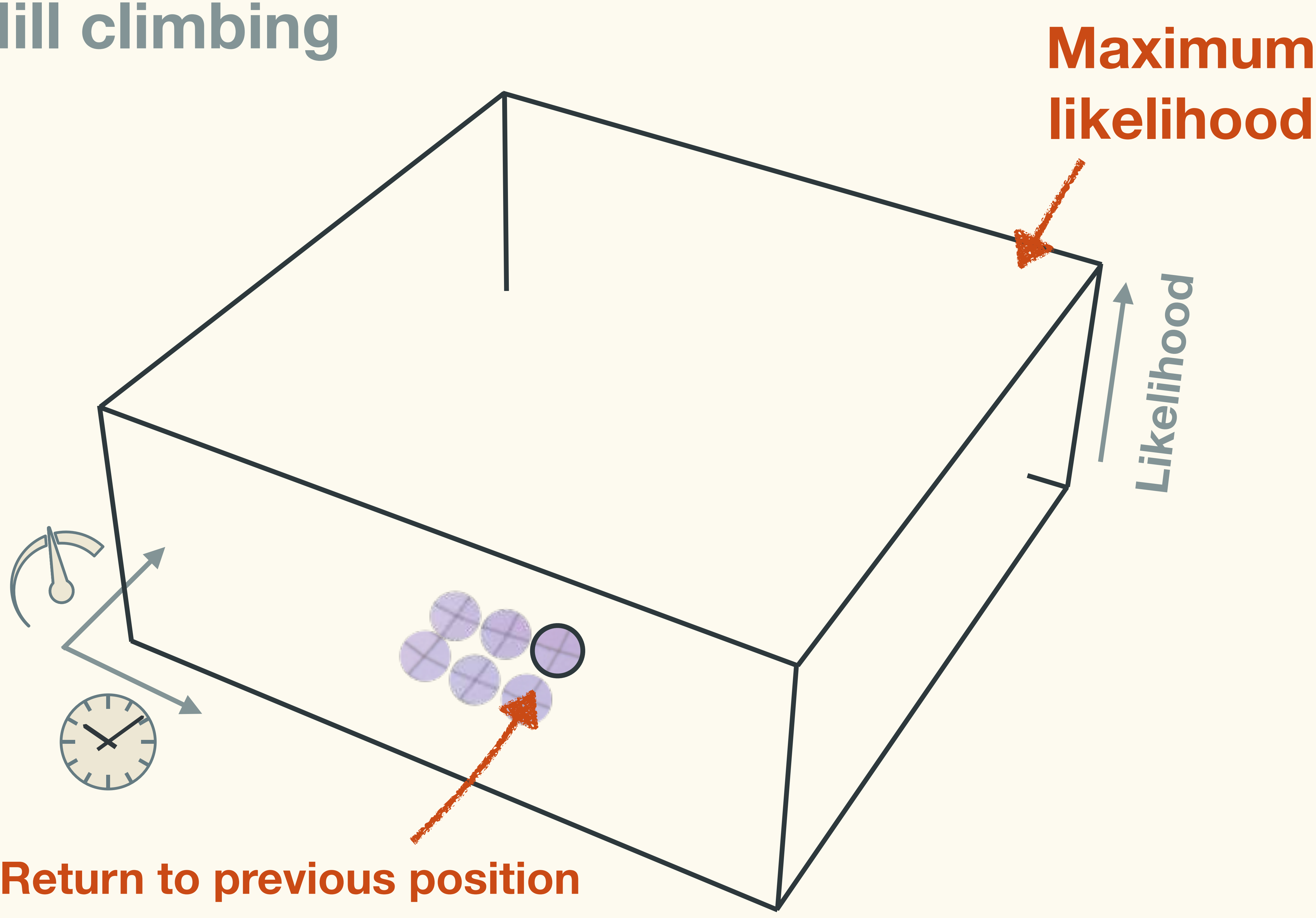
Heuristic search

Hill climbing



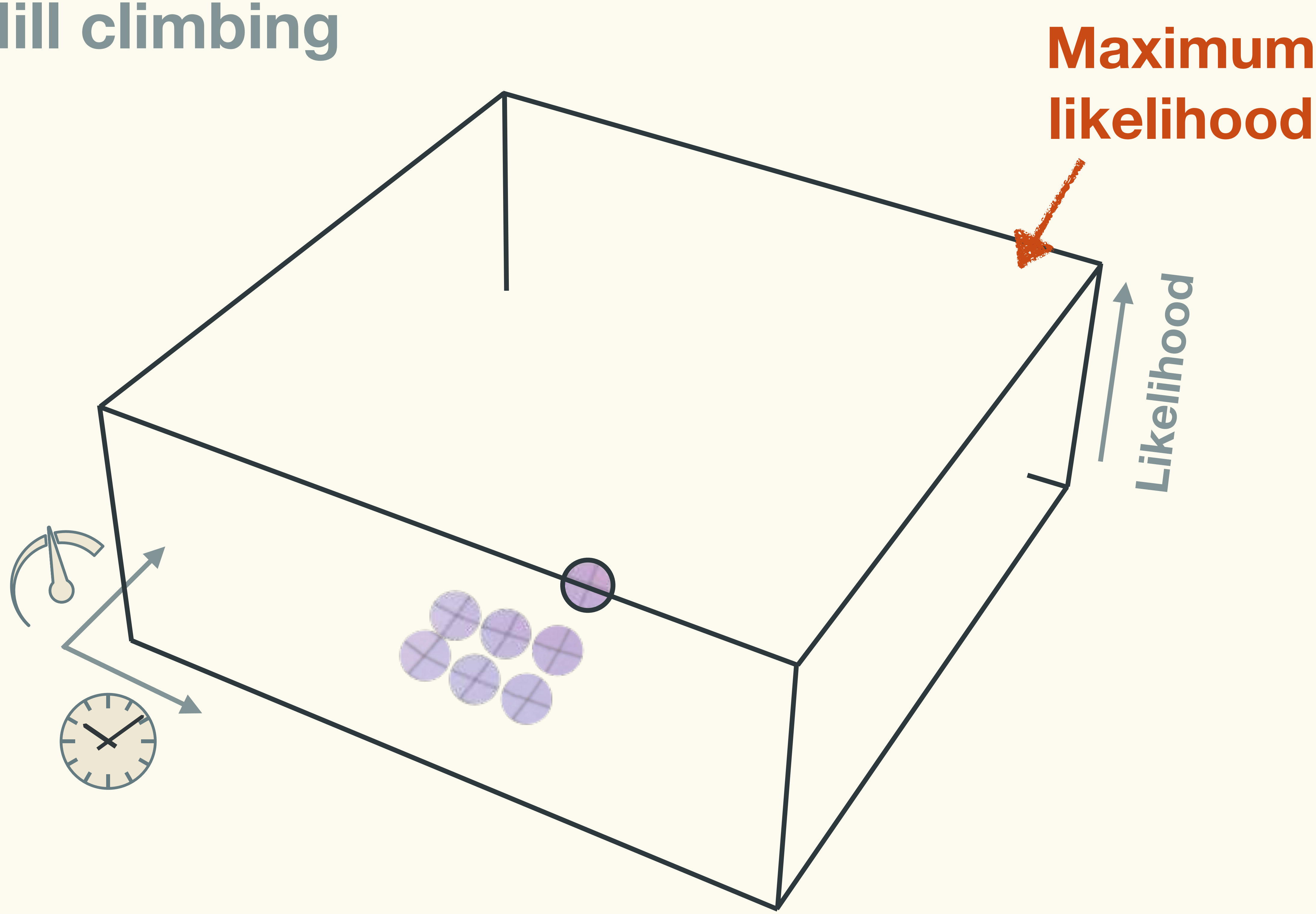
Heuristic search

Hill climbing



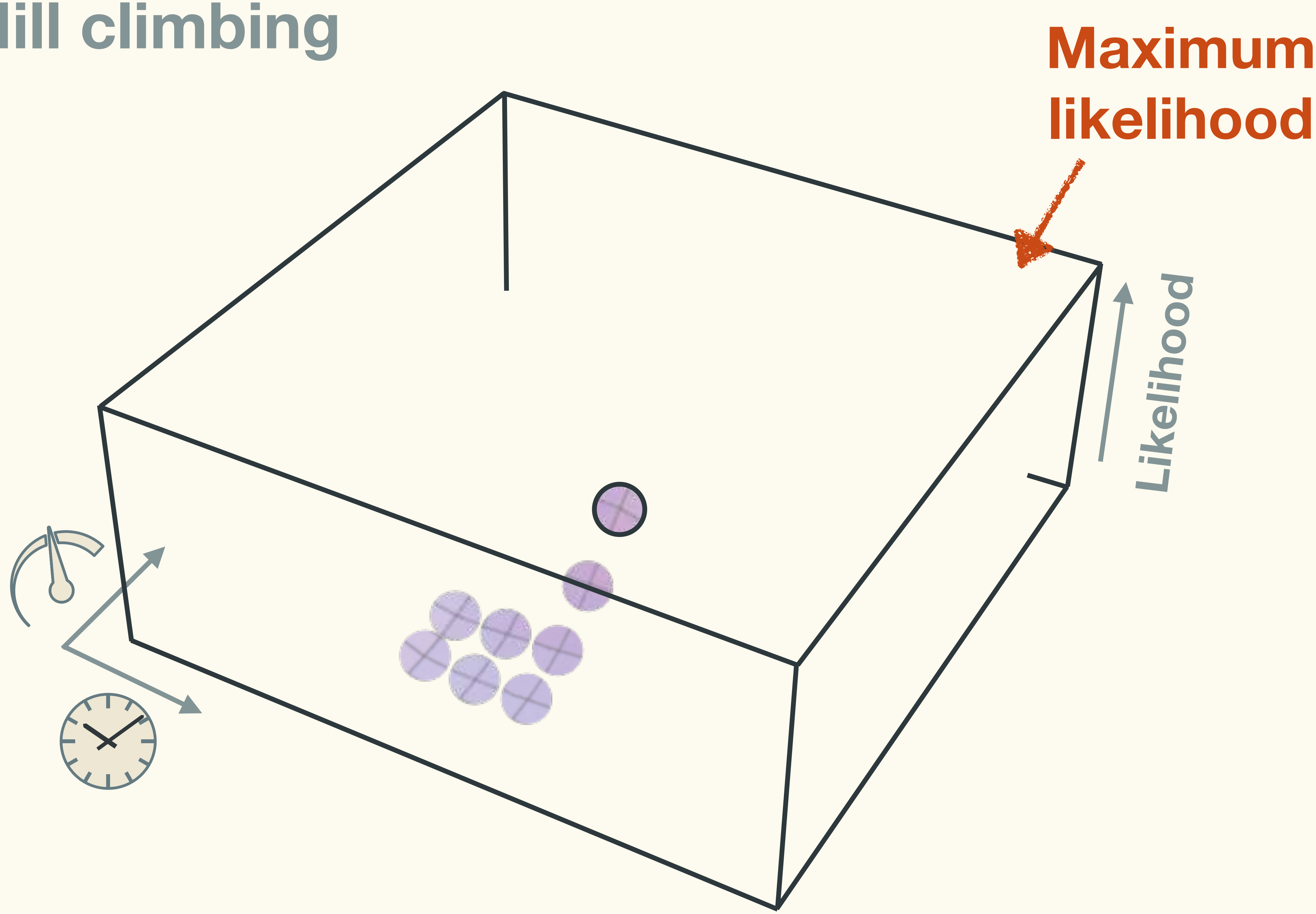
Heuristic search

Hill climbing



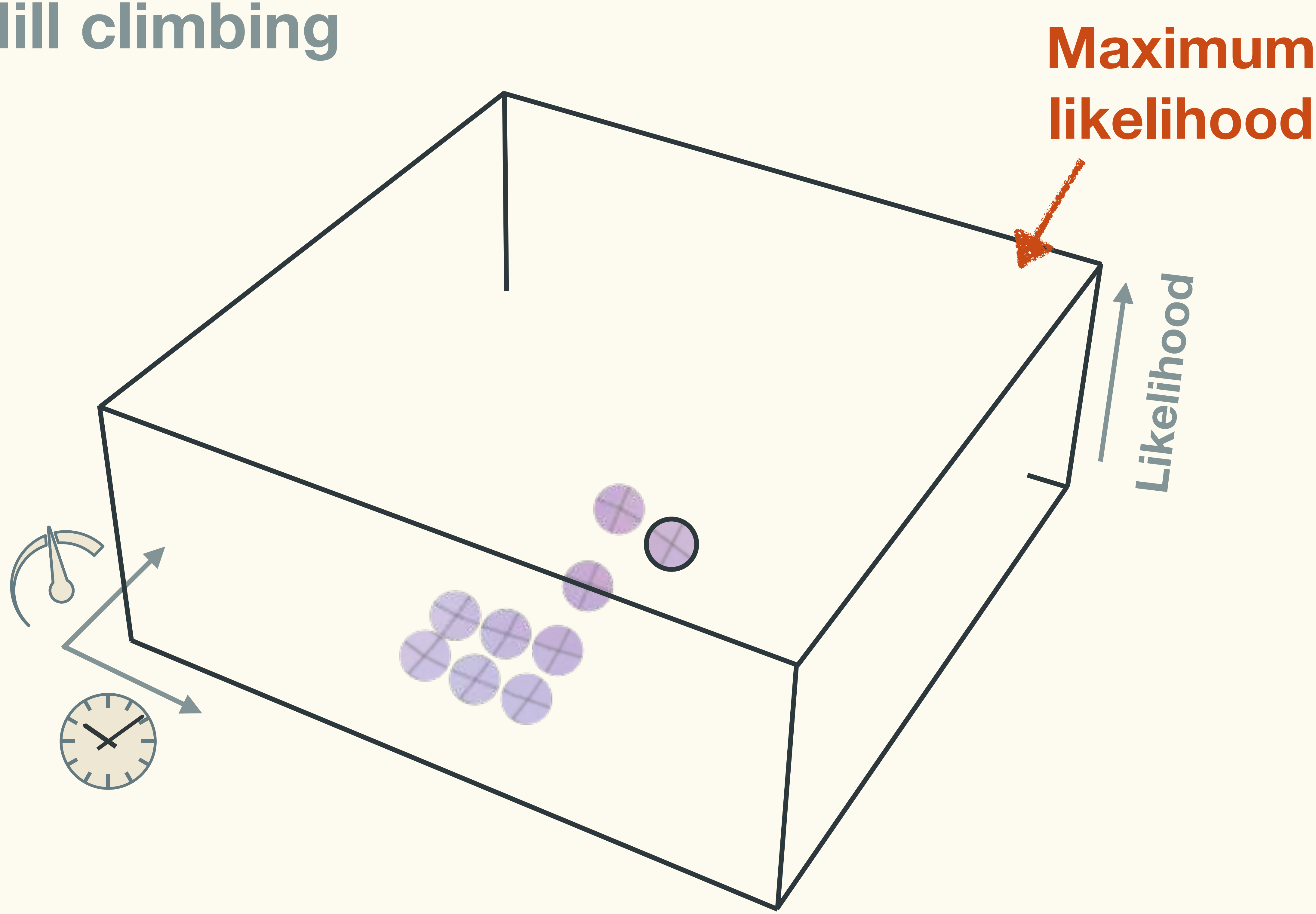
Heuristic search

Hill climbing



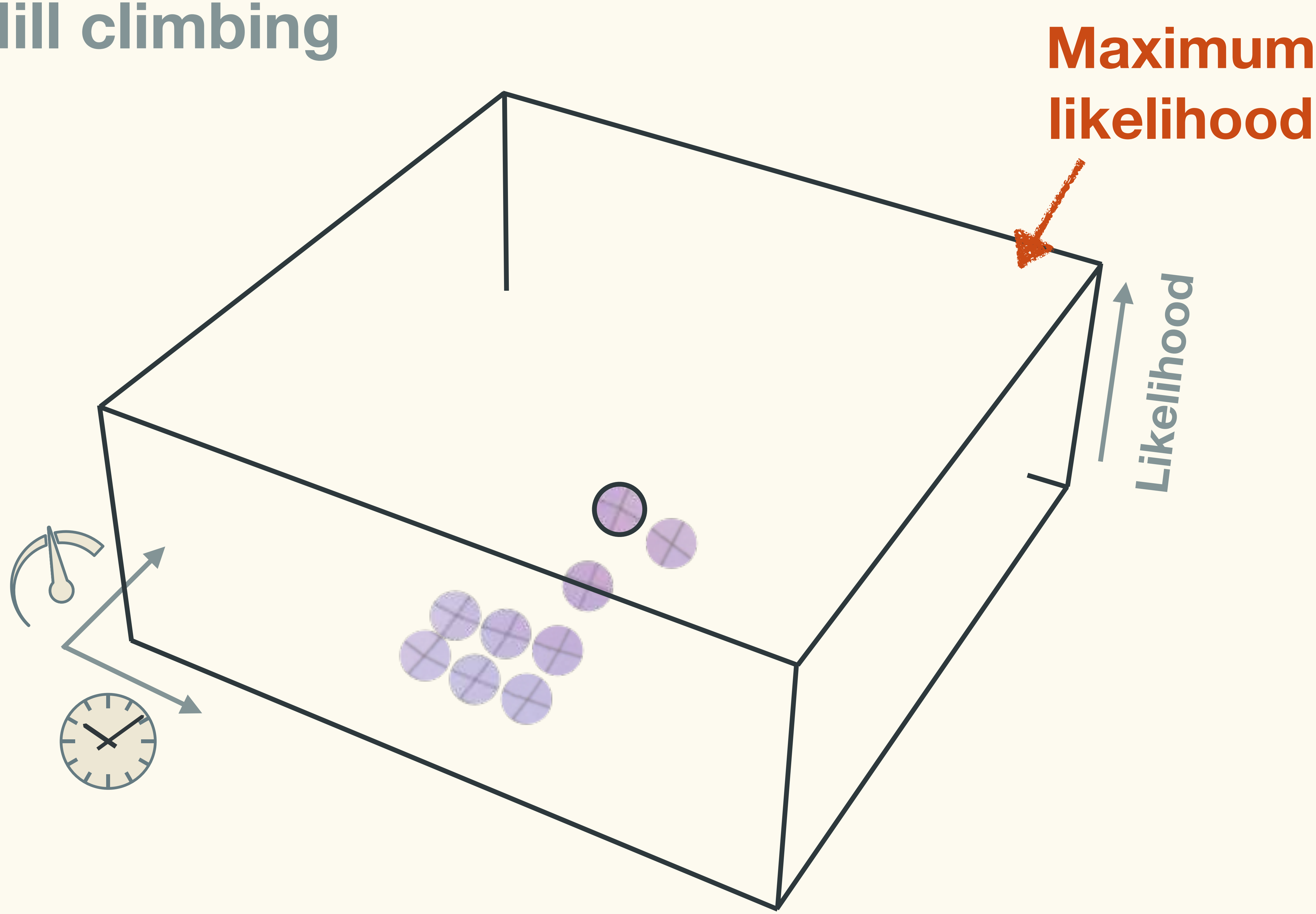
Heuristic search

Hill climbing



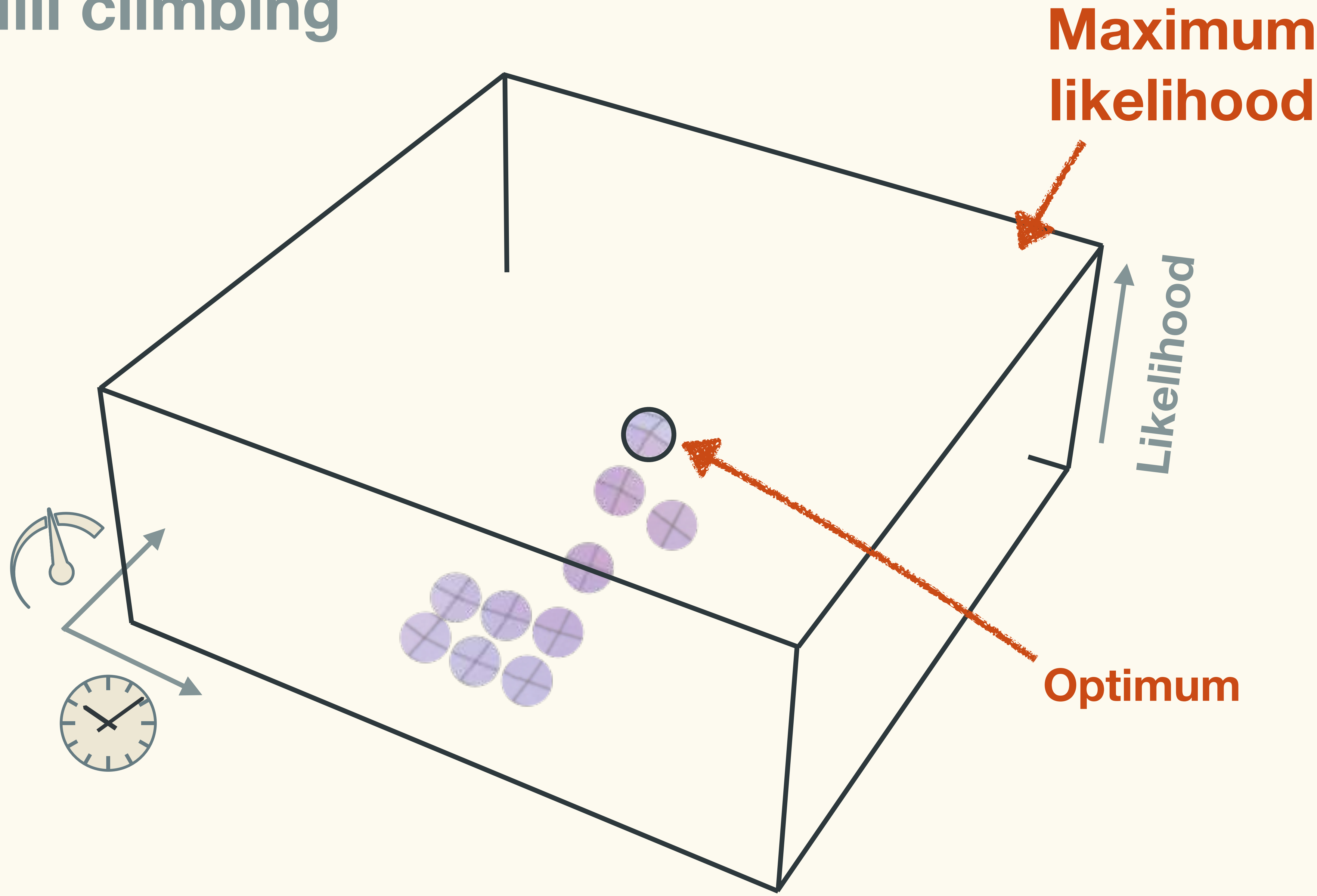
Heuristic search

Hill climbing



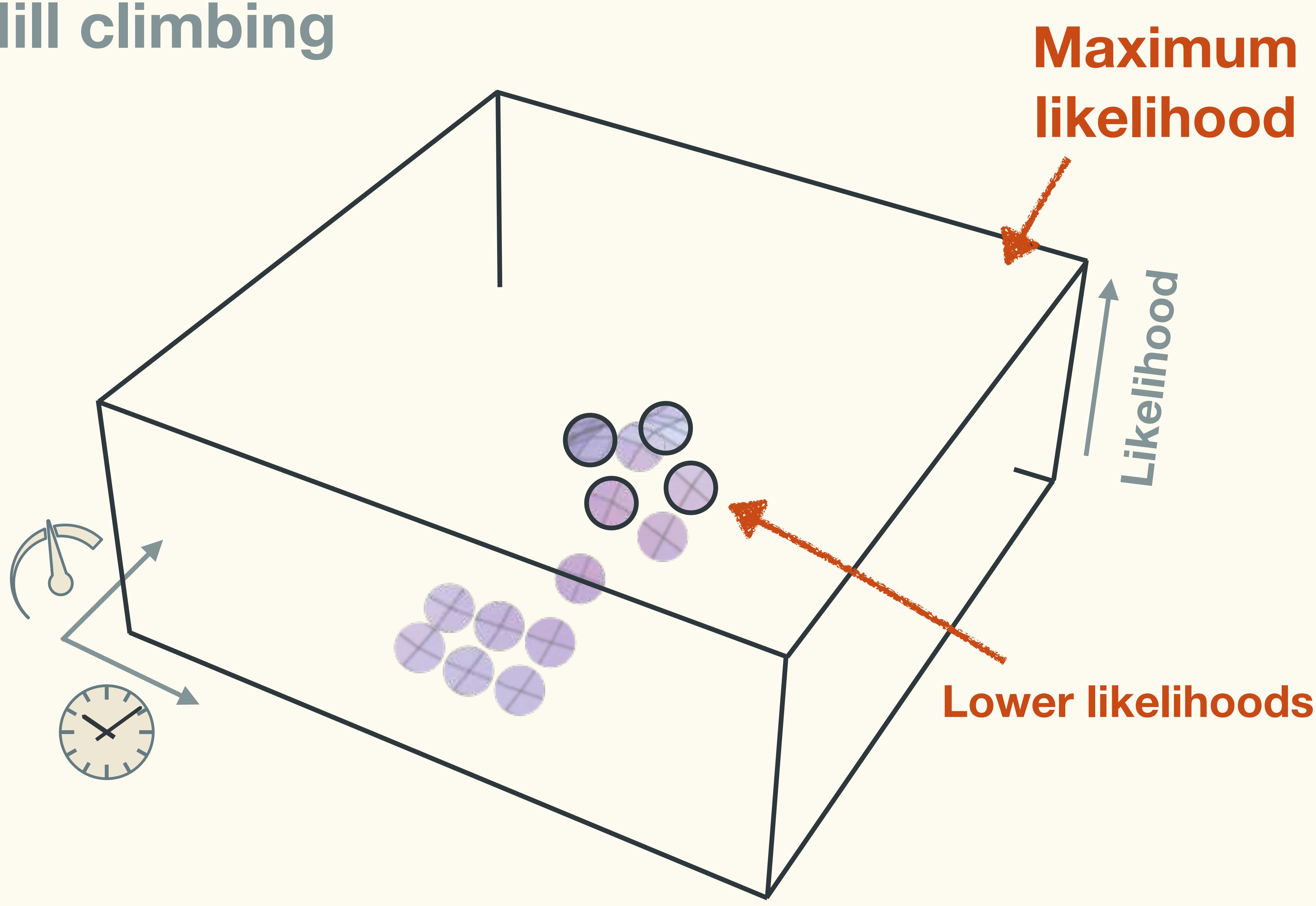
Heuristic search

Hill climbing



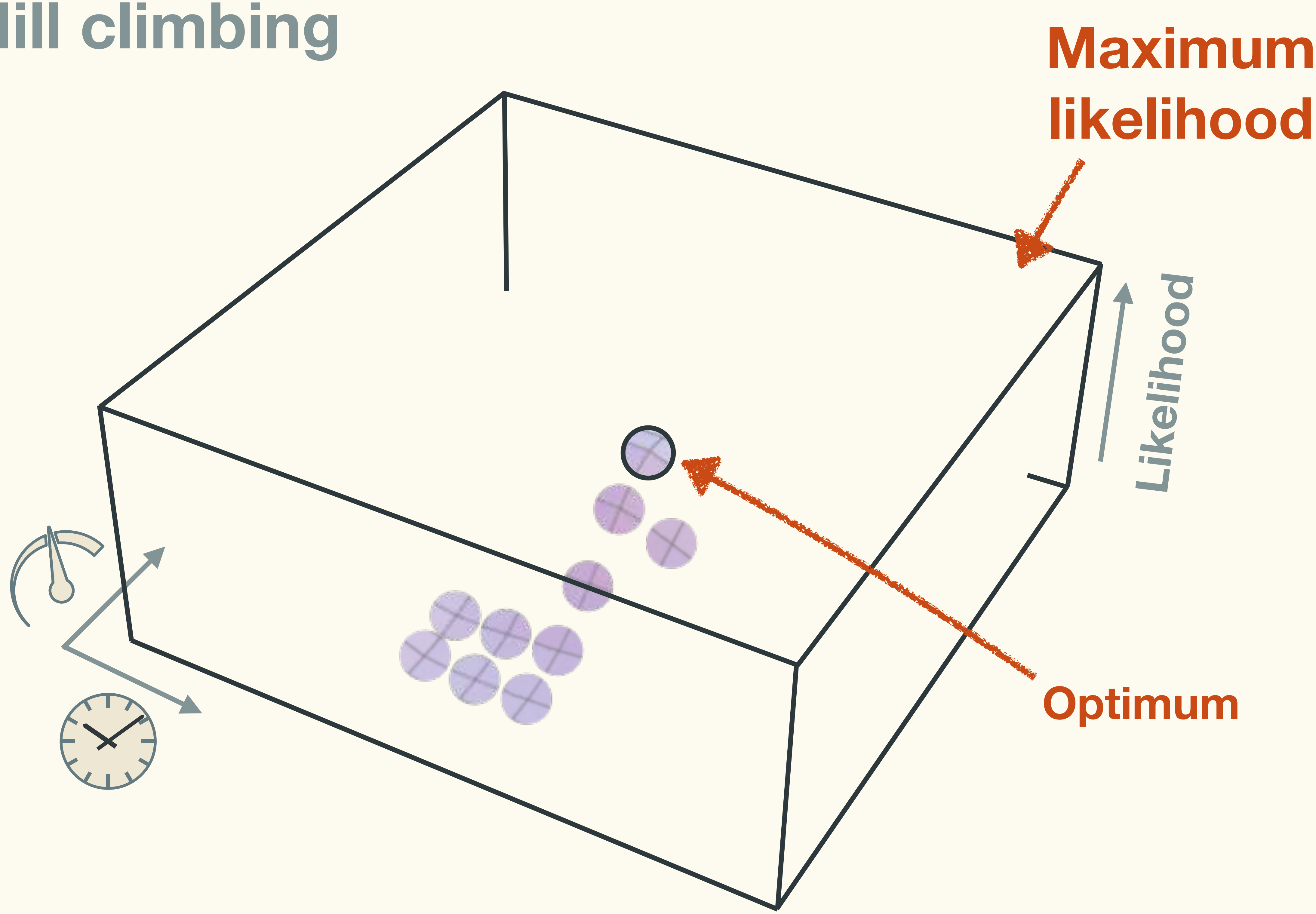
Heuristic search

Hill climbing



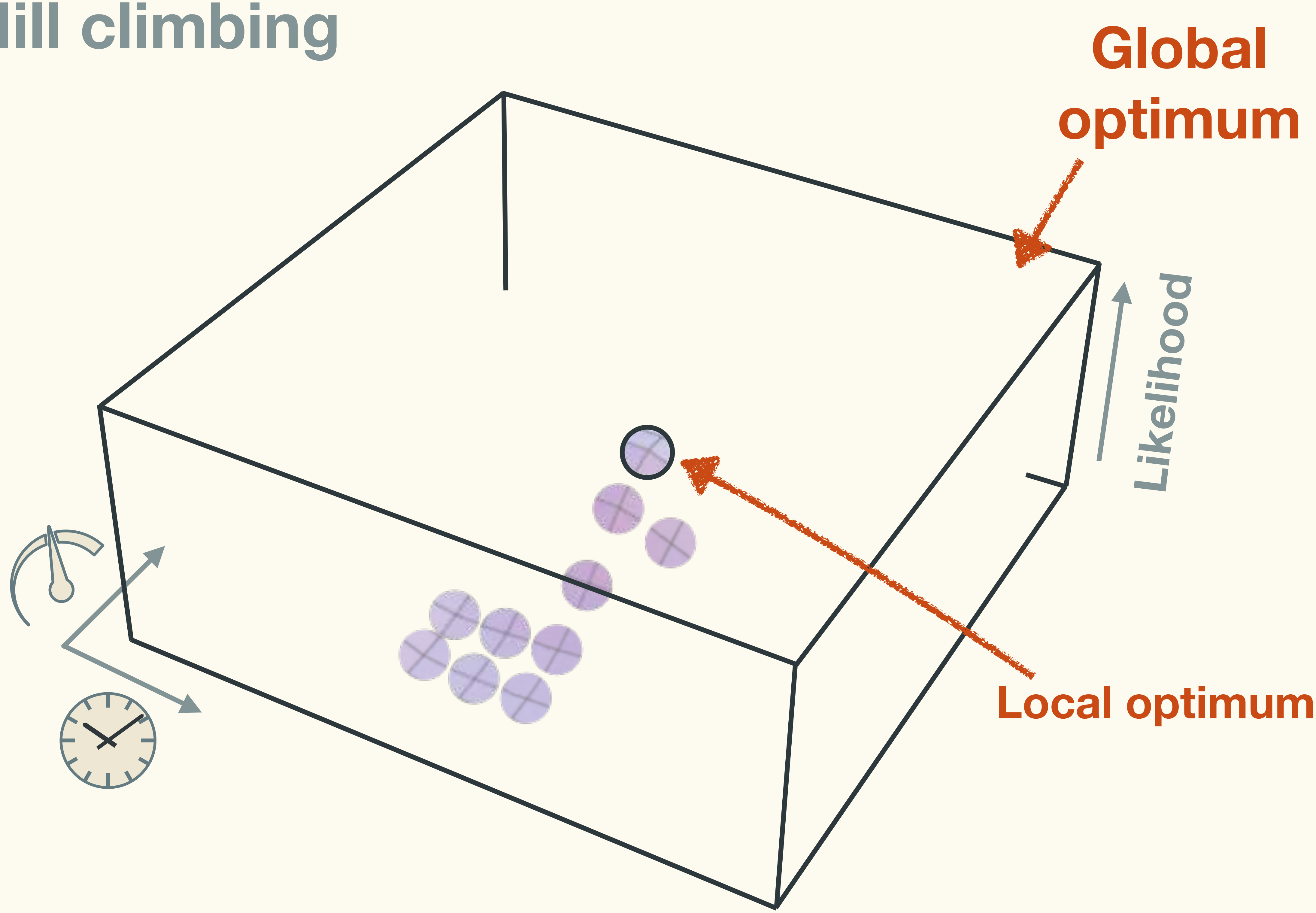
Heuristic search

Hill climbing



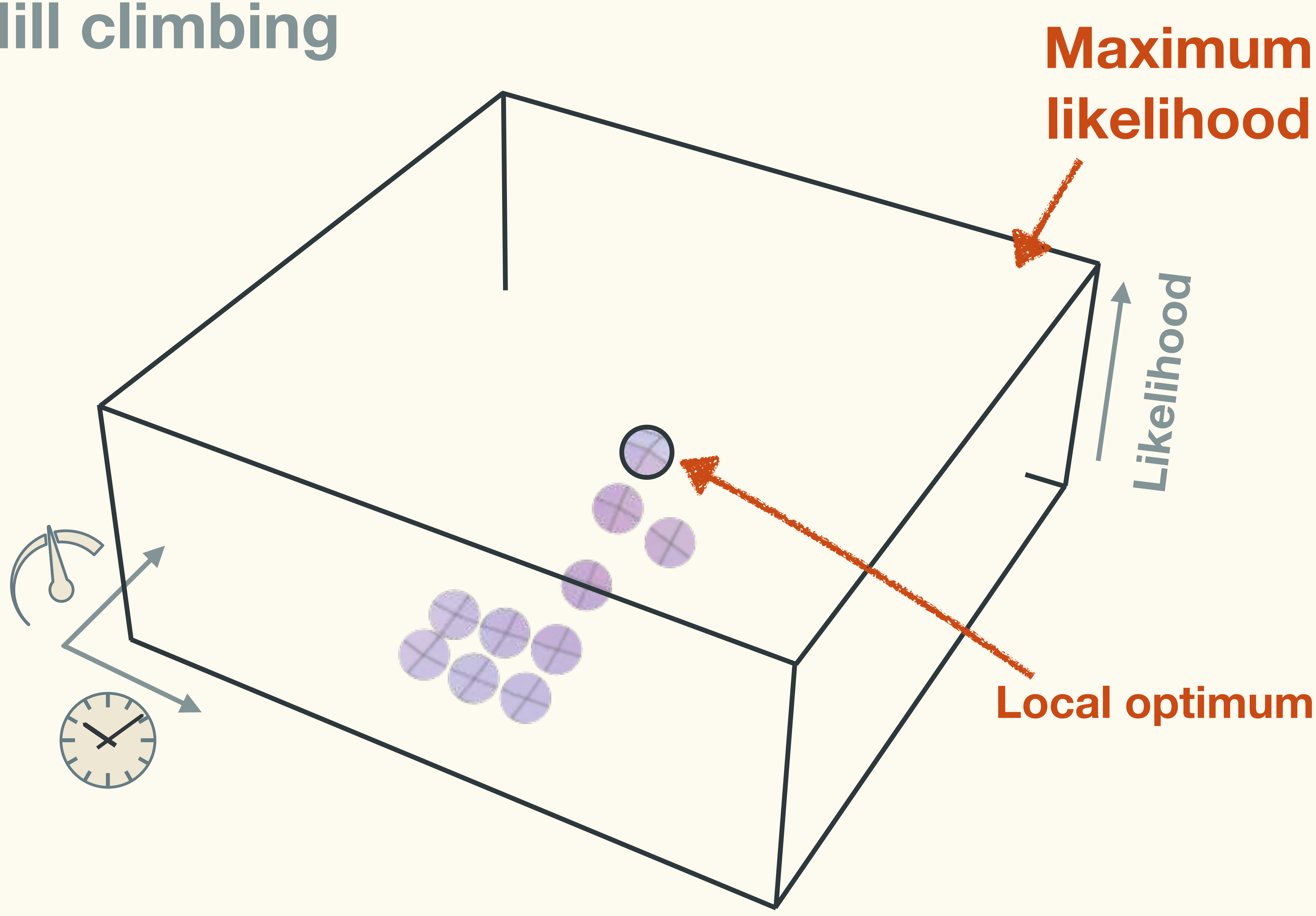
Heuristic search

Hill climbing



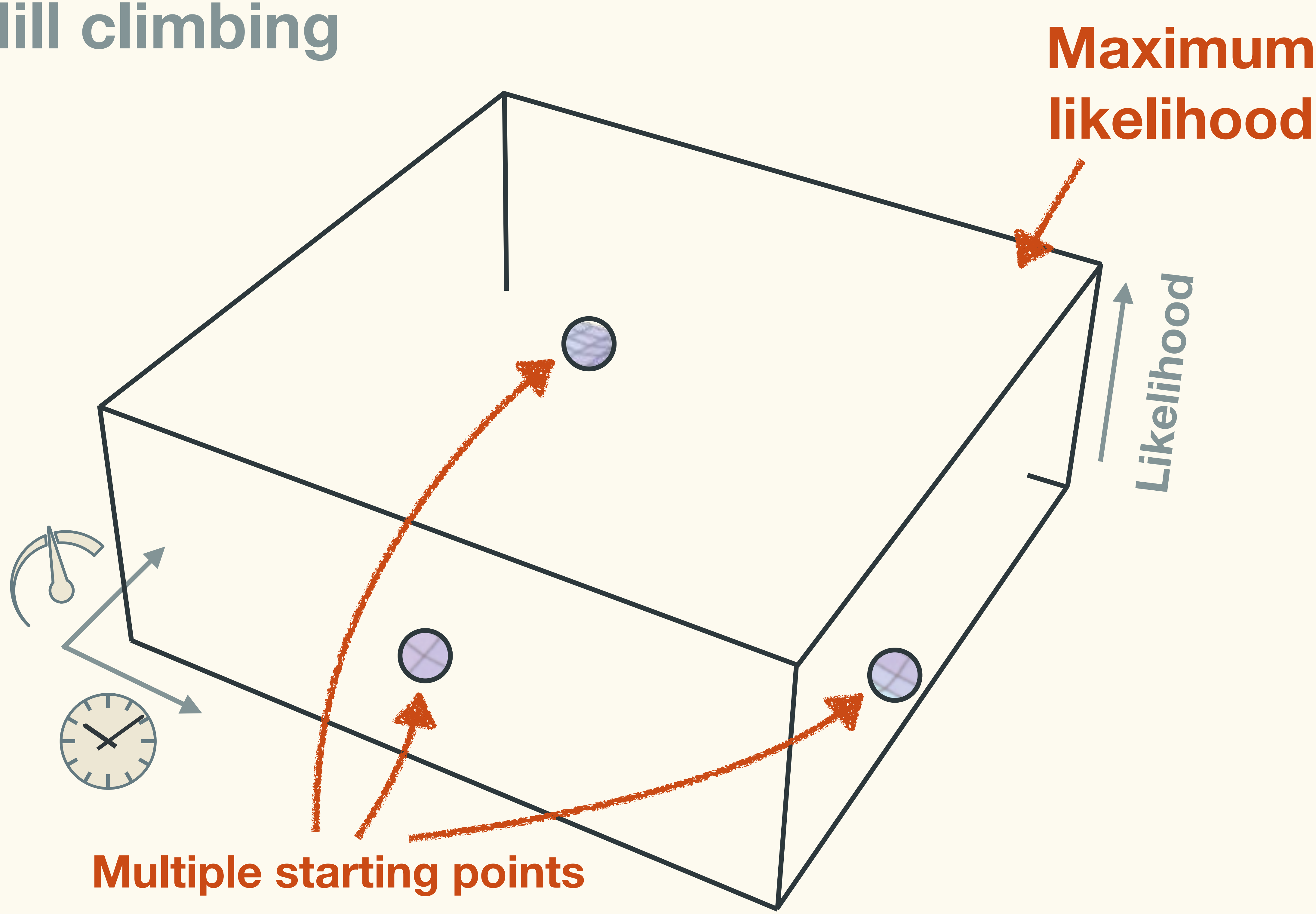
Heuristic search

Hill climbing



Heuristic search

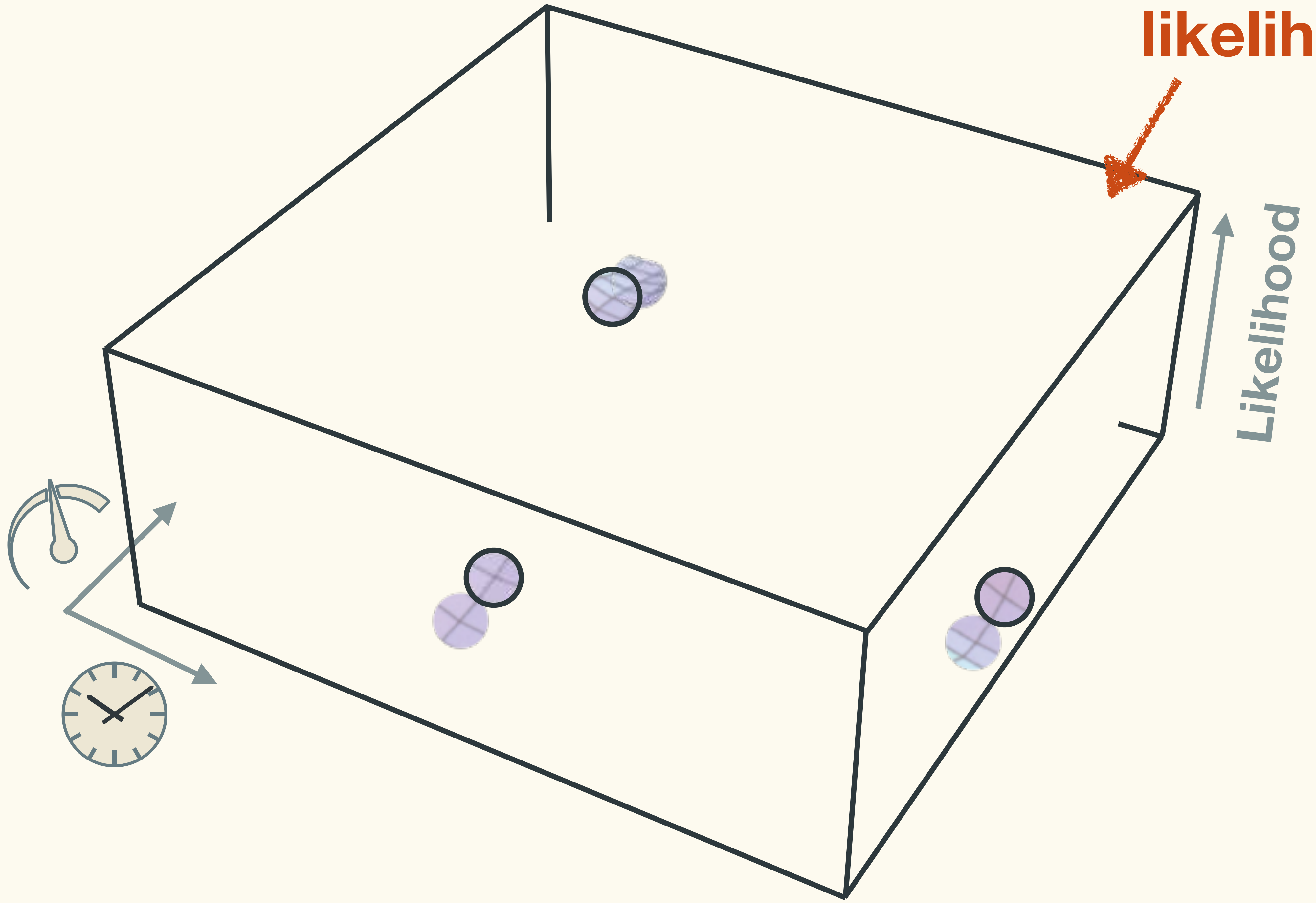
Hill climbing



Heuristic search

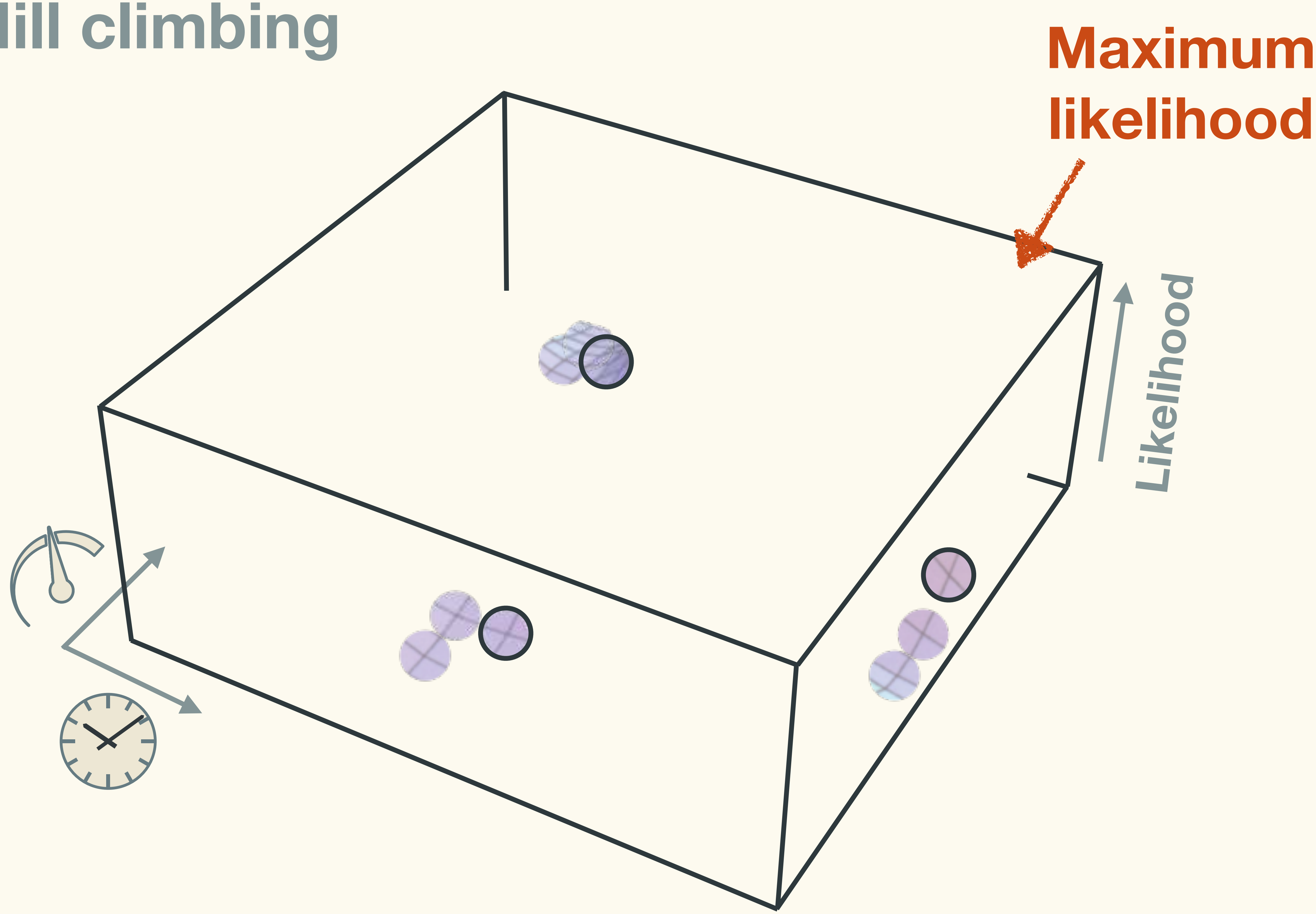
Hill climbing

Maximum likelihood



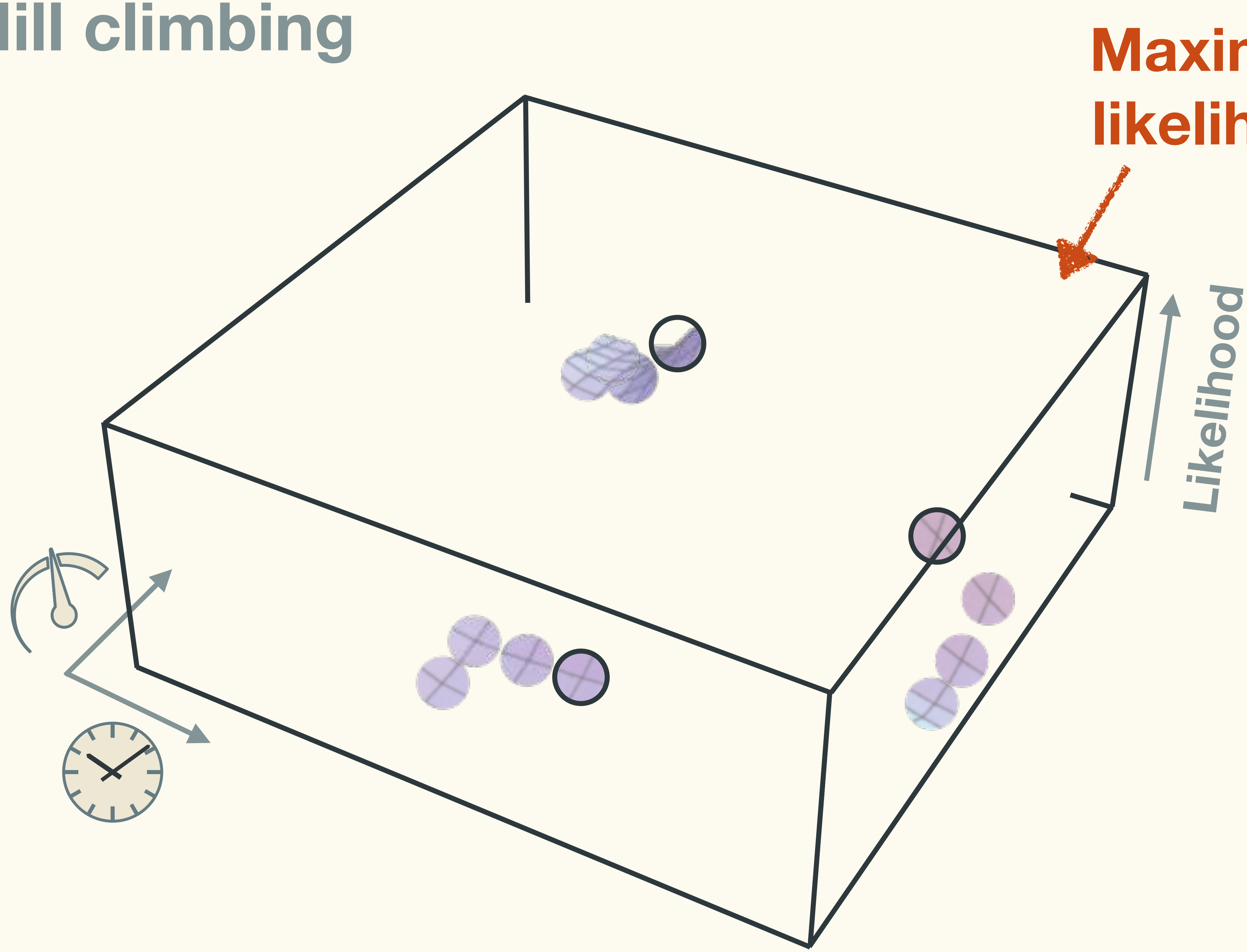
Heuristic search

Hill climbing



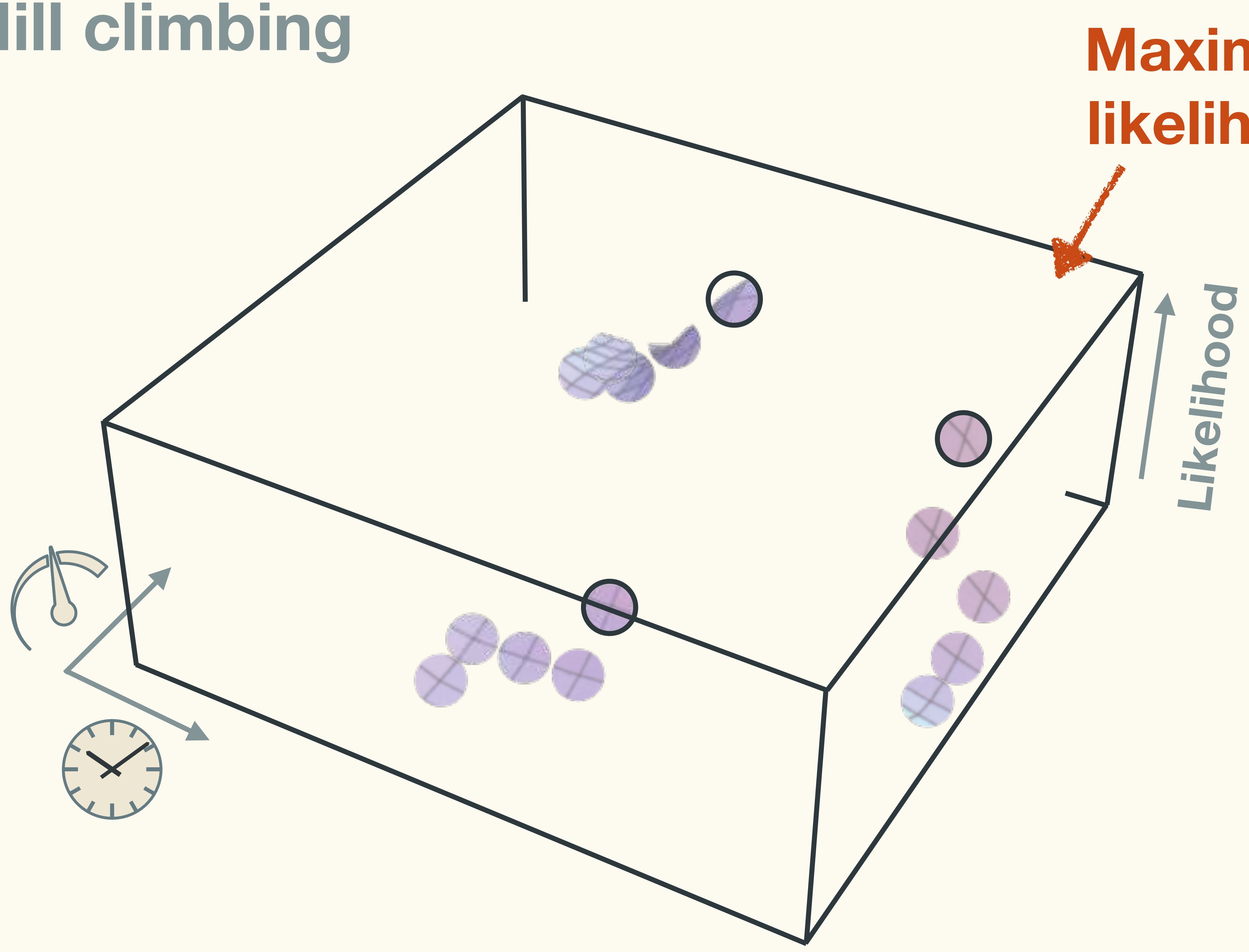
Heuristic search

Hill climbing



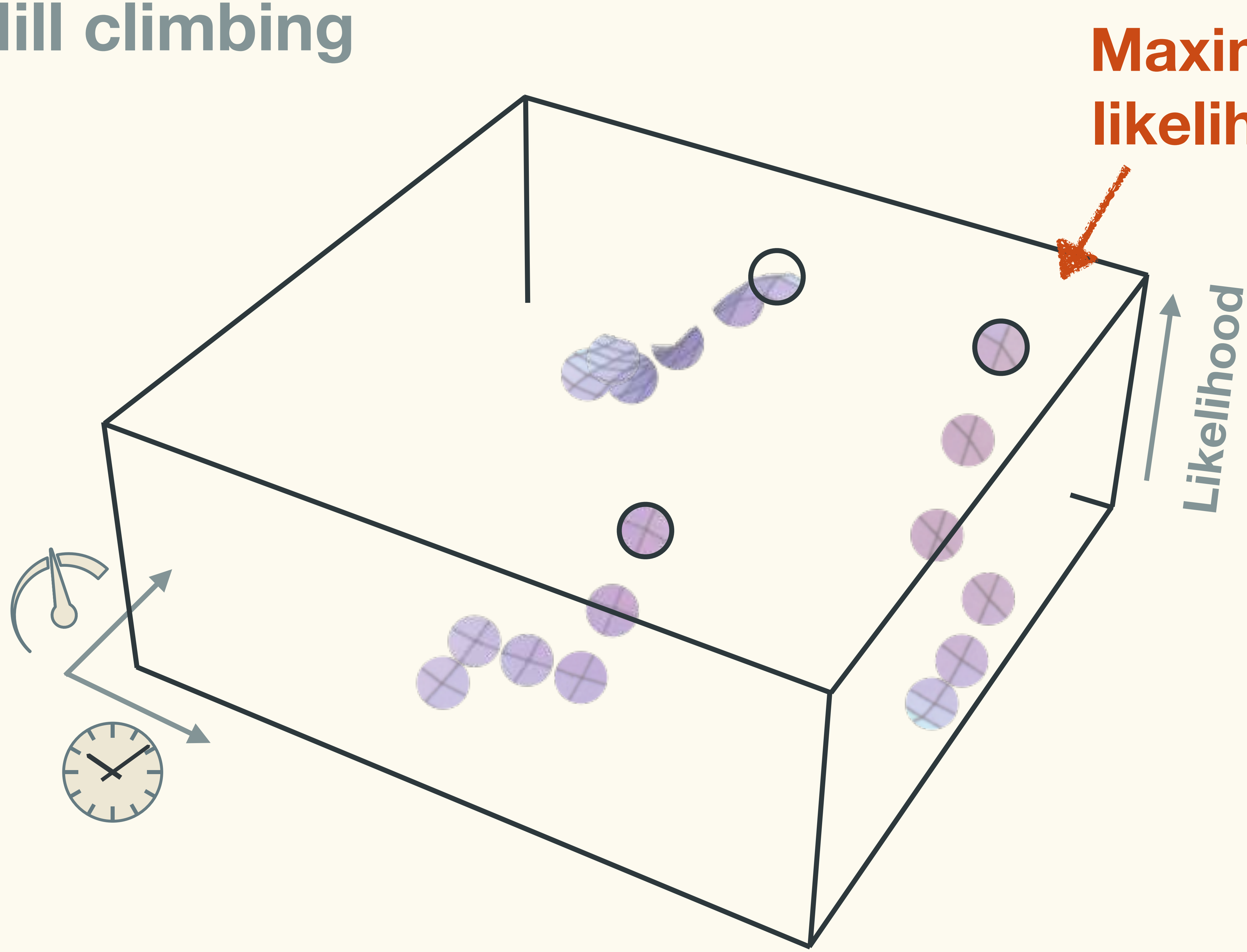
Heuristic search

Hill climbing

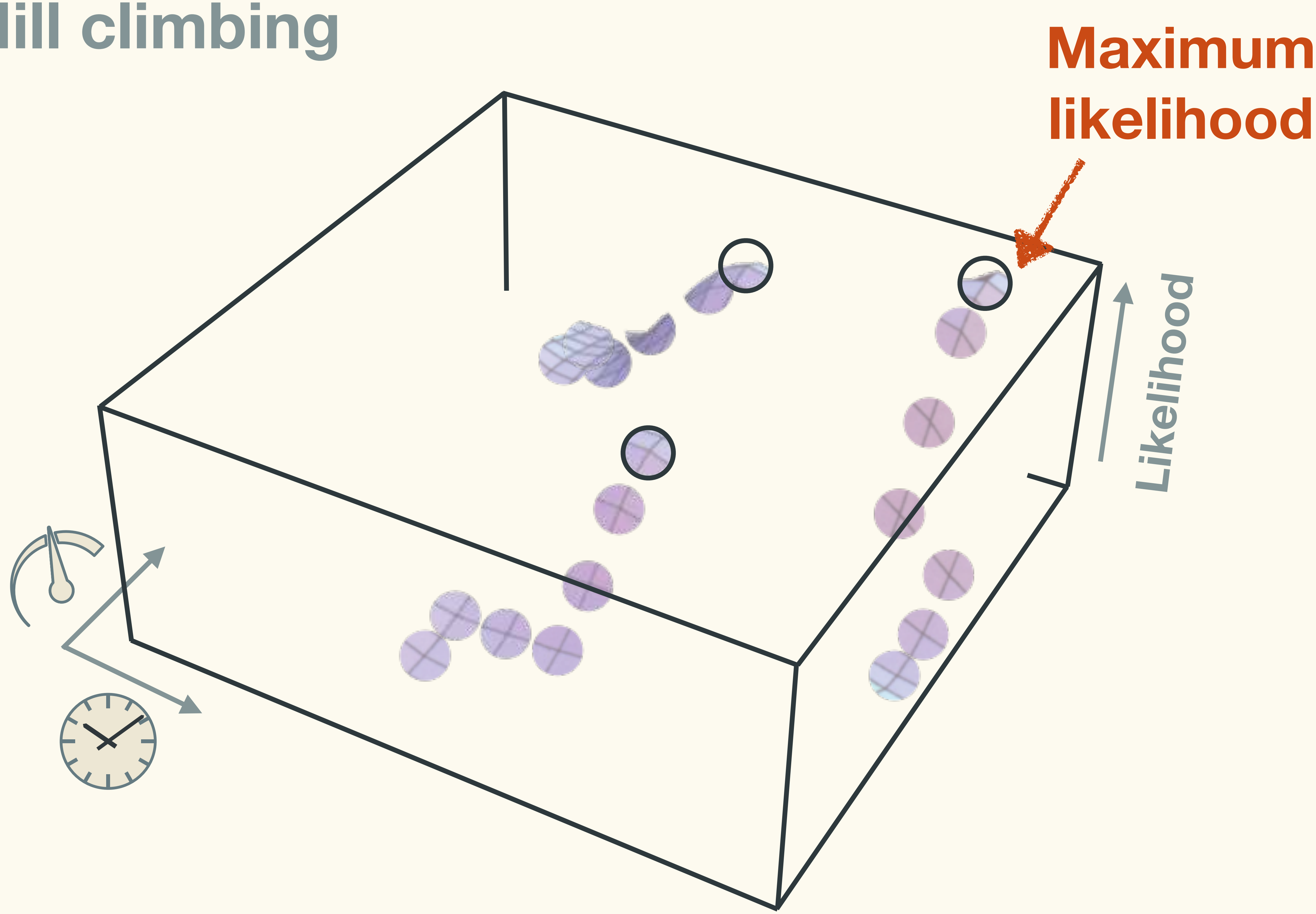


Heuristic search

Hill climbing

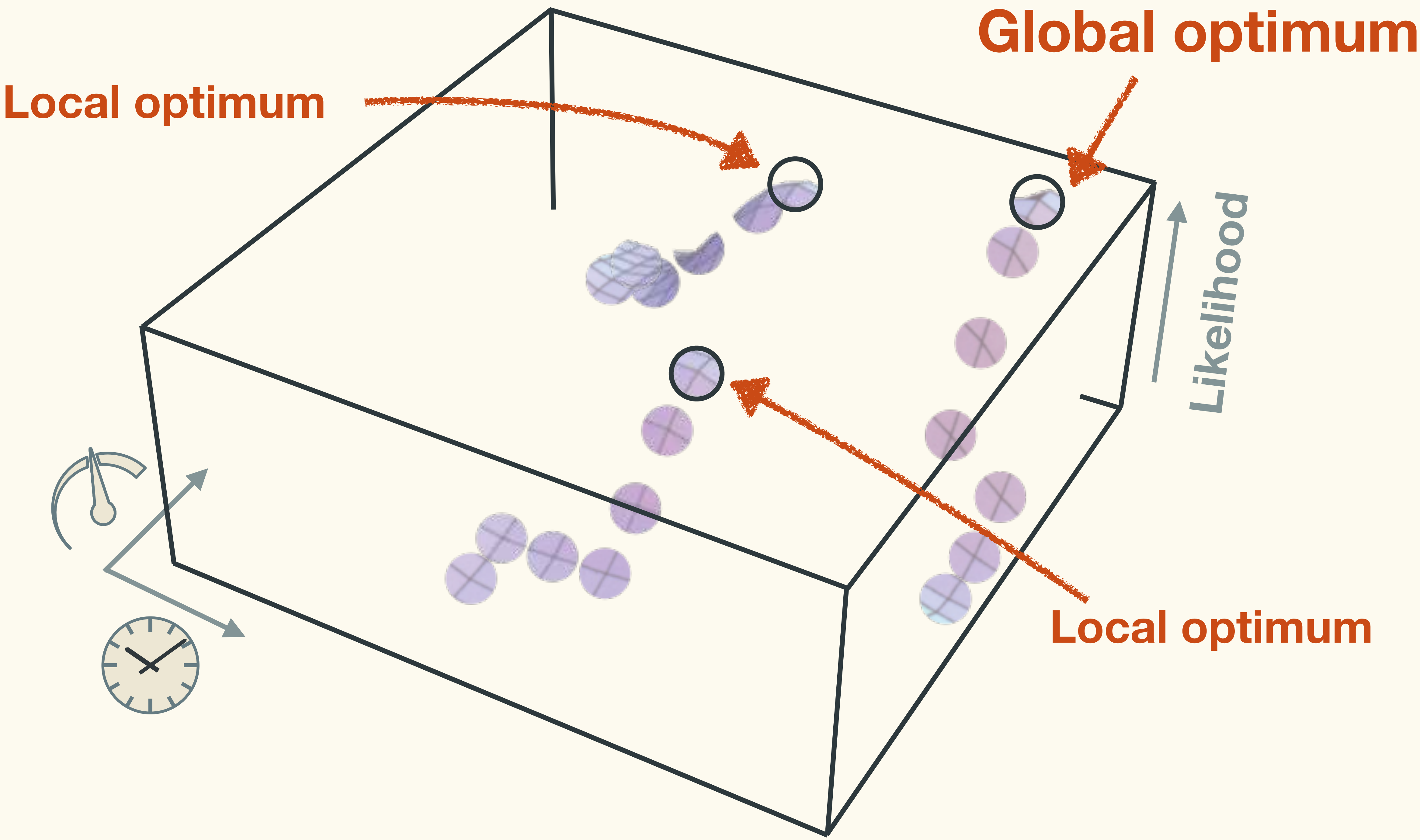


Heuristic search
Hill climbing



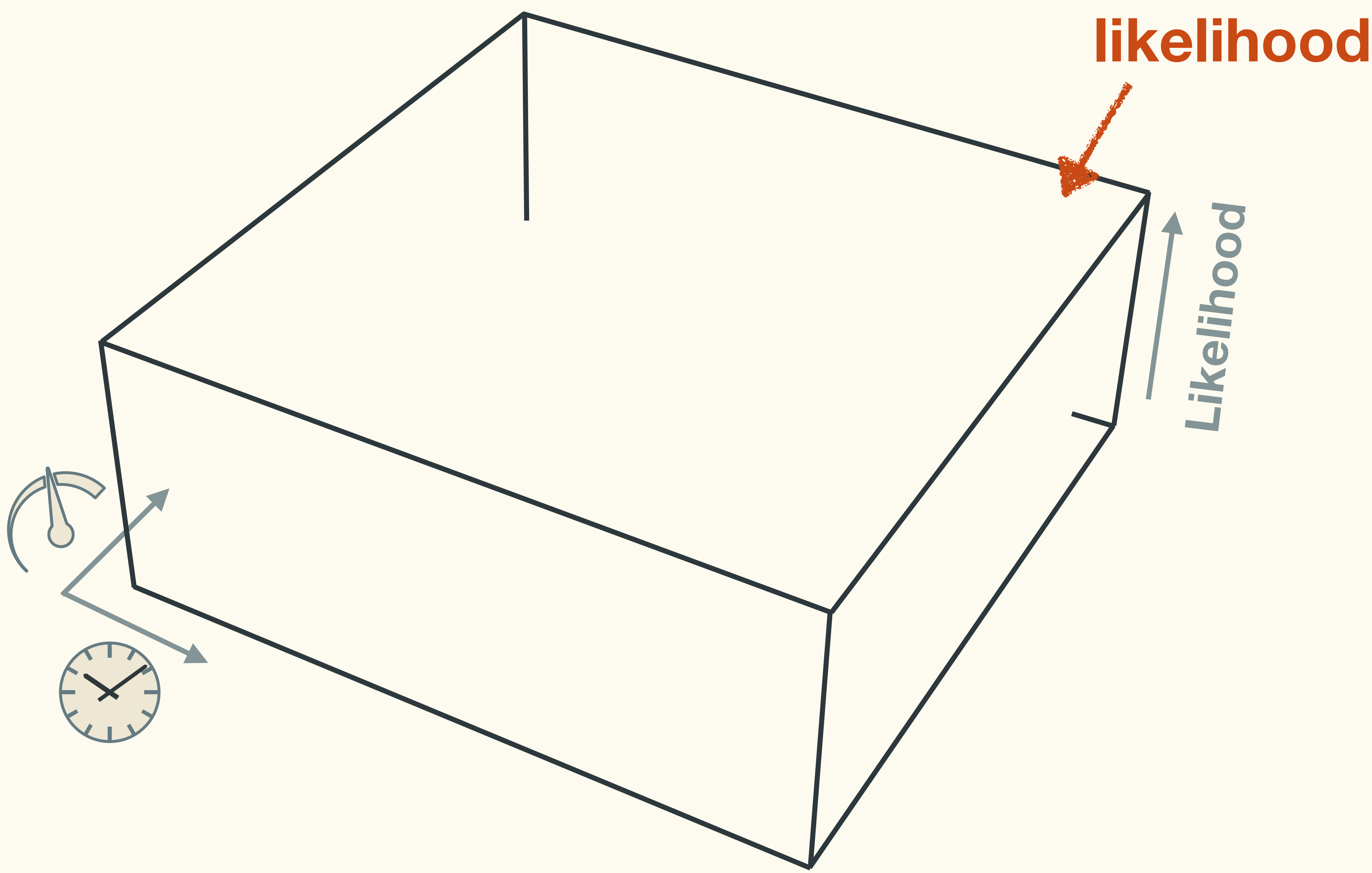
Heuristic search

Hill climbing



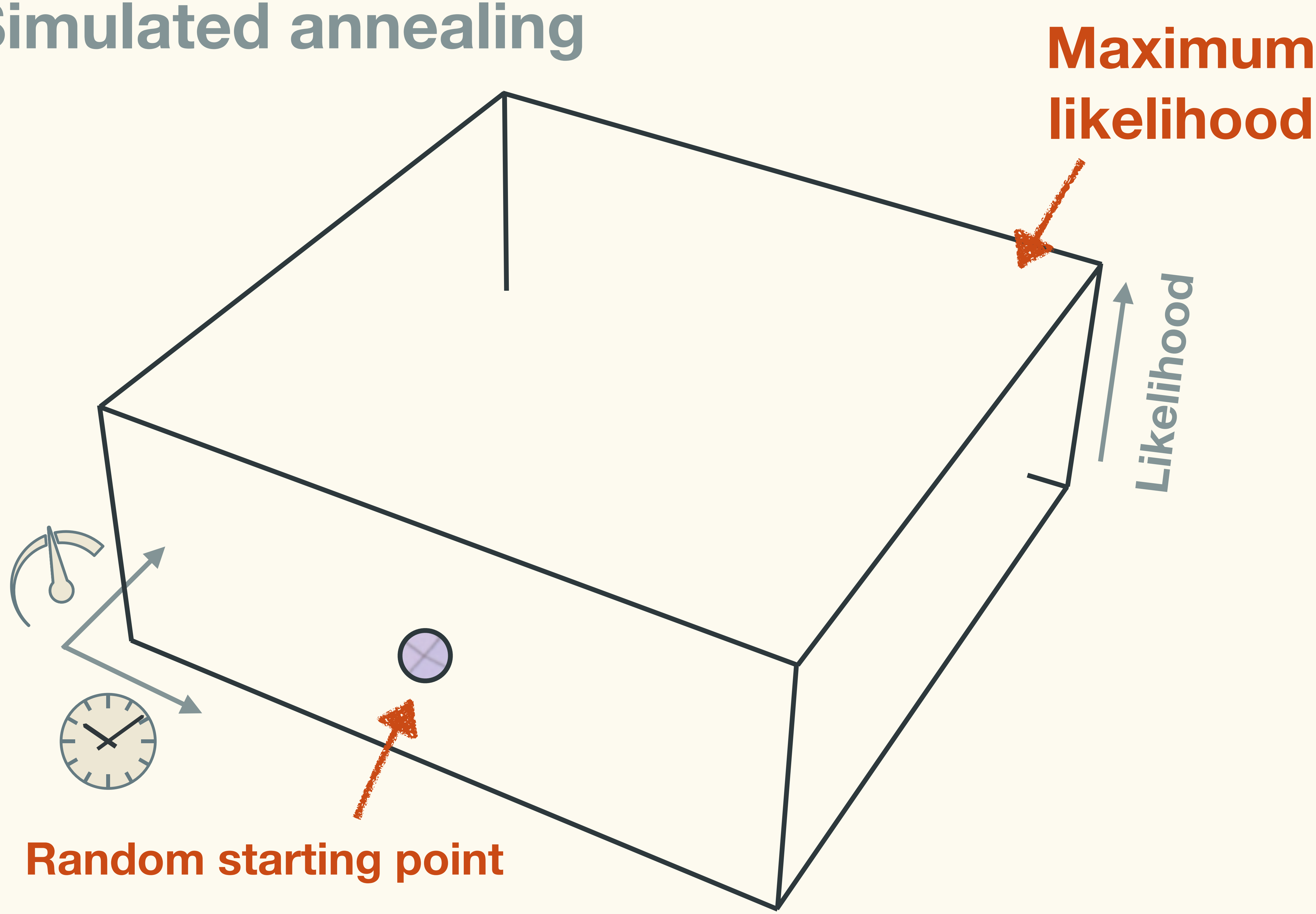
Heuristic search

Simulated annealing



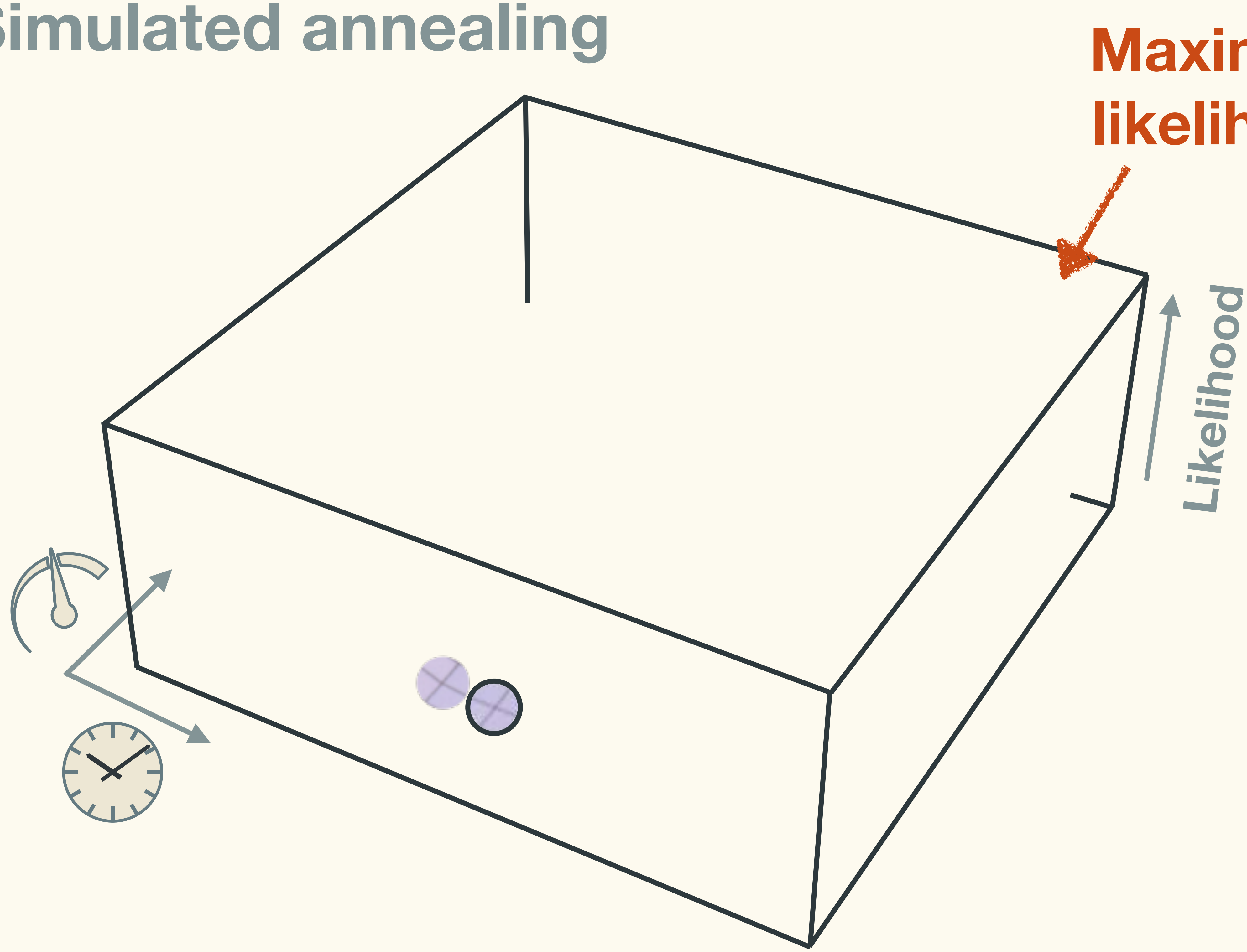
Heuristic search

Simulated annealing



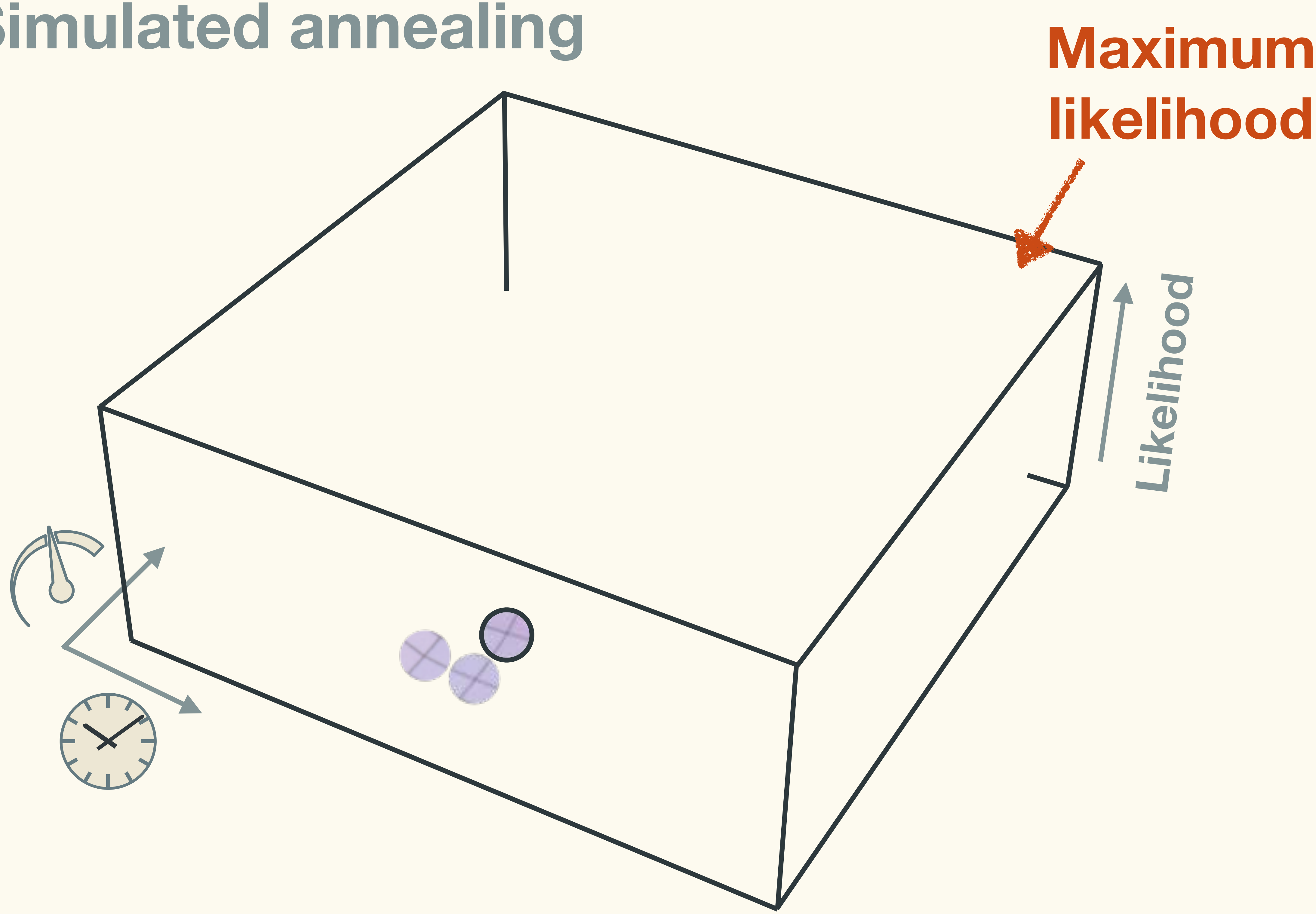
Heuristic search

Simulated annealing



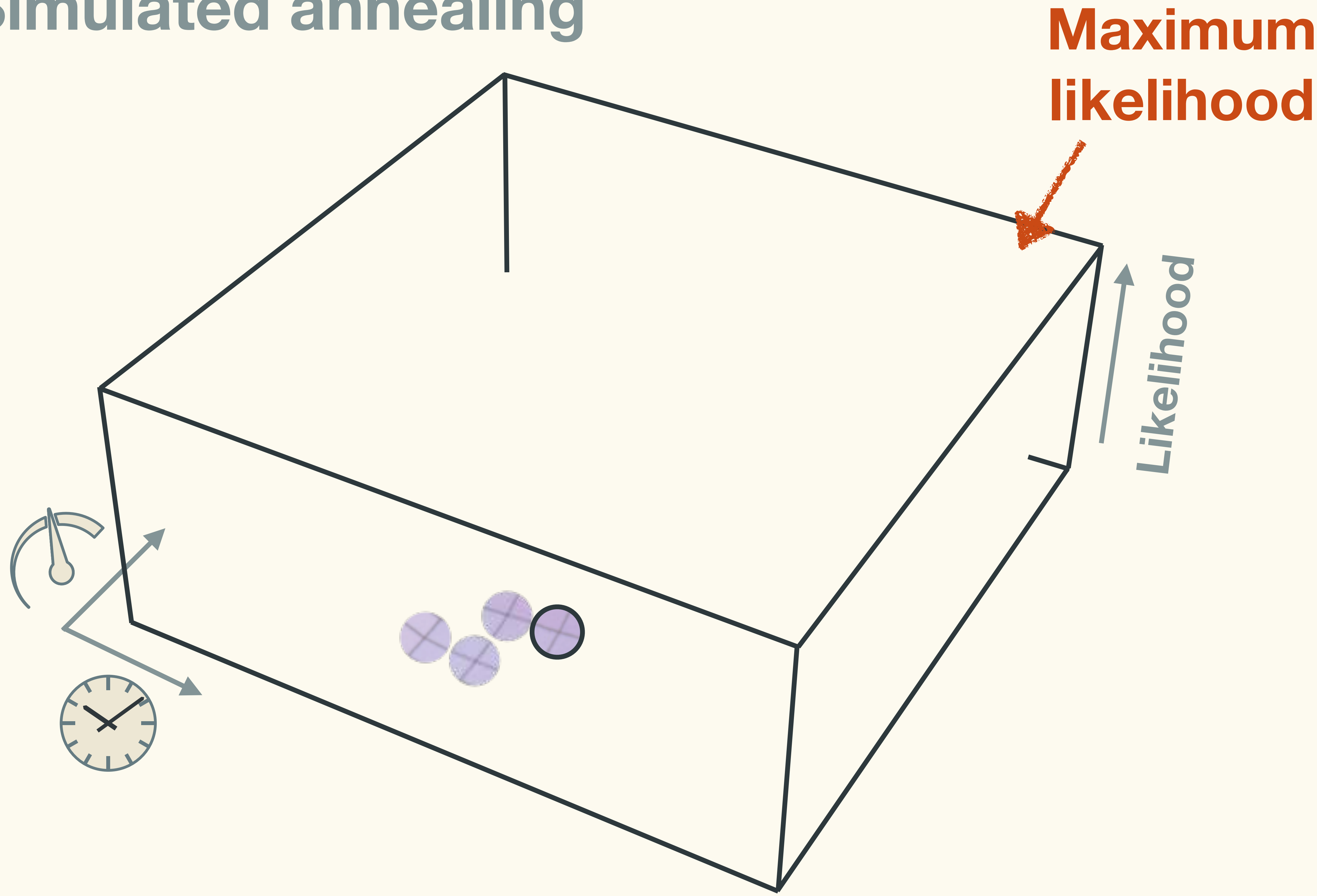
Heuristic search

Simulated annealing



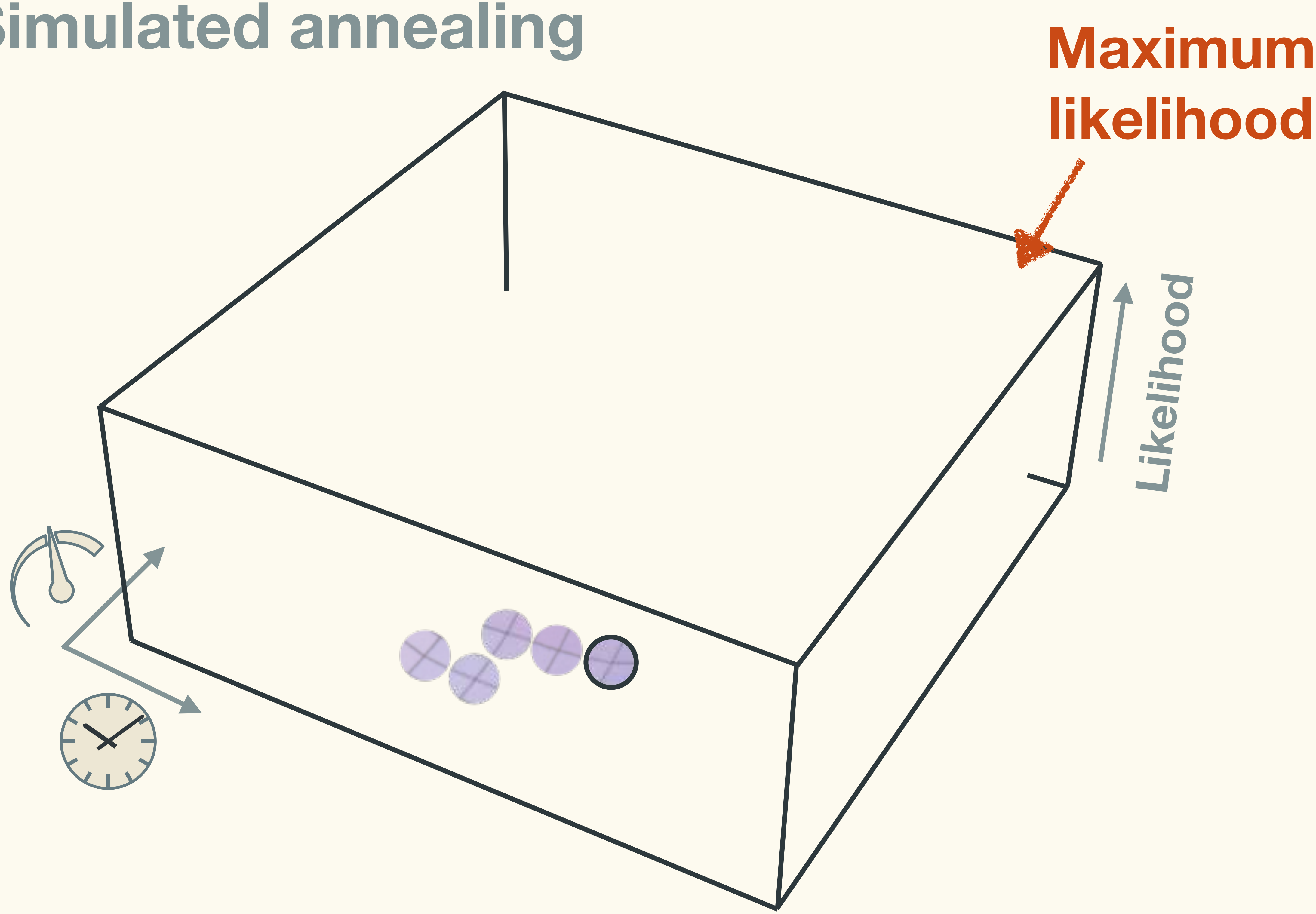
Heuristic search

Simulated annealing



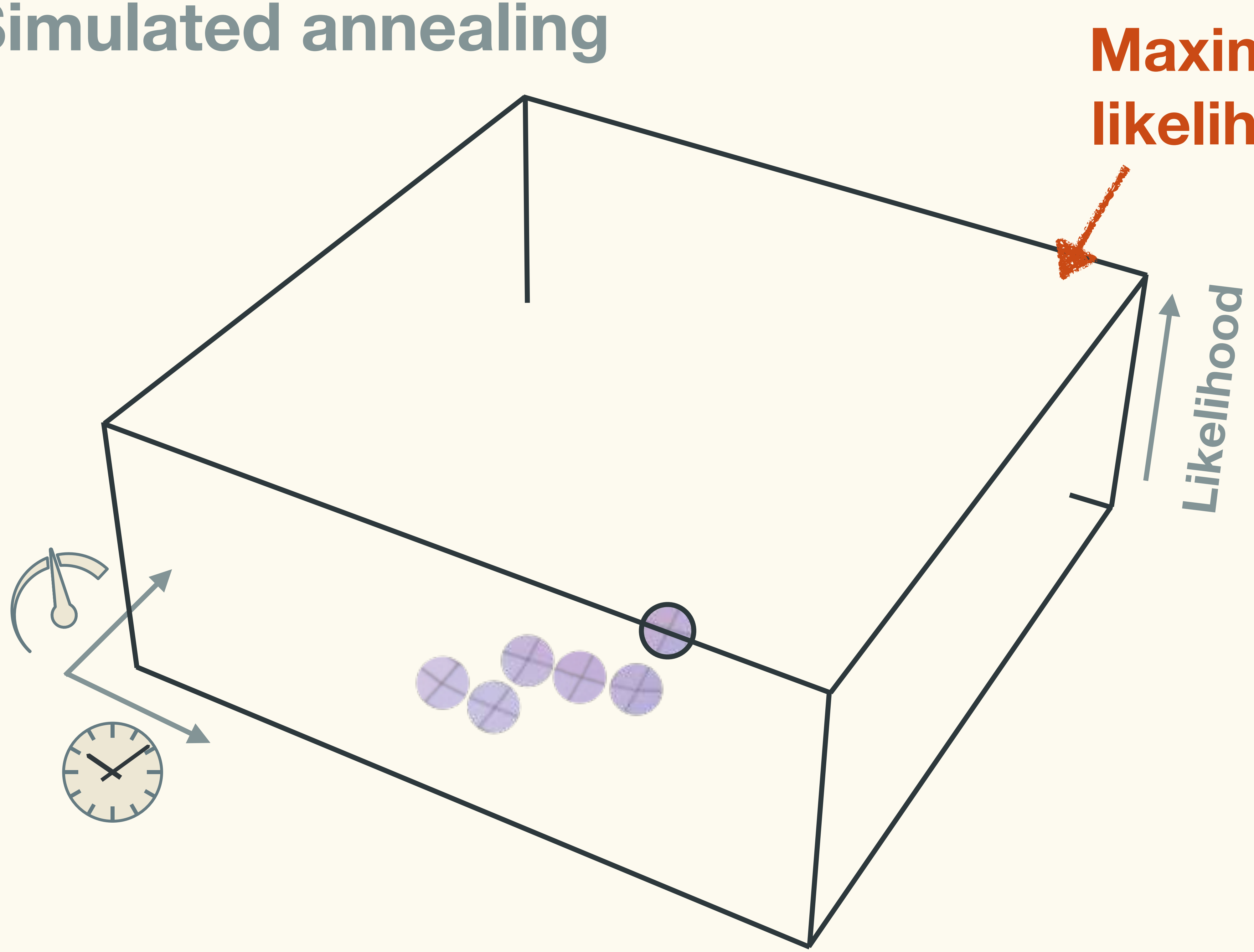
Heuristic search

Simulated annealing



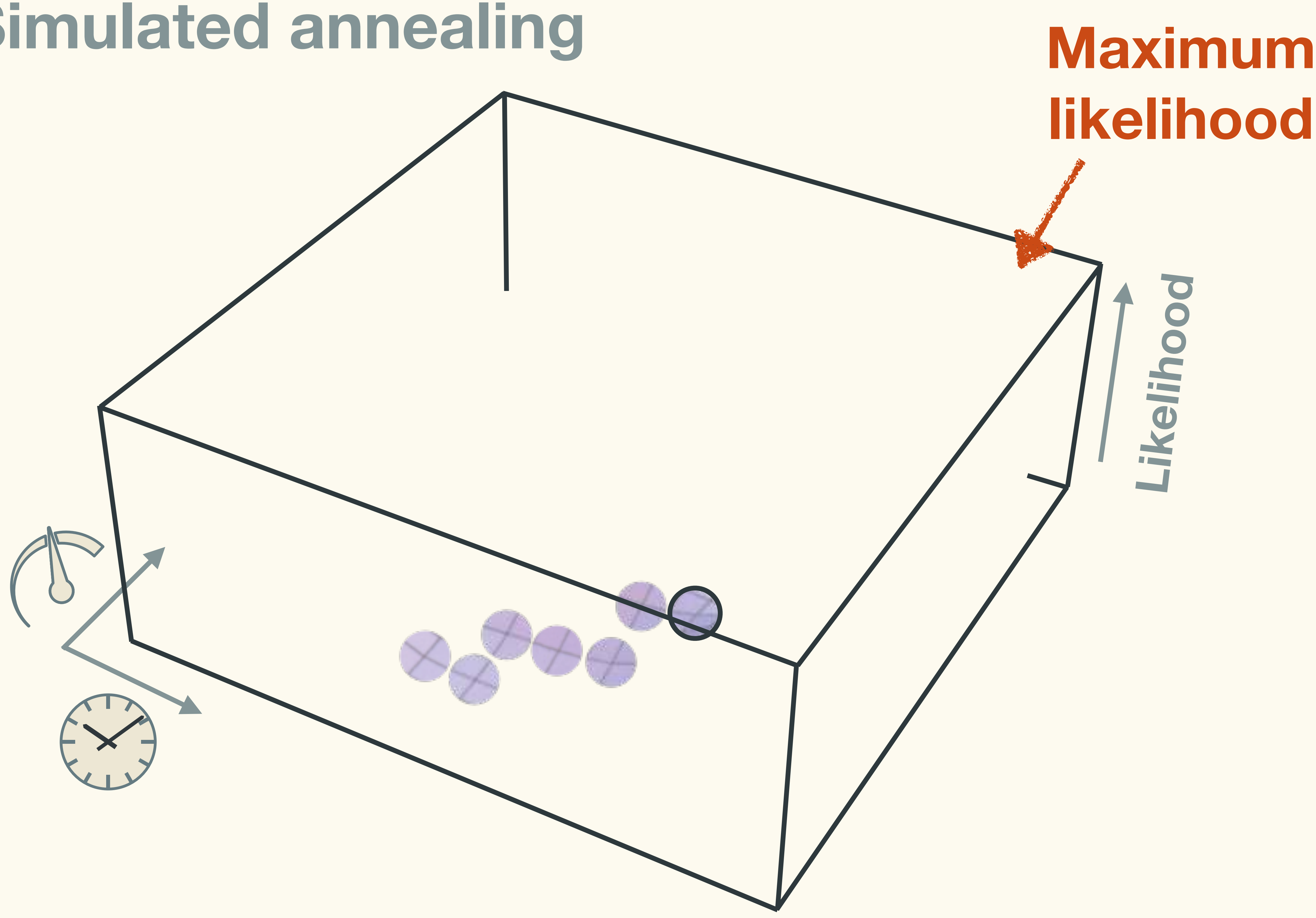
Heuristic search

Simulated annealing



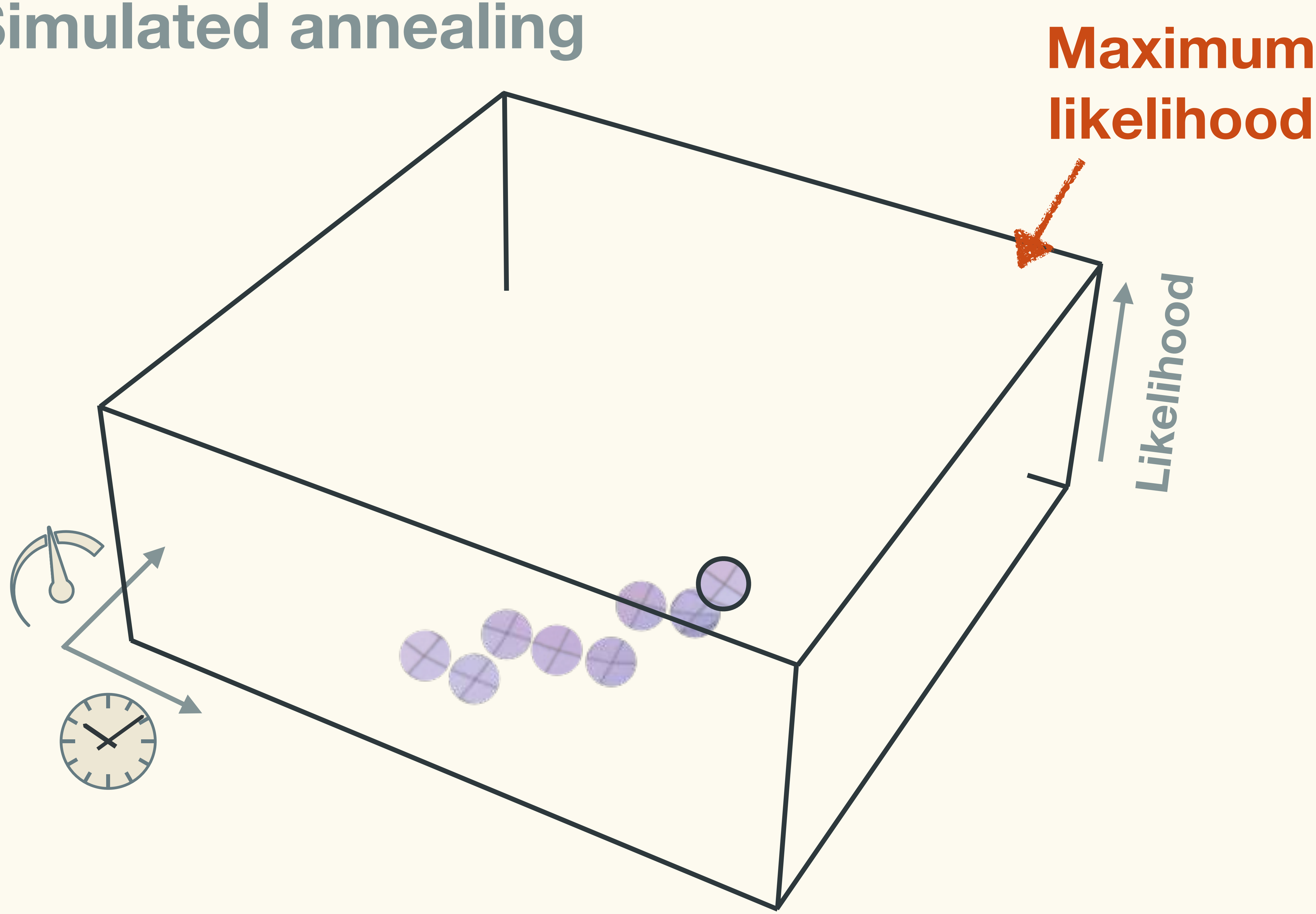
Heuristic search

Simulated annealing



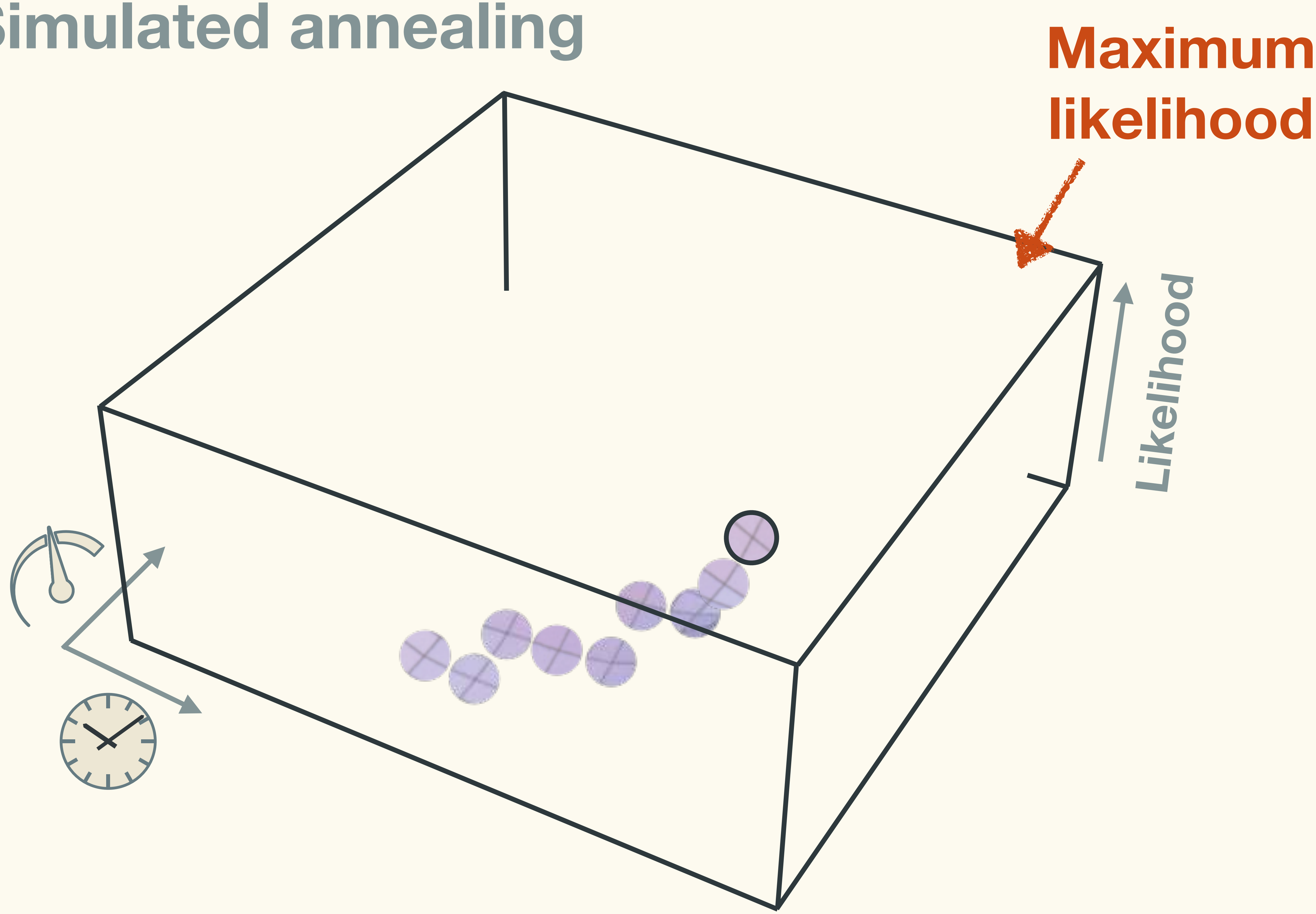
Heuristic search

Simulated annealing



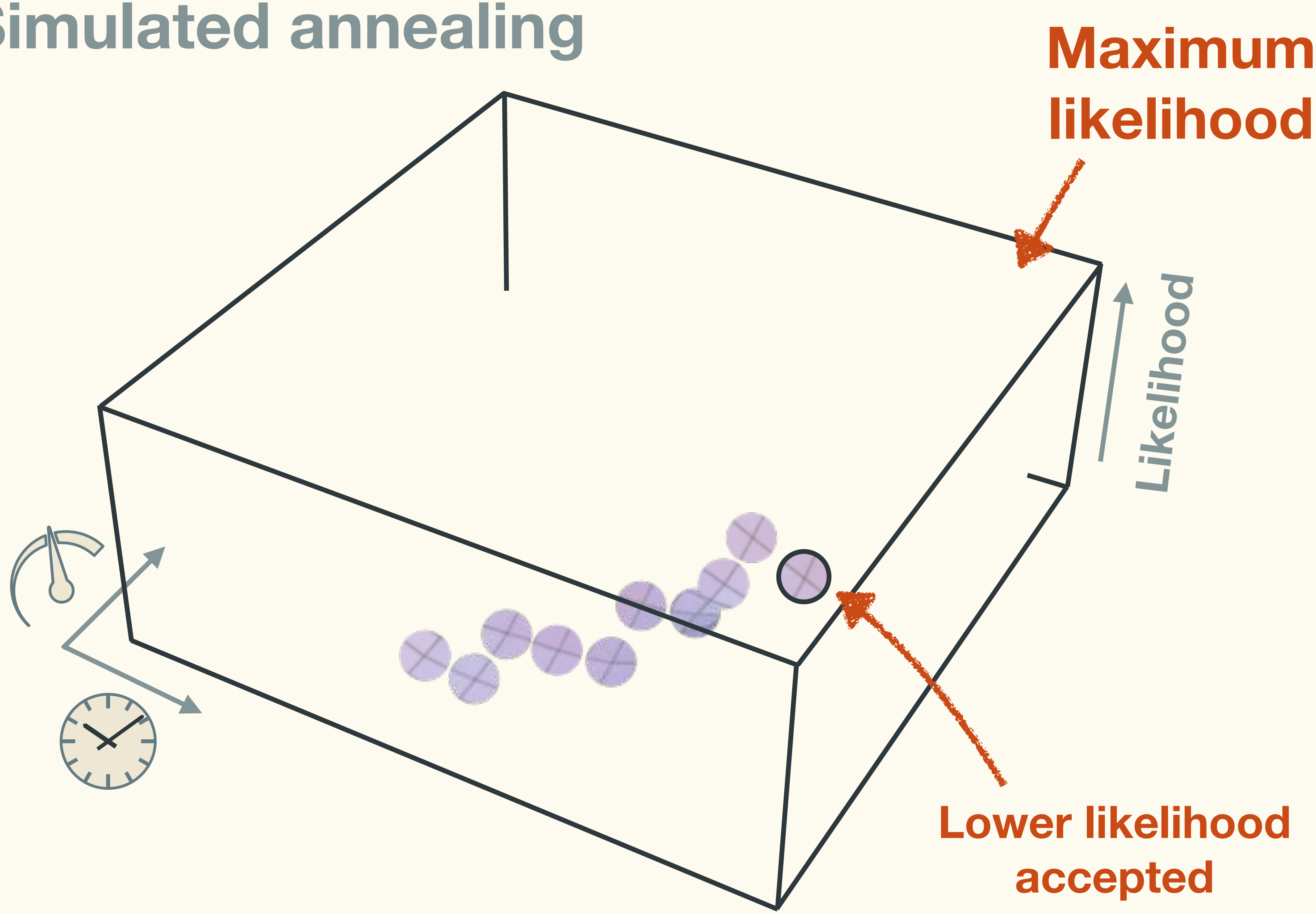
Heuristic search

Simulated annealing



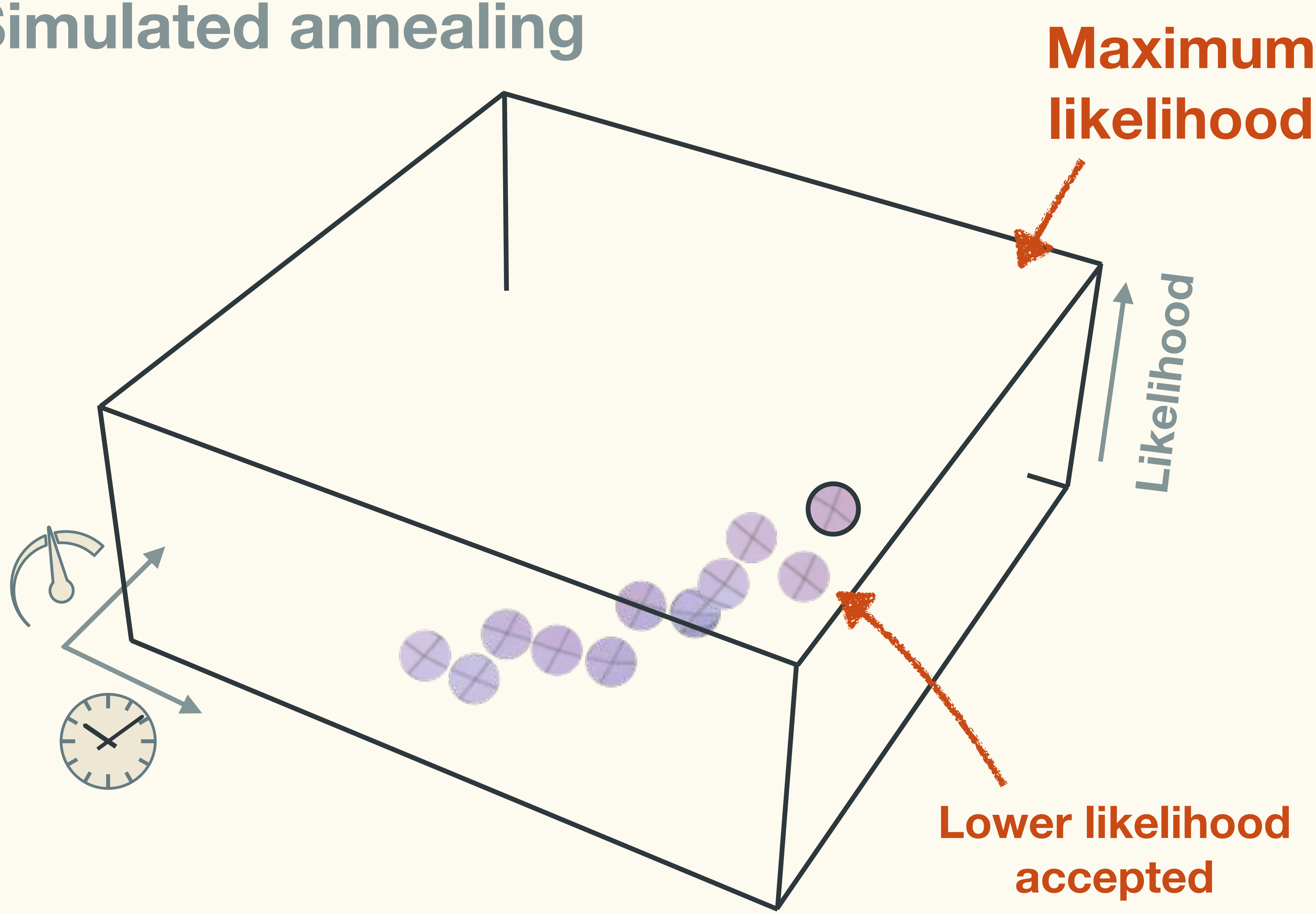
Heuristic search

Simulated annealing



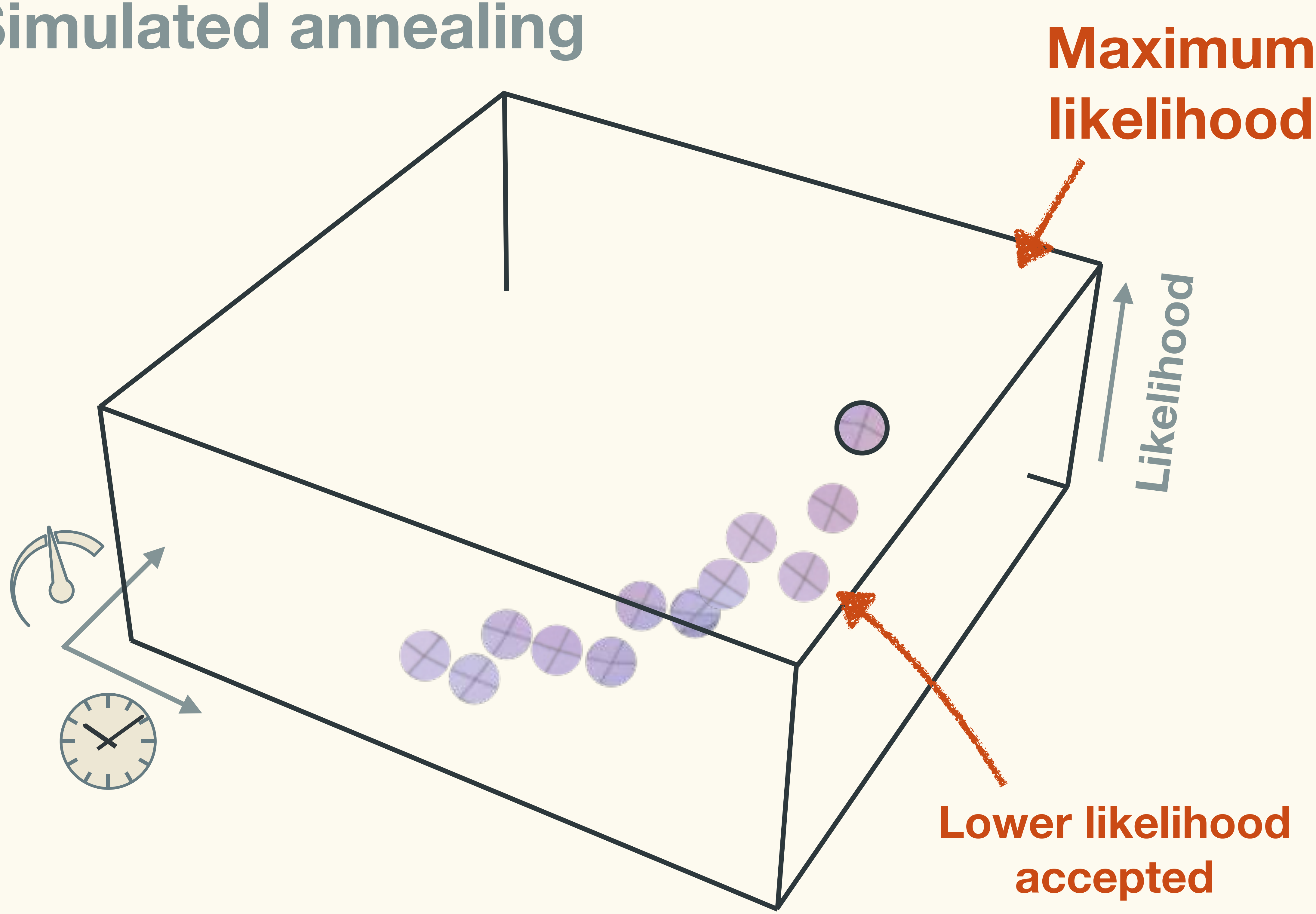
Heuristic search

Simulated annealing



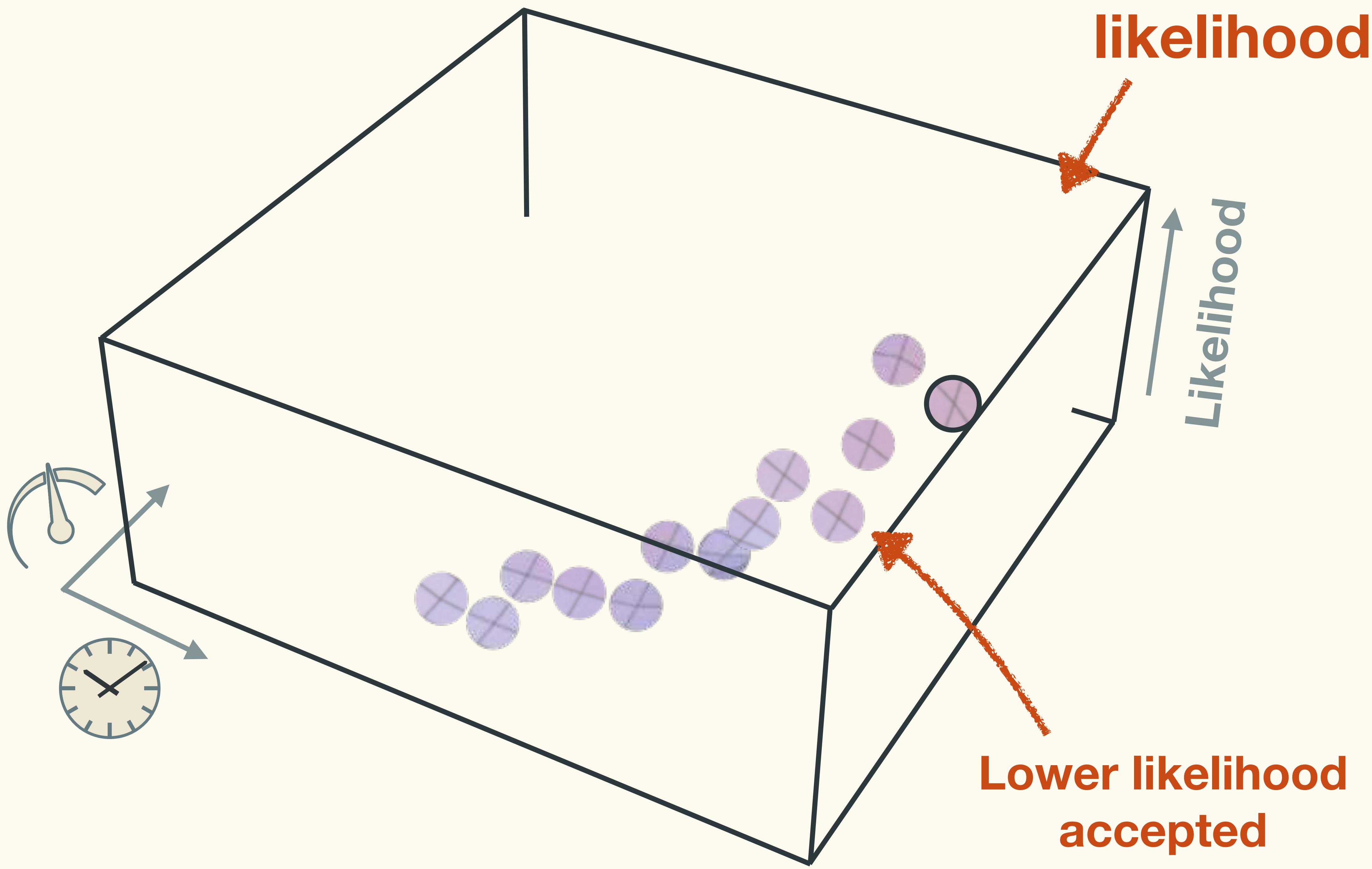
Heuristic search

Simulated annealing



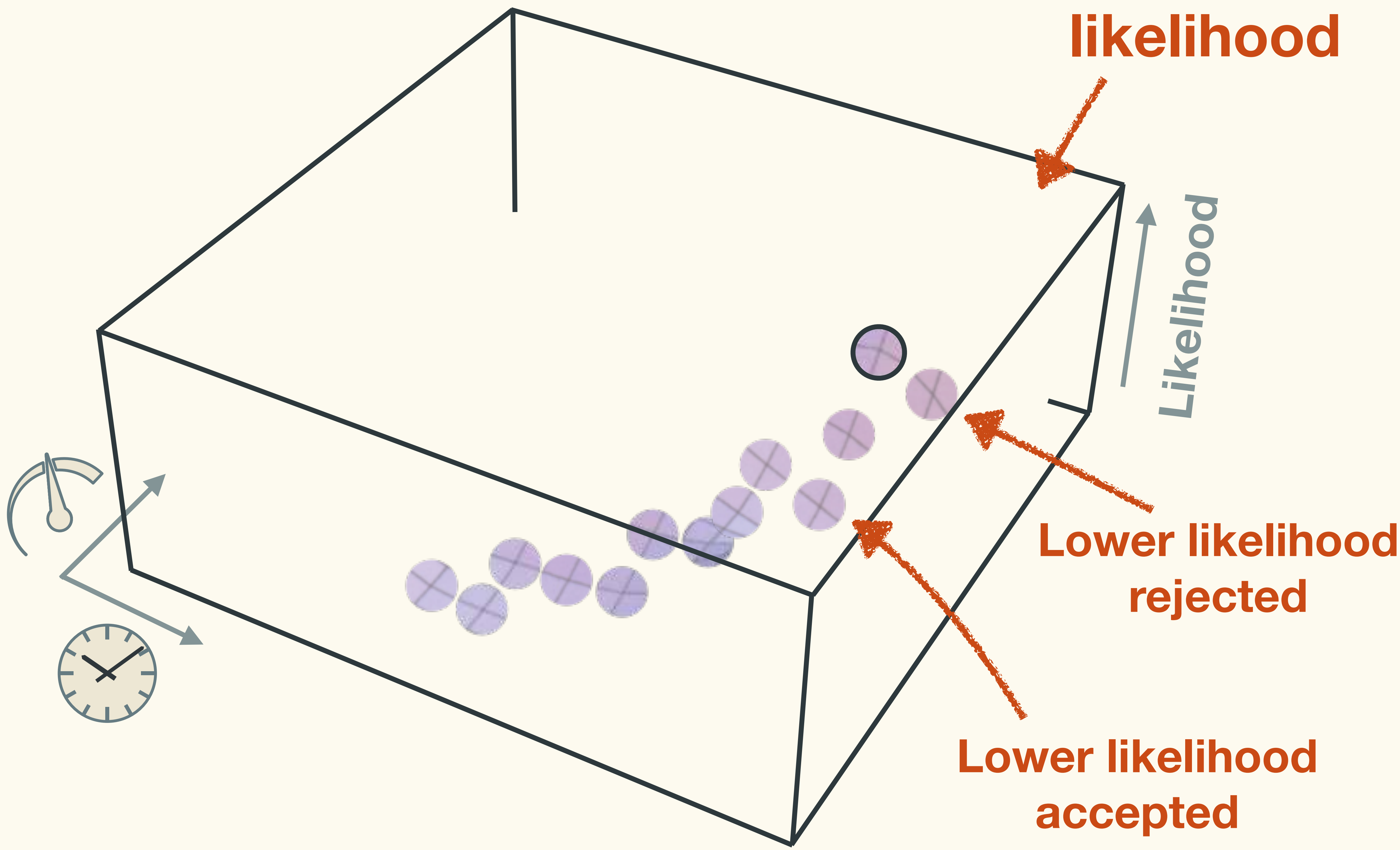
Heuristic search

Simulated annealing



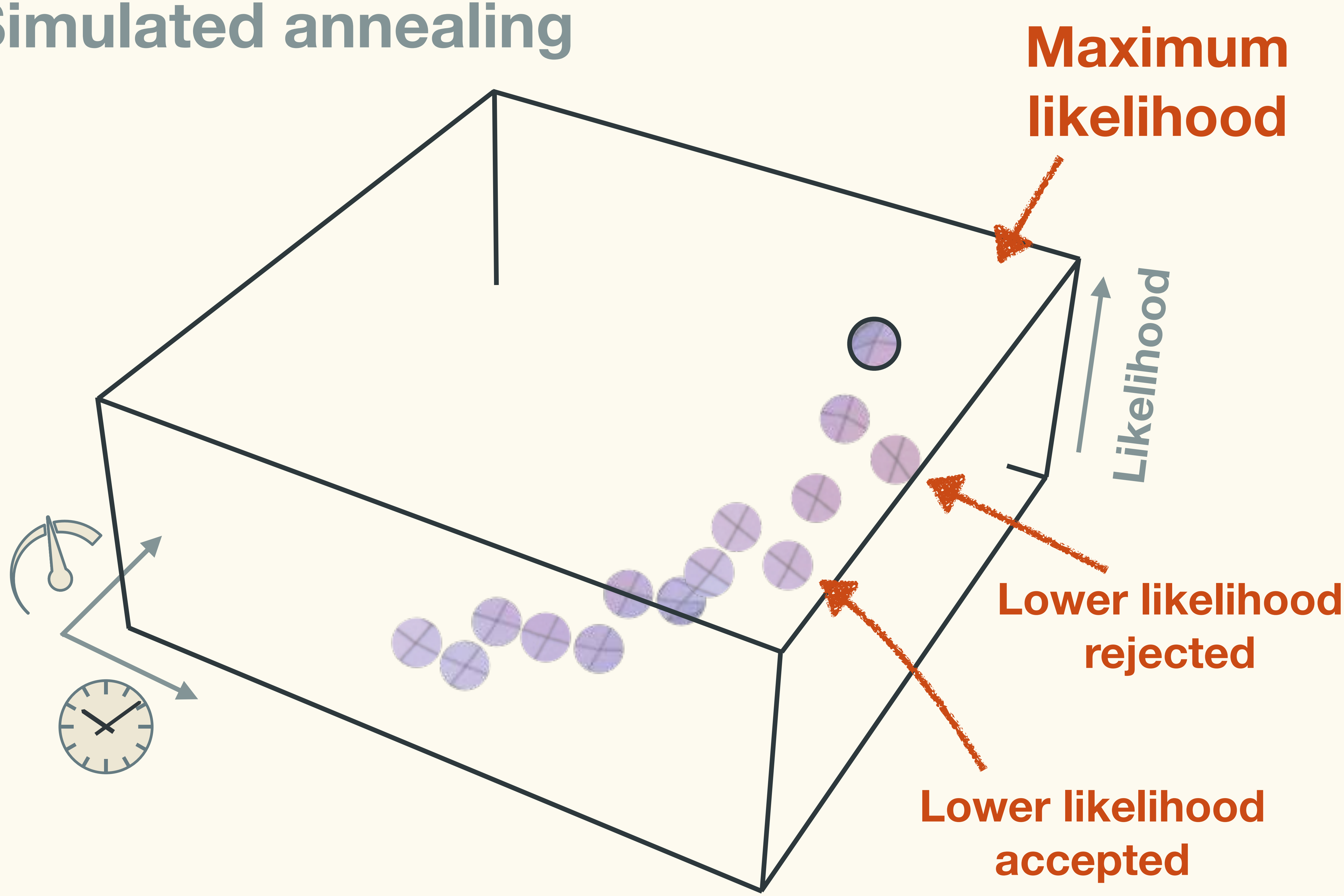
Heuristic search

Simulated annealing



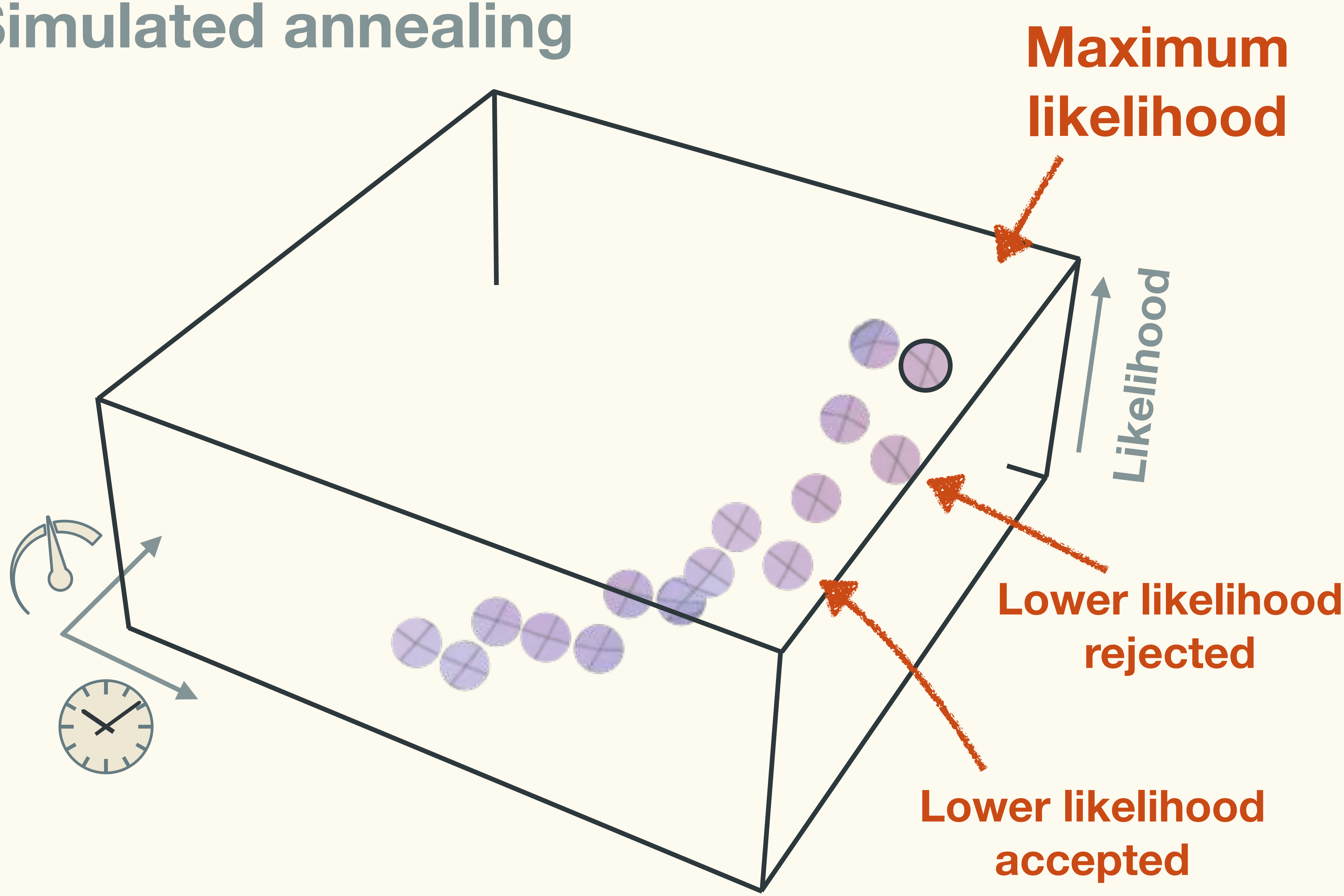
Heuristic search

Simulated annealing



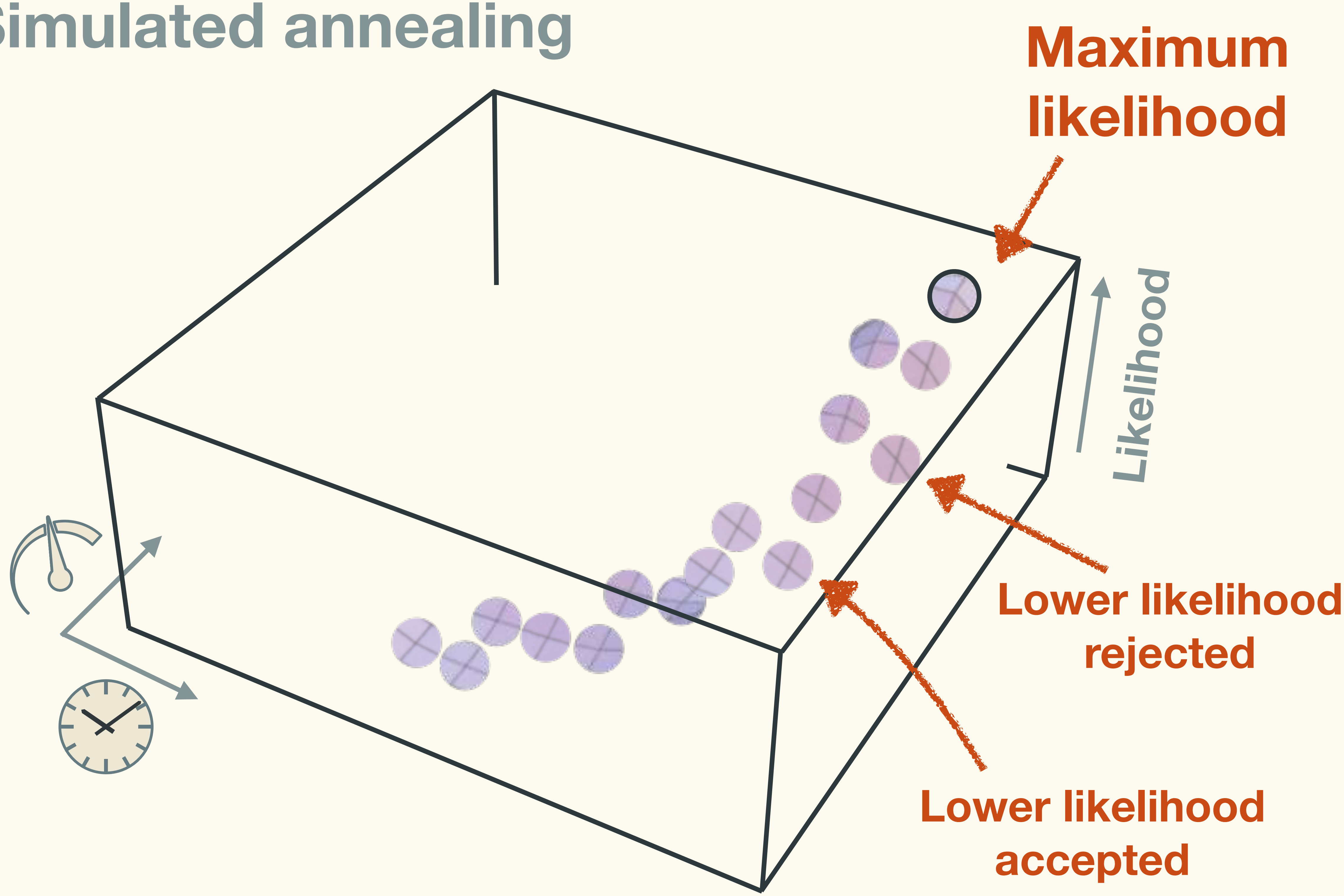
Heuristic search

Simulated annealing



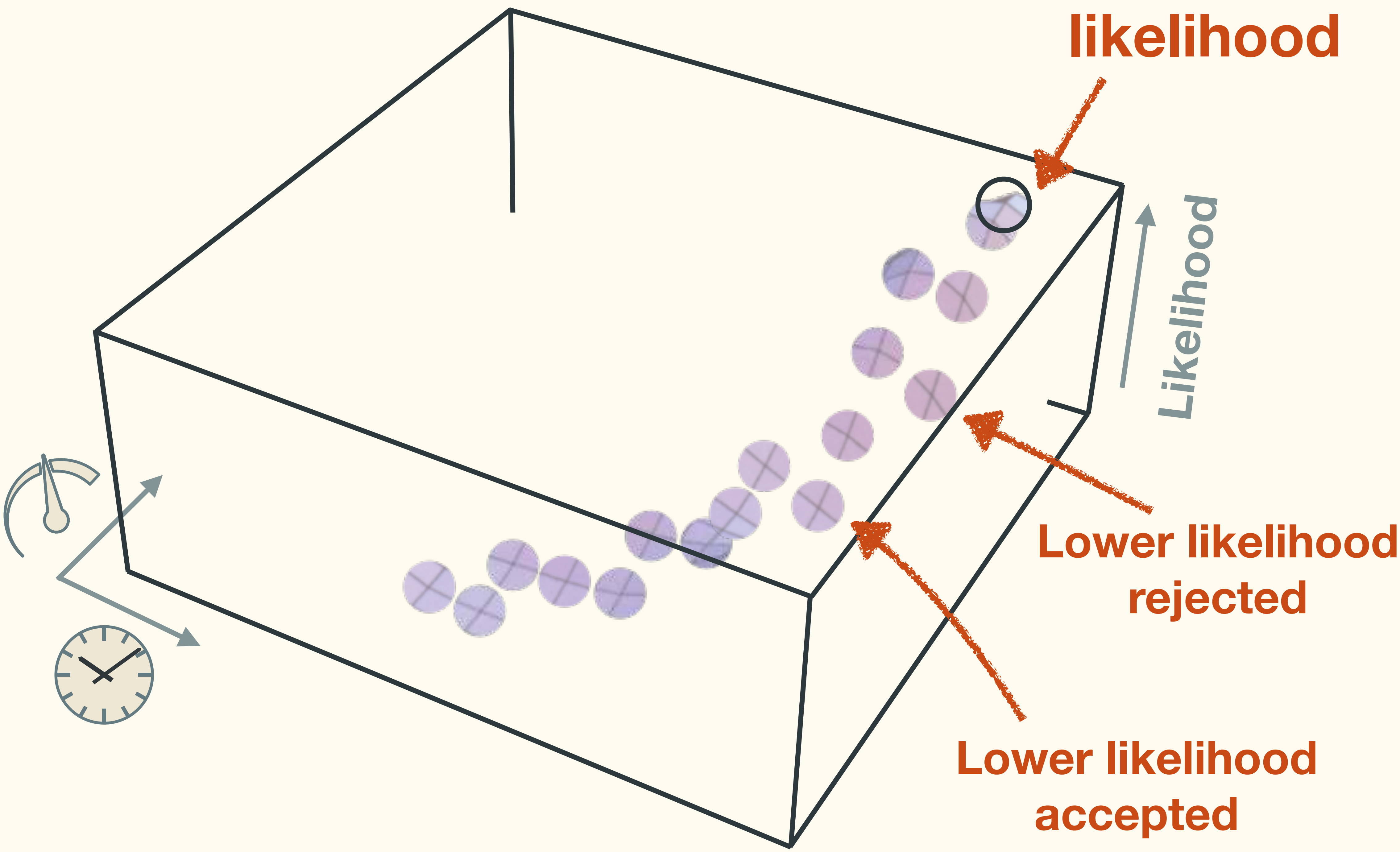
Heuristic search

Simulated annealing



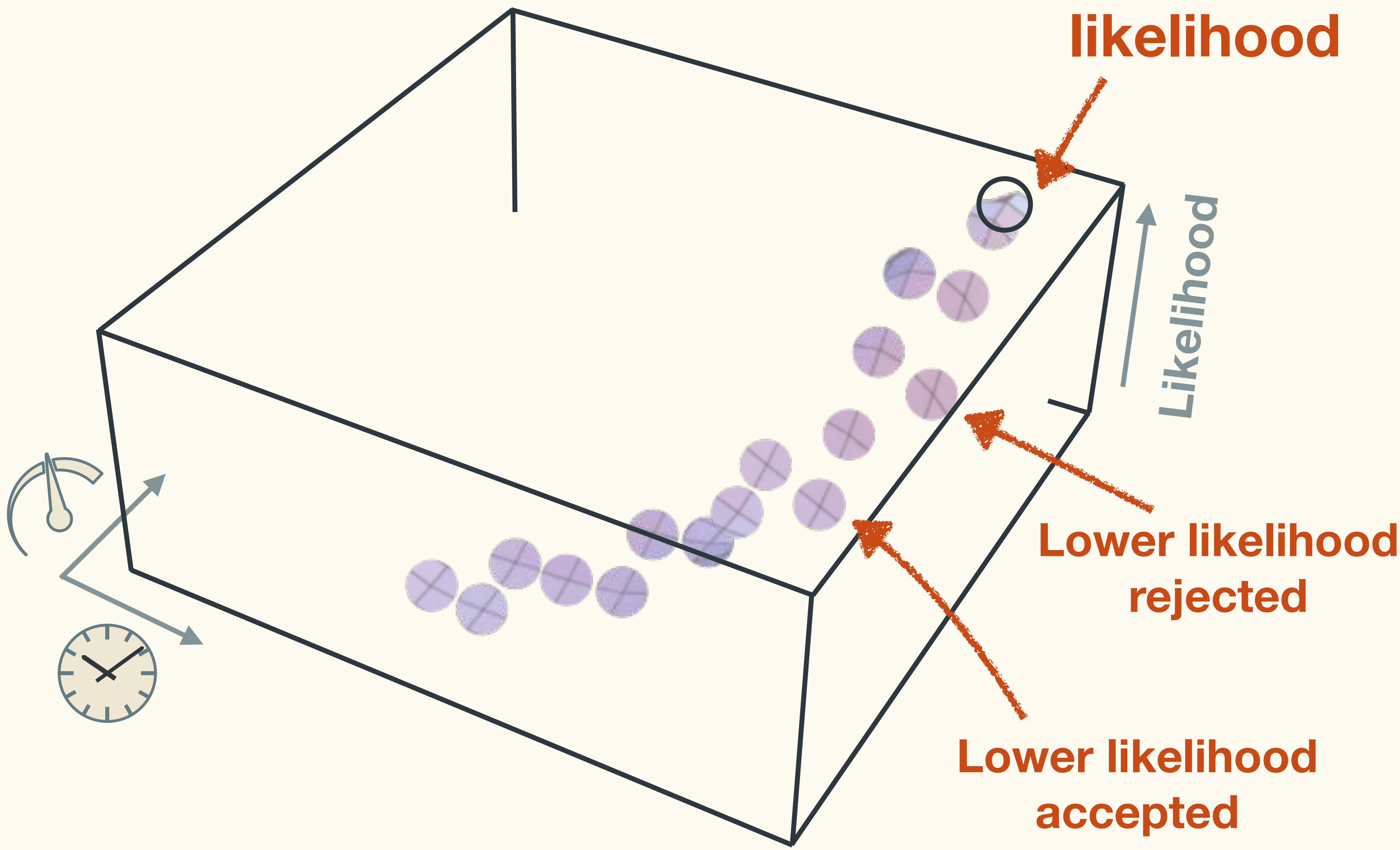
Heuristic search

Simulated annealing



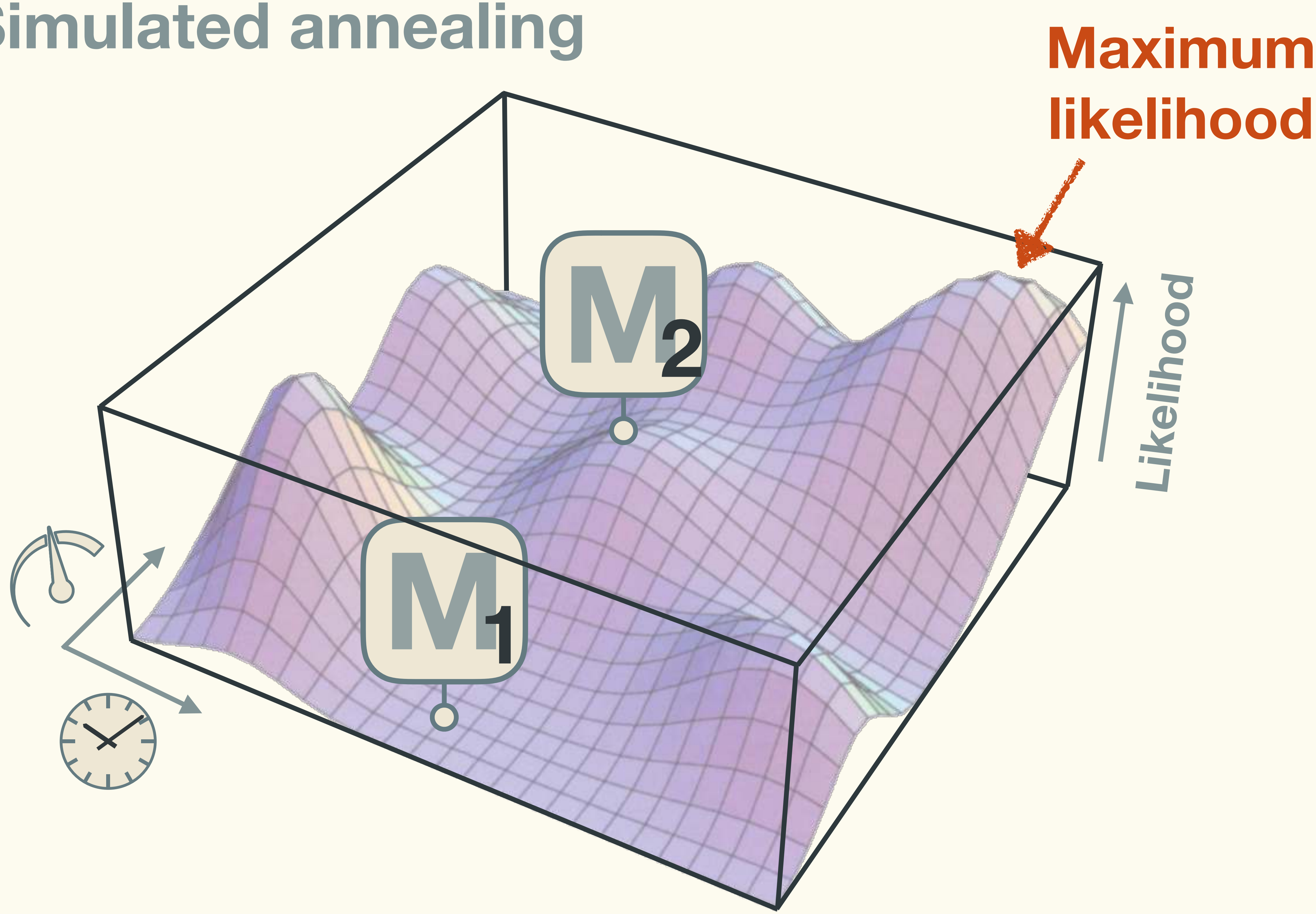
Heuristic search

Simulated annealing



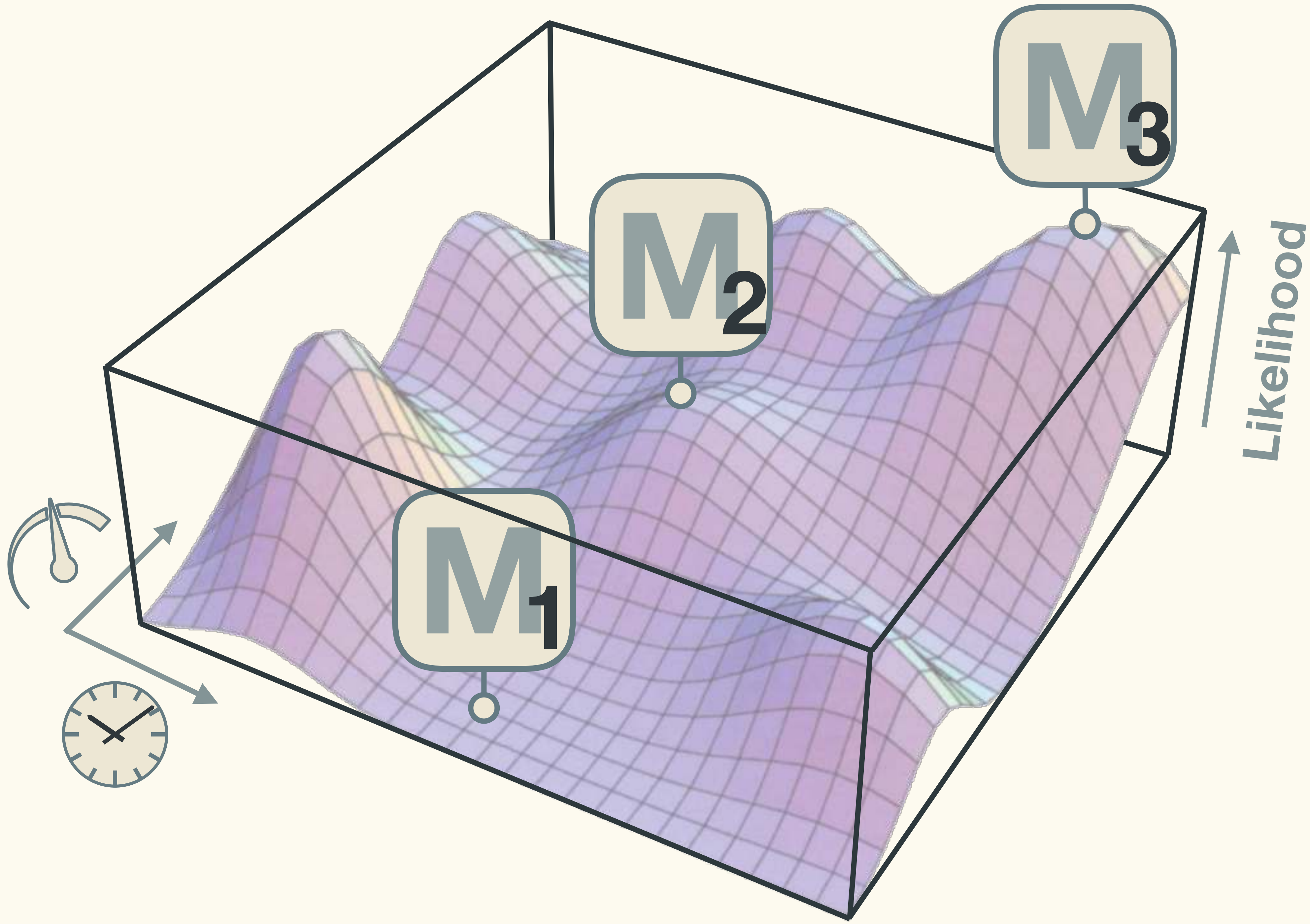
Heuristic search

Simulated annealing



Heuristic search

Simulated annealing



Likelihood

$$\log\left(L\left(\boxed{M_1} \;\middle|\; \begin{matrix}\text{ACTTG}\\ \text{ACTGG}\end{matrix}\right)\right) = -13.4$$

$$\log\left(L\left(\boxed{M_2} \;\middle|\; \begin{matrix}\text{ACTTG}\\ \text{ACTGG}\end{matrix}\right)\right) = -10.3$$

Likelihood

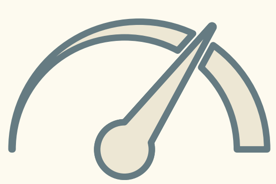
$$\log\left(L\left(\boxed{M_1} \mid \begin{matrix}\text{ACTTG}\\ \text{ACTGG}\end{matrix}\right)\right) = -13.4$$

$$\log\left(L\left(\boxed{M_2} \mid \begin{matrix}\text{ACTTG}\\ \text{ACTGG}\end{matrix}\right)\right) = -10.3$$

$$\log\left(L\left(\boxed{M_3} \mid \begin{matrix}\text{ACTTG}\\ \text{ACTGG}\end{matrix}\right)\right) = -5.4$$

Likelihood

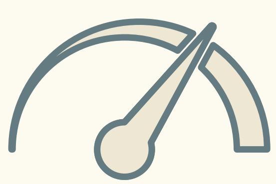
$$\log\left(L\left(\boxed{M_3} \,\middle|\, \begin{matrix}\text{ACTTG}\\ \text{ACTGG}\end{matrix}\right)\right) = -5.4$$

 ,  = 0.2 ,  = 0.01

Likelihood

$$\log\left(L \left(\boxed{M}_3 \;\middle|\; \begin{matrix} \mathrm{ACTTG} \\ \mathrm{ACTGG} \end{matrix} \right) \right) = -5.4$$

,



$= 0.2$,

$= 0.01$

Likelihood

$$\log\left(L\left(\boxed{M_3} \,\middle|\, \begin{matrix}\text{ACTTG}\\ \text{ACTGG}\end{matrix}\right)\right) = -5.4$$

Likelihood

$$\log\left(L\left(\boxed{M_3} \,\middle|\, \begin{matrix}\text{ACTTG}\\ \text{ACTGG}\end{matrix}\right)\right) = -5.4$$

Likelihood

$$\log\left(L\left(\boxed{M_3} \,\middle|\, \begin{matrix}\mathrm{ACTTG}\\ \mathrm{ACTGG}\end{matrix}\right)\right) = -5.4$$

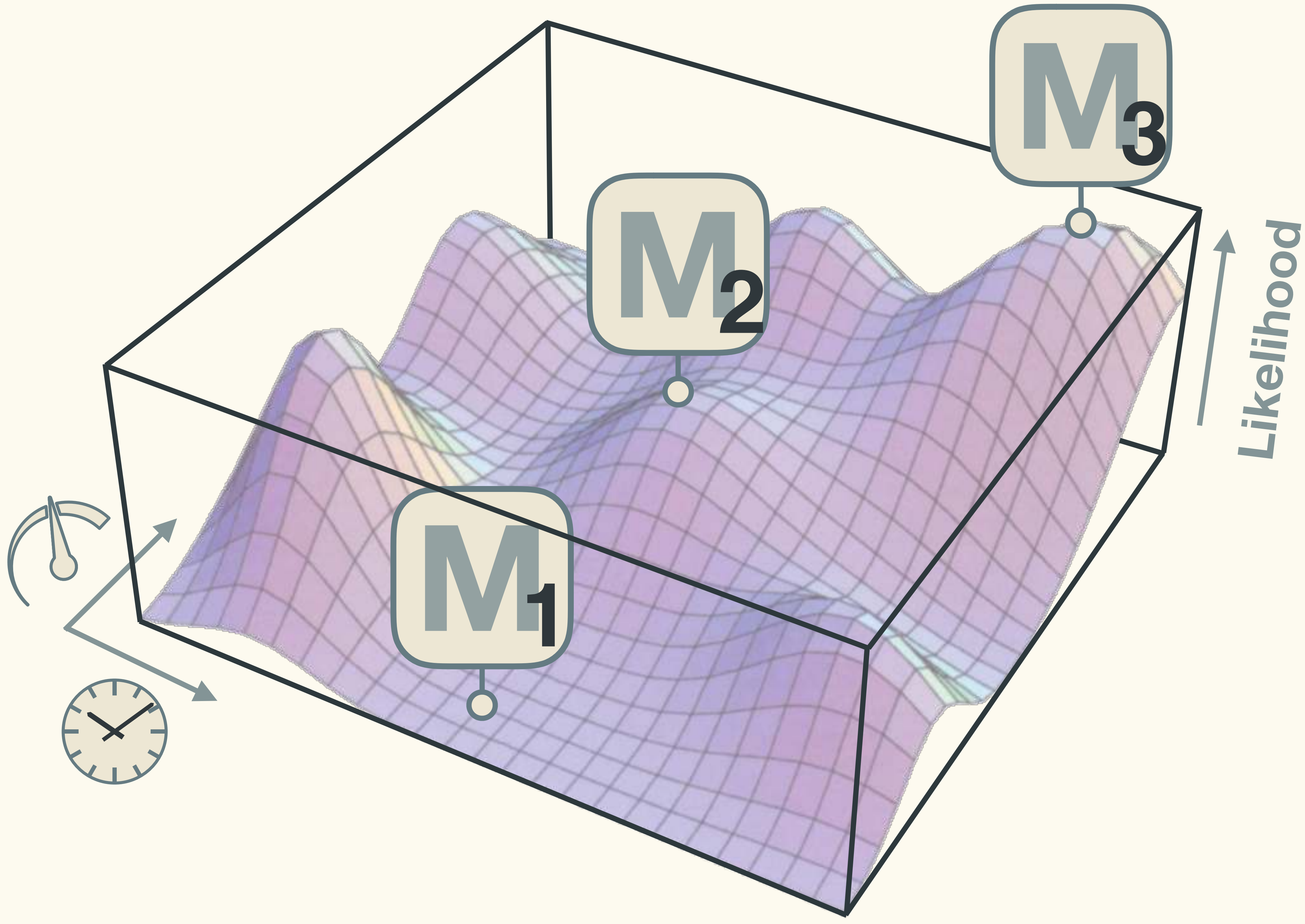
Likelihood

$$\log\left(L\left(\boxed{M_4} \,\middle|\, \begin{matrix}\text{ACTTG}\\ \text{ACTGG}\end{matrix}\right)\right) = -5.4$$

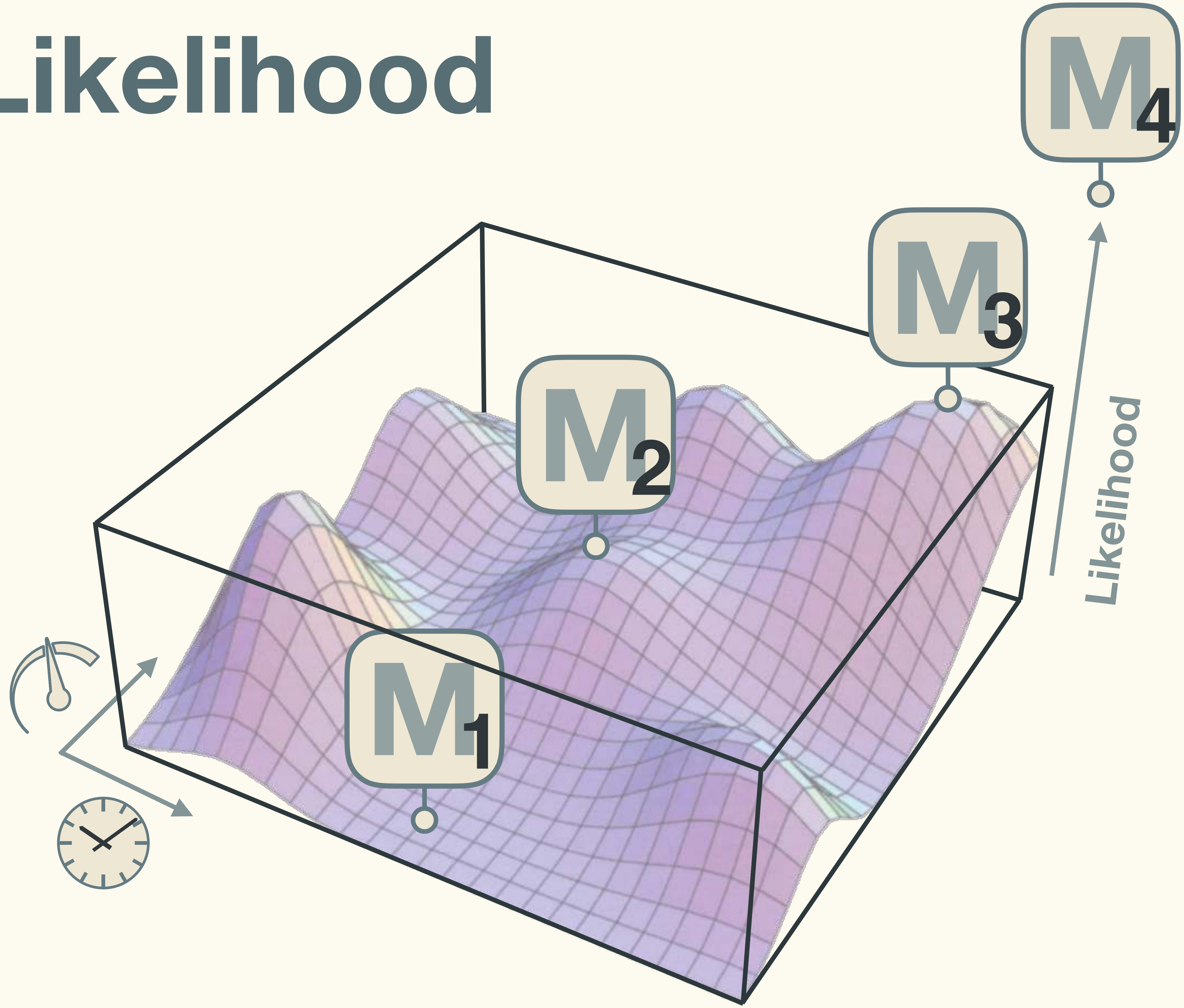
Likelihood

$$\log\left(L\left(\boxed{M}_4 \mid \begin{matrix}\text{ACTTG}\\ \text{ACTGG}\end{matrix}\right)\right) = -5.0$$

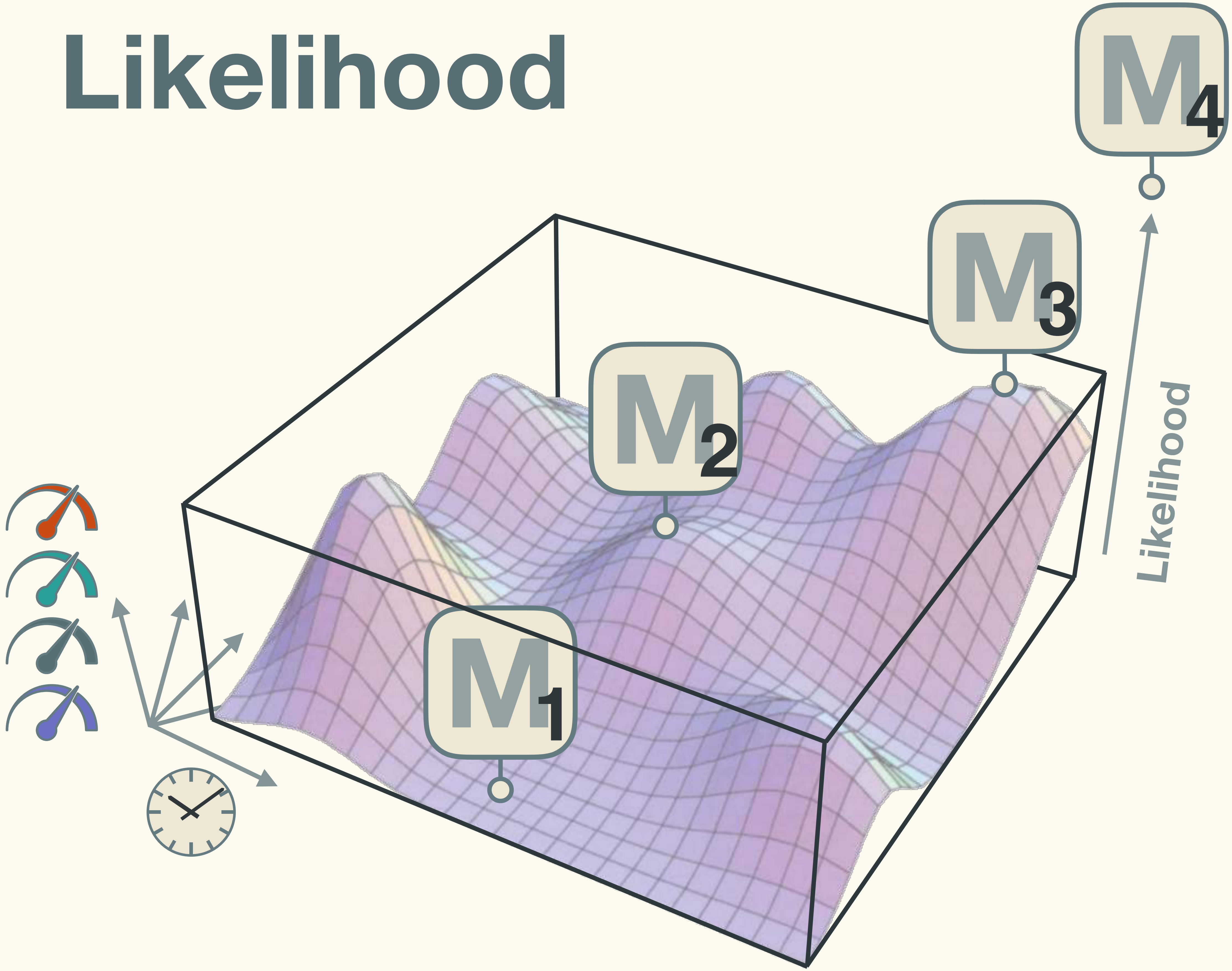
Likelihood



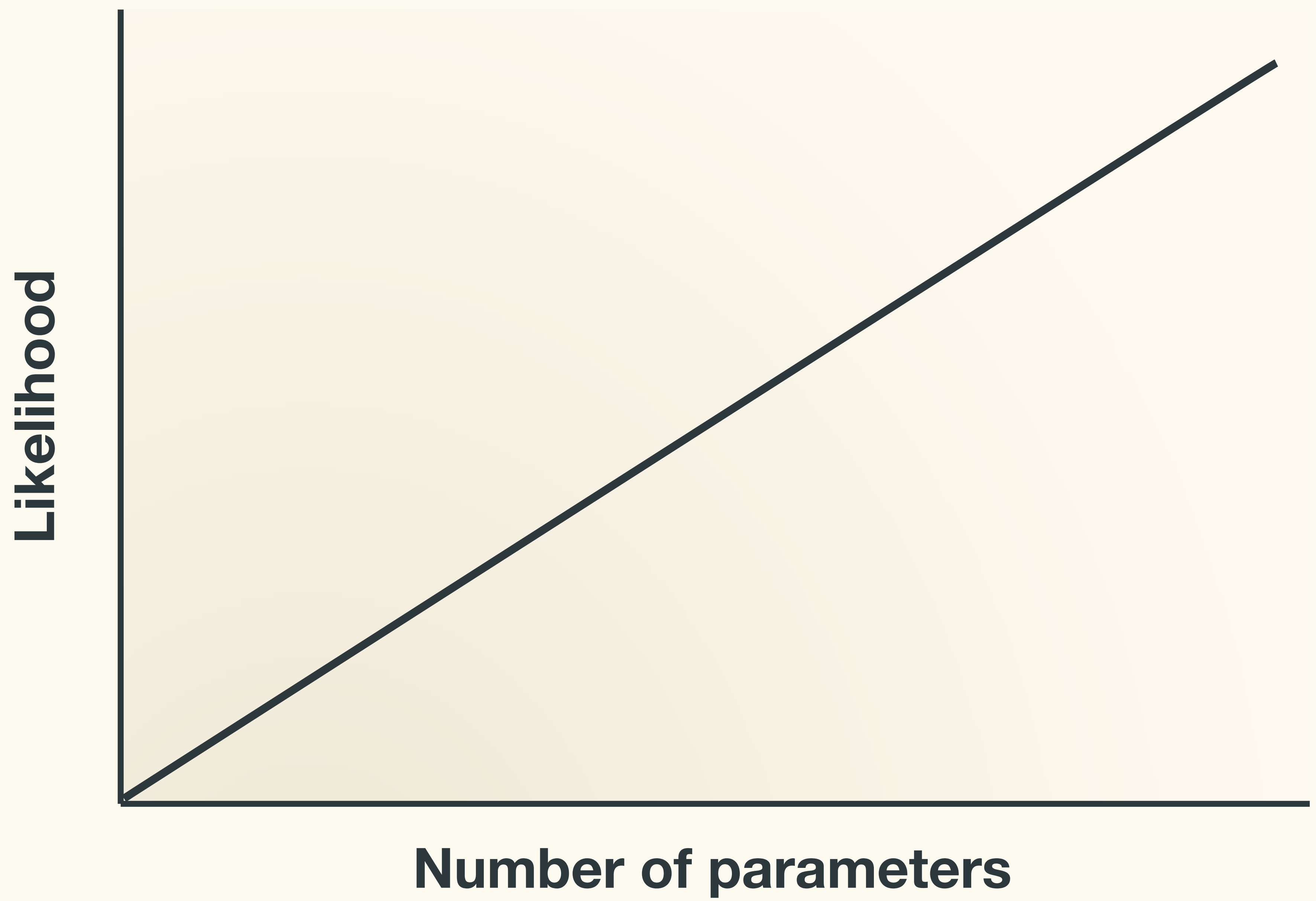
Likelihood



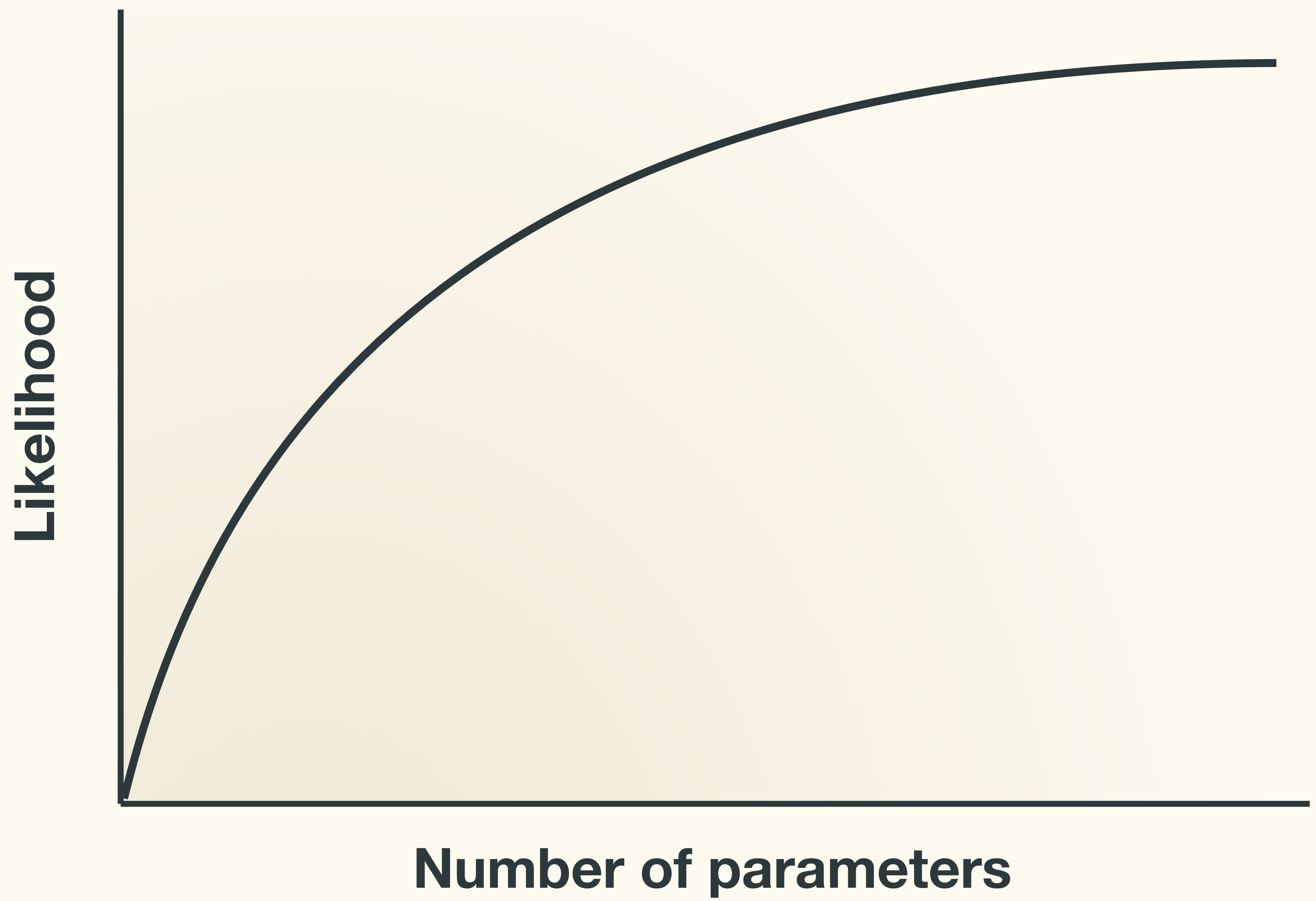
Likelihood



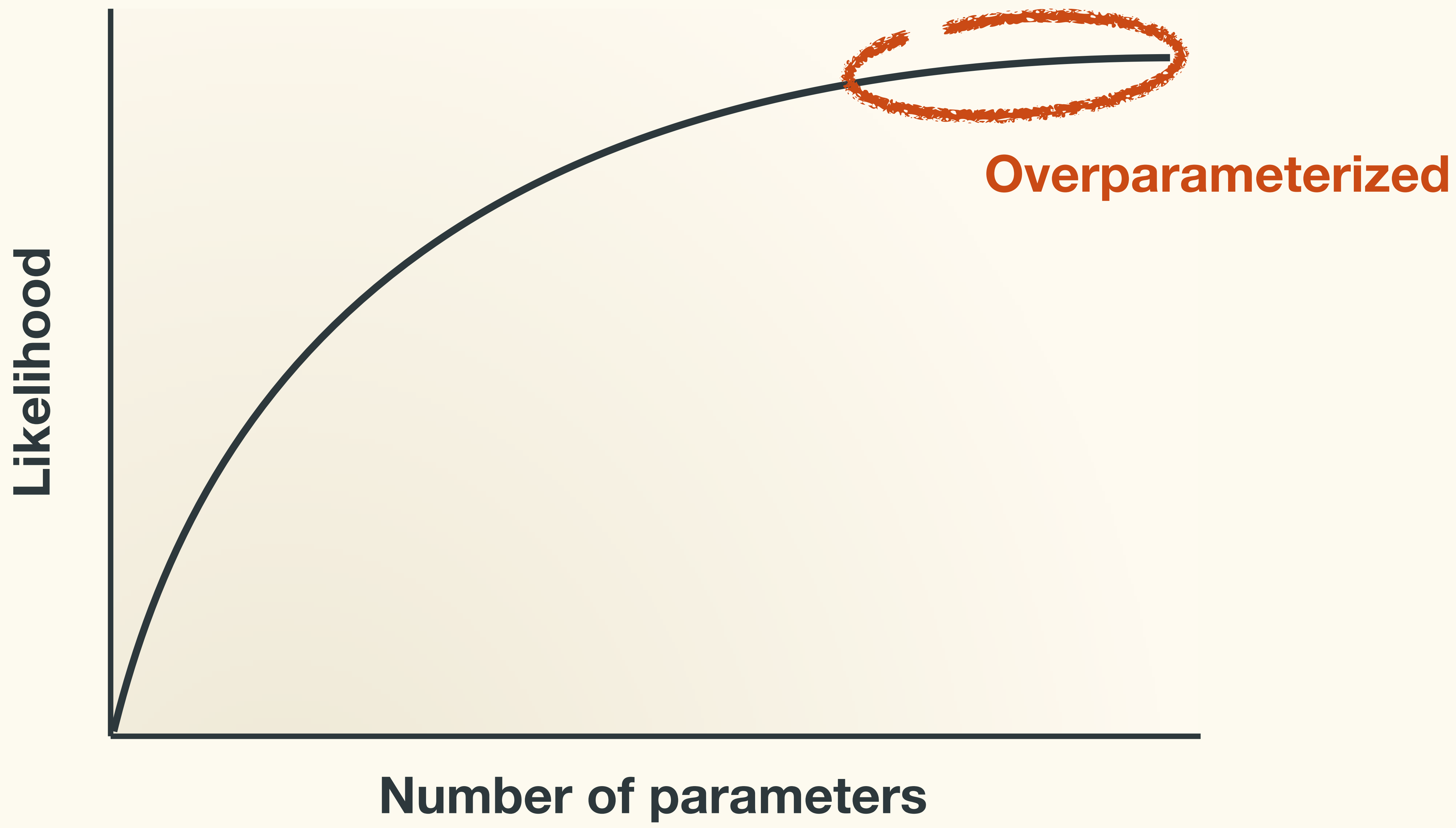
Likelihood



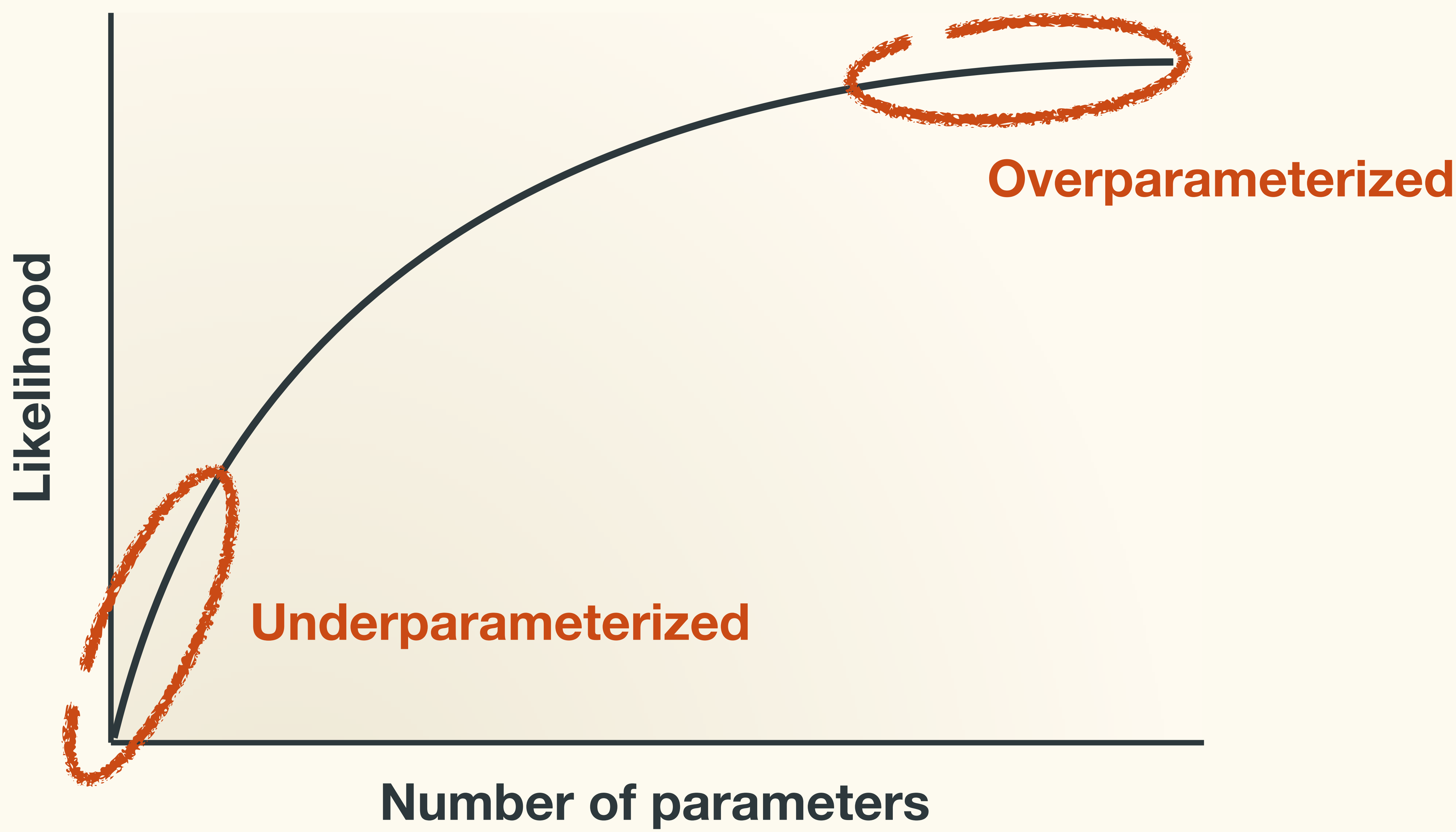
Likelihood



Likelihood



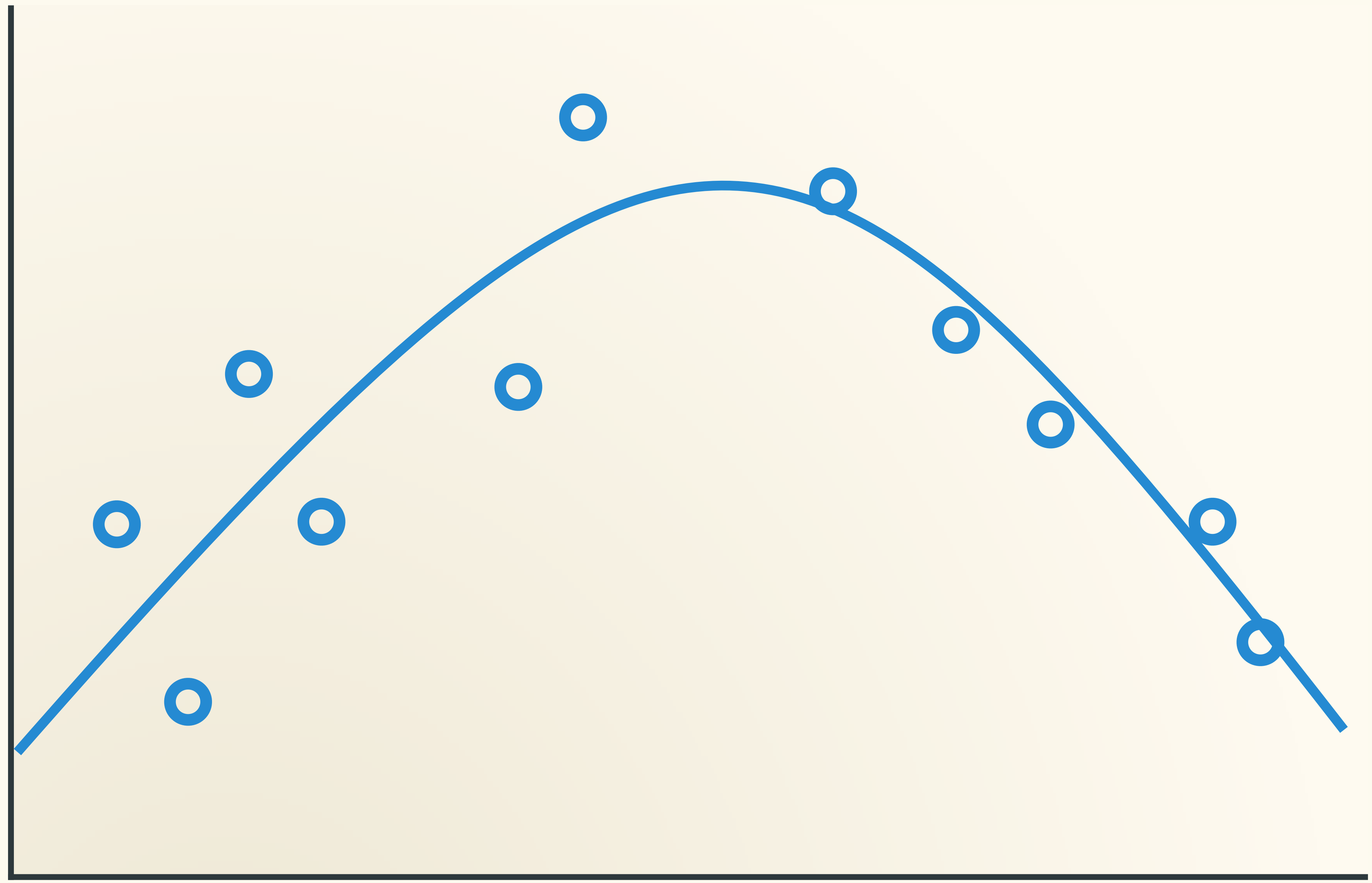
Likelihood



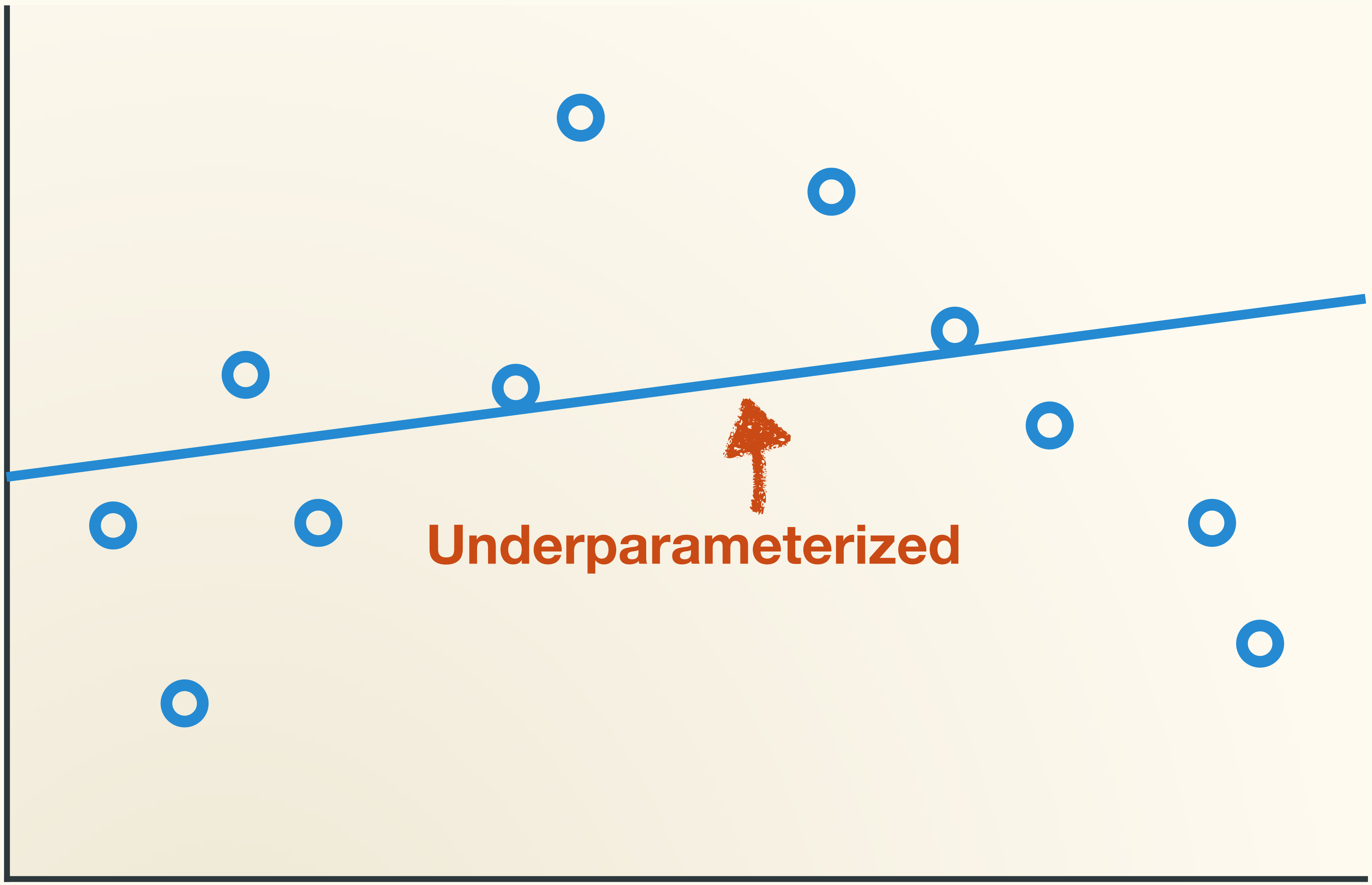
Overparameterized
Likelihood
Underparameterized
Number of parameters

Parameterization

Parameterization



Parameterization



Parameterization

Likelihood ratio test

$$LRT = 2\log\left(\frac{L_{(\text{Complex model})}}{L_{(\text{Simple model})}}\right)$$

Likelihood ratio test

$$LRT = 2\log\left(\frac{L_{(\text{Complex model})}}{L_{(\text{Simple model})}}\right)$$

Likelihood ratio test

$$LRT = 2 \log \left(\frac{L(\boxed{M_4})}{L(\boxed{M_3})} \right)$$

Likelihood ratio test

$$LRT = 2 \log \left(\frac{L(\boxed{M_4})}{L(\boxed{M_3})} \right)$$

Compared to
Chi-square score

Akaike information criterion

$$\mathrm{AIC} = 2k - 2\left(\log(L)\right)$$

Akaike information criterion

$$\text{AIC} = 2k - 2\left(\log(L)\right)$$

Number of parameters

Akaike information criterion

$$\mathrm{AIC}(\boxed{M_4}) = 2k - 2\left(\log(L \mid \boxed{M_4})\right)$$

Akaike information criterion

$$\mathrm{AIC}\left(\boxed{M_4}\right) = 2k - 2\left(\log\left(L \mid \boxed{M_4}\right)\right)$$

$$\mathrm{AIC}\left(\boxed{M_3}\right) = 2k - 2\left(\log\left(L \mid \boxed{M_3}\right)\right)$$

Akaike information criterion

$$\mathrm{AIC}(\boxed{M_4}) = 2k - 2\left(\log(L \mid \boxed{M_4})\right)$$

$$\mathrm{AIC}(\boxed{M_3}) = 2k - 2\left(\log(L \mid \boxed{M_3})\right)$$

Akaike information criterion

$$\delta AIC = AIC(\boxed{M_4}) - AIC(\boxed{M_3})$$

Akaike information criterion

$$\delta AIC = AIC(\boxed{M_4}) - AIC(\boxed{M_3})$$

Bayesian inference

Bayesian inference

- **A T-rex outside the door?**

Bayesian inference

- A T-rex outside the door?
- Somebody pretending to be a T-rex?

Likelihood

- **A T-rex outside the door?**
- **Somebody pretending to be a T-rex?**

Likelihood

- A T-rex outside the door?
- Somebody pretending to be a T-rex?

Likelihood

- A T-rex outside the door?
- Somebody pretending to be a T-rex?

Likelihood

L( | ROAR) = *P*(ROAR | )

L( | ROAR) = *P*(ROAR | )

Likelihood

L( | ROAR) = P(ROAR | )

L( | ROAR) = P(ROAR | )

Likelihood

$P($ROAR $|$  $) \approx 1$

$P($ROAR $|$  $) \approx 1$

Likelihood

L ( | ROAR) ≈ 1

L ( | ROAR) ≈ 1

Likelihood

L ( | ROAR)

Likelihood

L ( | ROAR)

P (ROAR | )

Likelihood

L ( | ROAR)

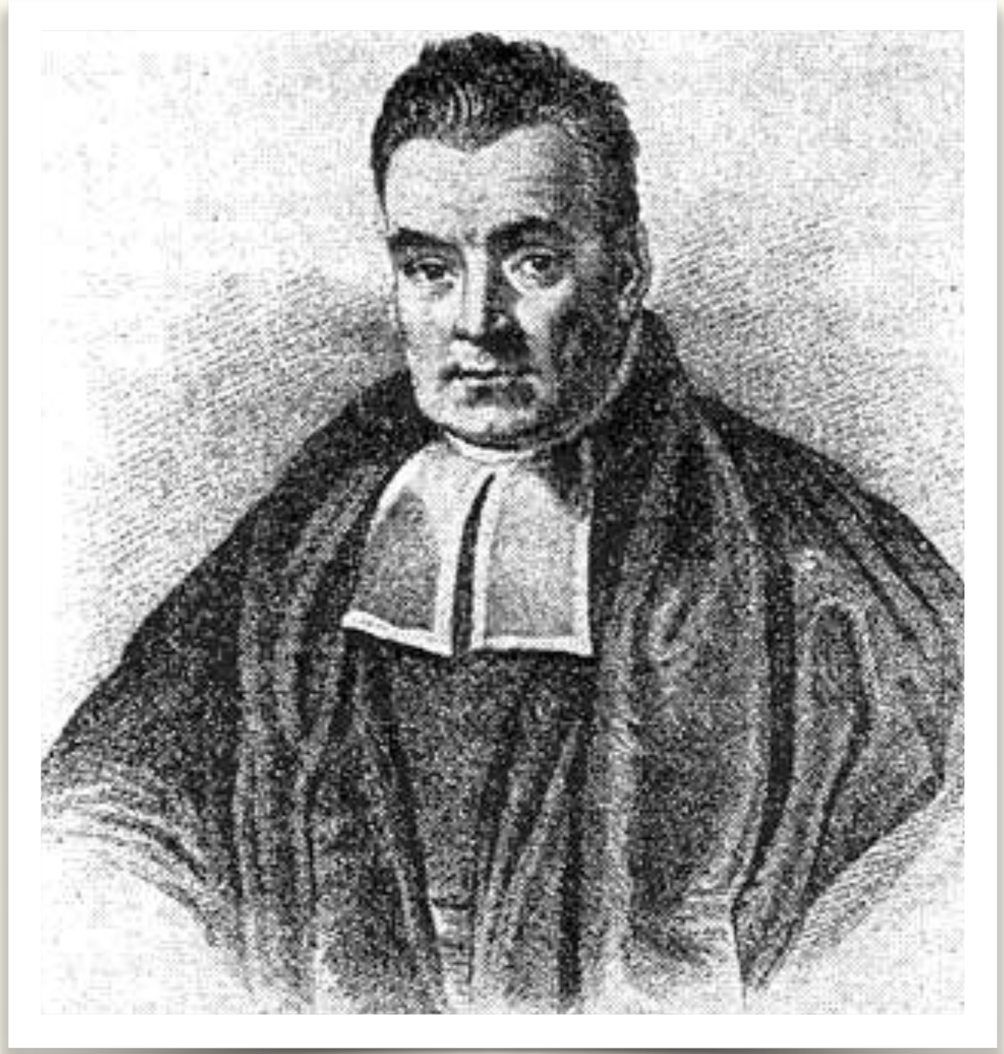
P (ROAR | )

P ( | ROAR)

Probability

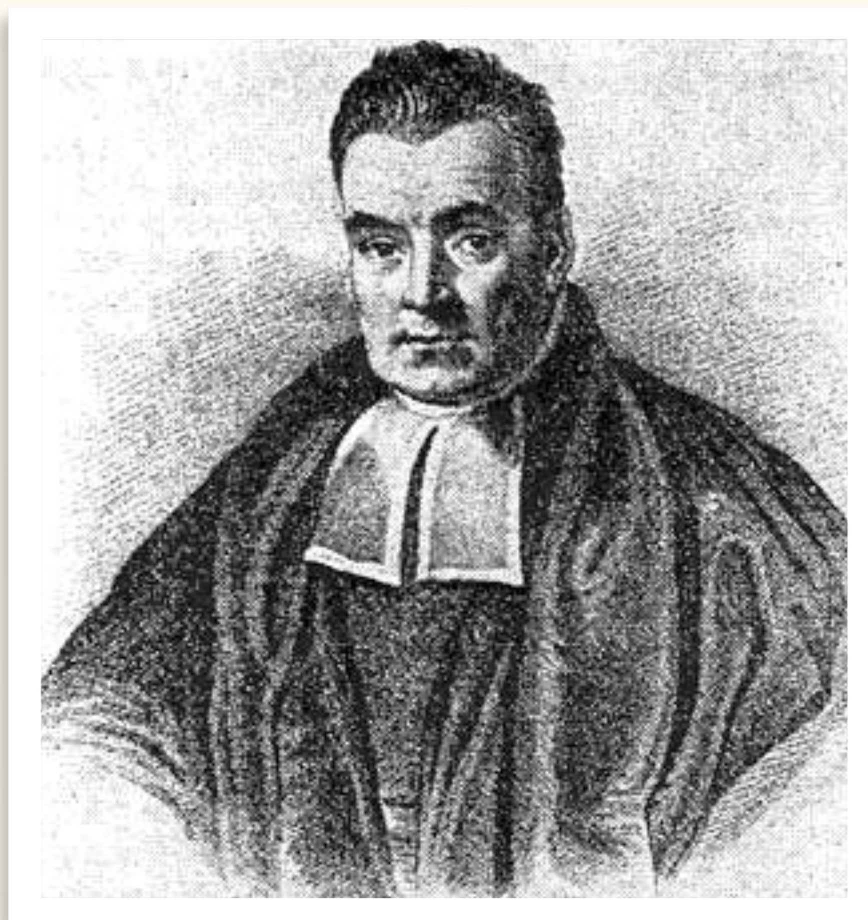
$P($  $|$ ROAR $) = ?$

Bayesian inference

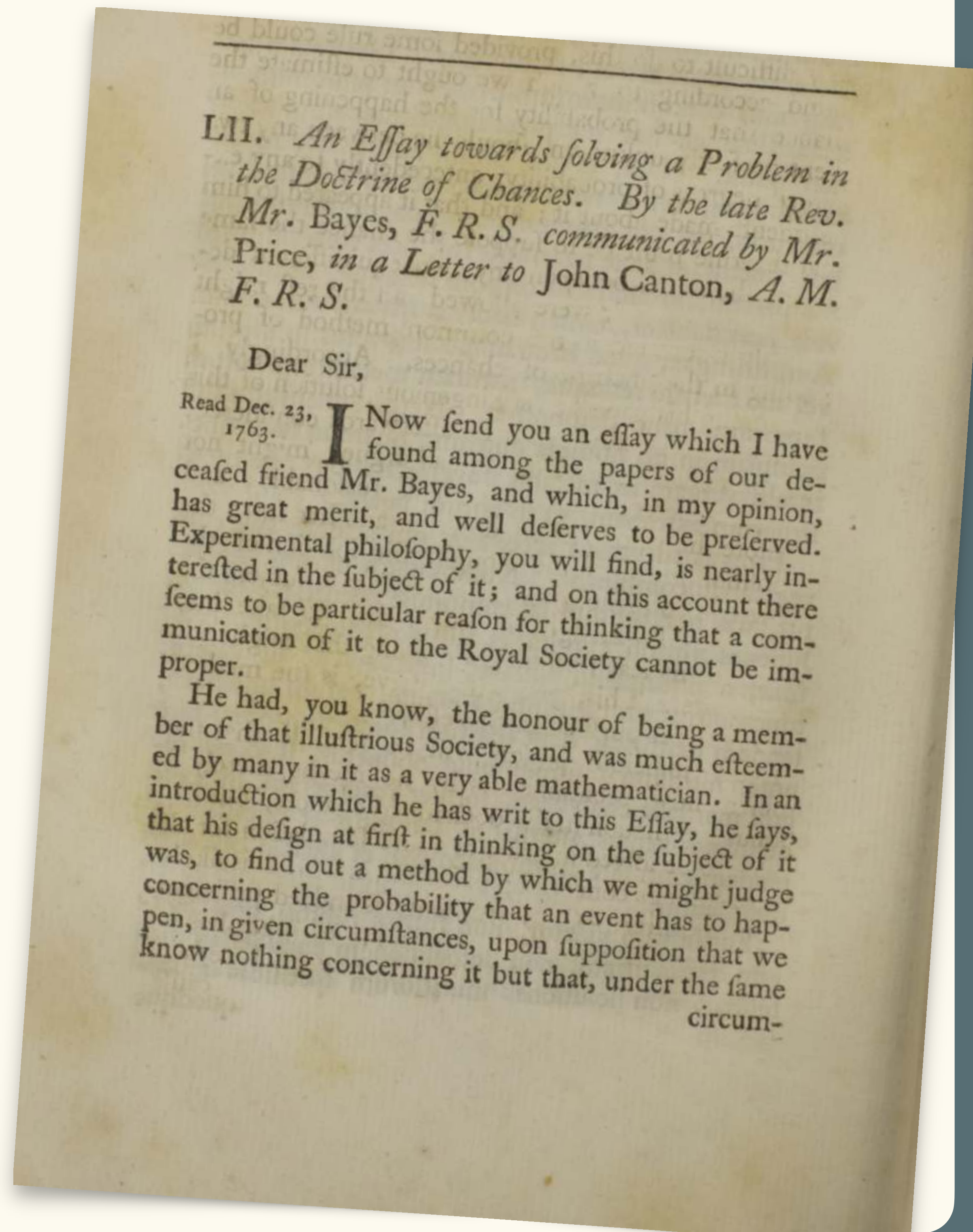


Thomas Bayes
(1701–1761)

Bayesian inference



Thomas Bayes
(1701–1761)



LII. *An Eſſay towards ſolving a Problem in the Doctrine of Chances. By the late Rev. Mr.* Bayes, *F. R. S. communicated by Mr.* Price, *in a Letter to* John Canton, *A. M. F. R. S.*

Dear Sir,

Read Dec. 23, 1763.

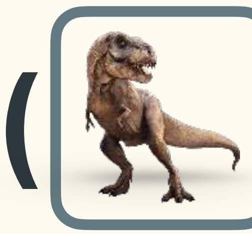
I Now ſend you an eſſay which I have found among the papers of our deceaſed friend Mr. Bayes, and which, in my opinion, has great merit, and well deſerves to be preſerved. Experimental philoſophy, you will find, is nearly intereſted in the ſubject of it; and on this account there ſeems to be particular reaſon for thinking that a communication of it to the Royal Society cannot be improper.

He had, you know, the honour of being a member of that illuſtrious Society, and was much eſteemed by many in it as a very able mathematician. In an introduction which he has writ to this Eſſay, he ſays, that his deſign at firſt in thinking on the ſubject of it was, to find out a method by which we might judge concerning the probability that an event has to happen, in given circumſtances, upon ſuppoſition that we know nothing concerning it but that, under the ſame circum-

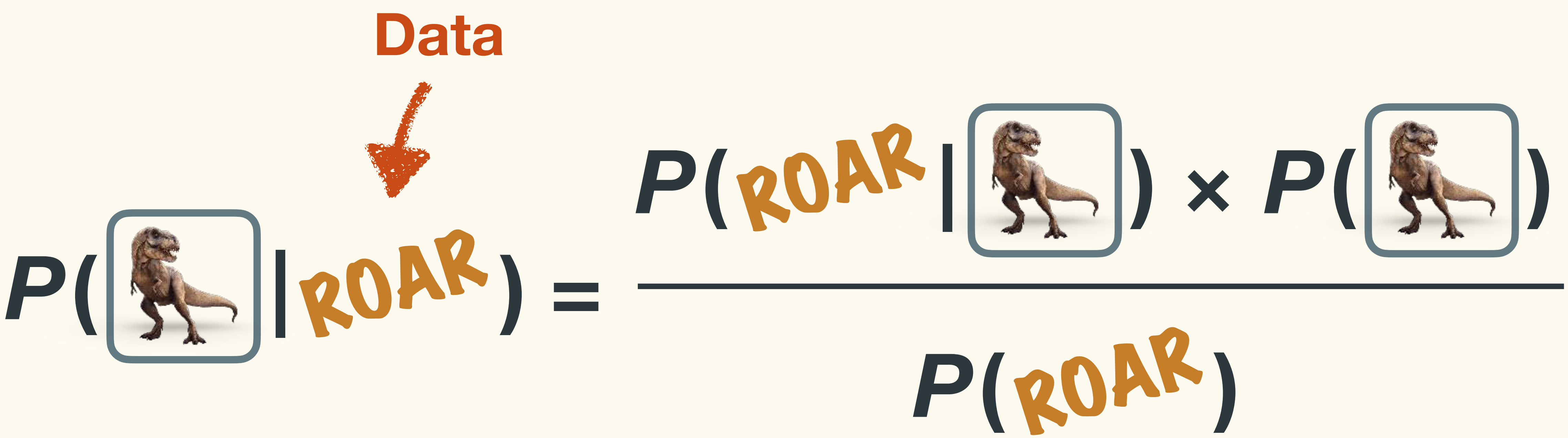
Price (1763) *Philosophical Transactions of the Royal Society* 53, 370–418

Bayes’ theorem

Model

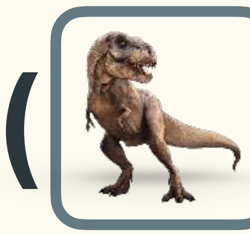
P( | ROAR) = *P*(ROAR | ) × *P*() / *P*(ROAR)

Bayes' theorem

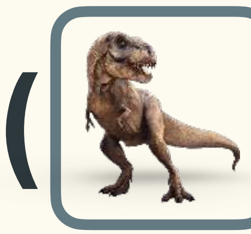


Bayes' theorem

Posterior probability

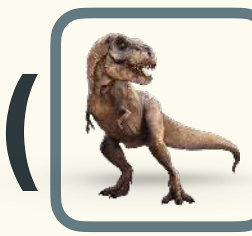
P( **| ROAR) =** **P(ROAR |** **) × P(****)** / **P(ROAR)**

Bayes' theorem

P( | ROAR) = P(ROAR | ) × P() / P(ROAR)

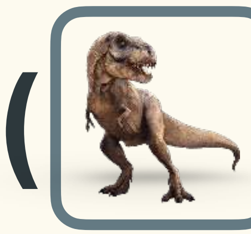
Bayes' theorem

Likelihood

P( | ROAR) = [L( | ROAR) × P()] / P(ROAR)

Bayes' theorem

Prior probability

P( | ROAR) = L( | ROAR) × P() / P(ROAR)

Bayes' theorem

P( | ROAR) = L( | ROAR) × P() / P(ROAR)

Constant

Bayes' theorem

P( | ROAR) =

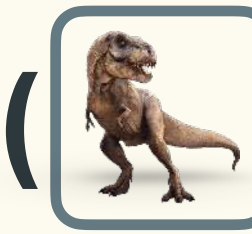
L( | ROAR) × P()

―――――――――――――――――

P(ROAR)

Bayes' theorem

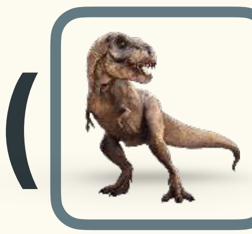
Posterior probability

P( | ROAR) = L( | ROAR) × P() / P(ROAR)

Bayes' theorem

Posterior probability **Likelihood**

P( | ROAR) =

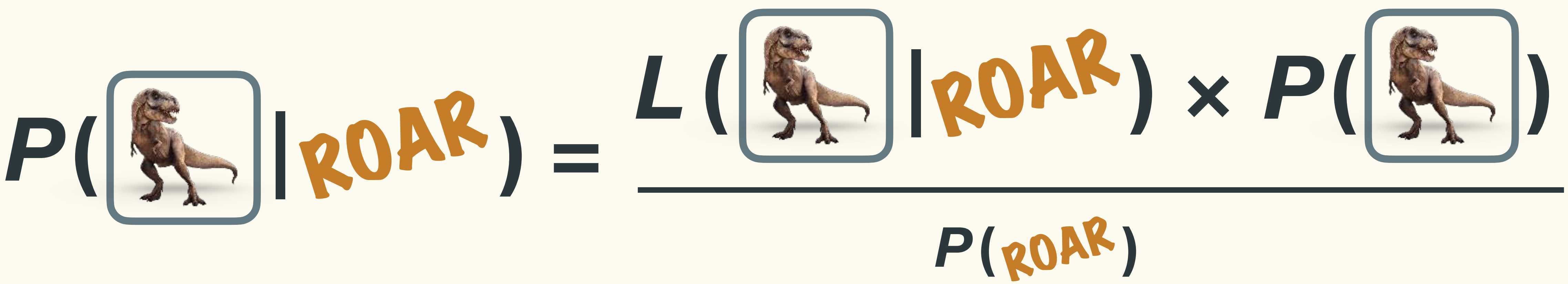
L( | ROAR) × P()

───────────────

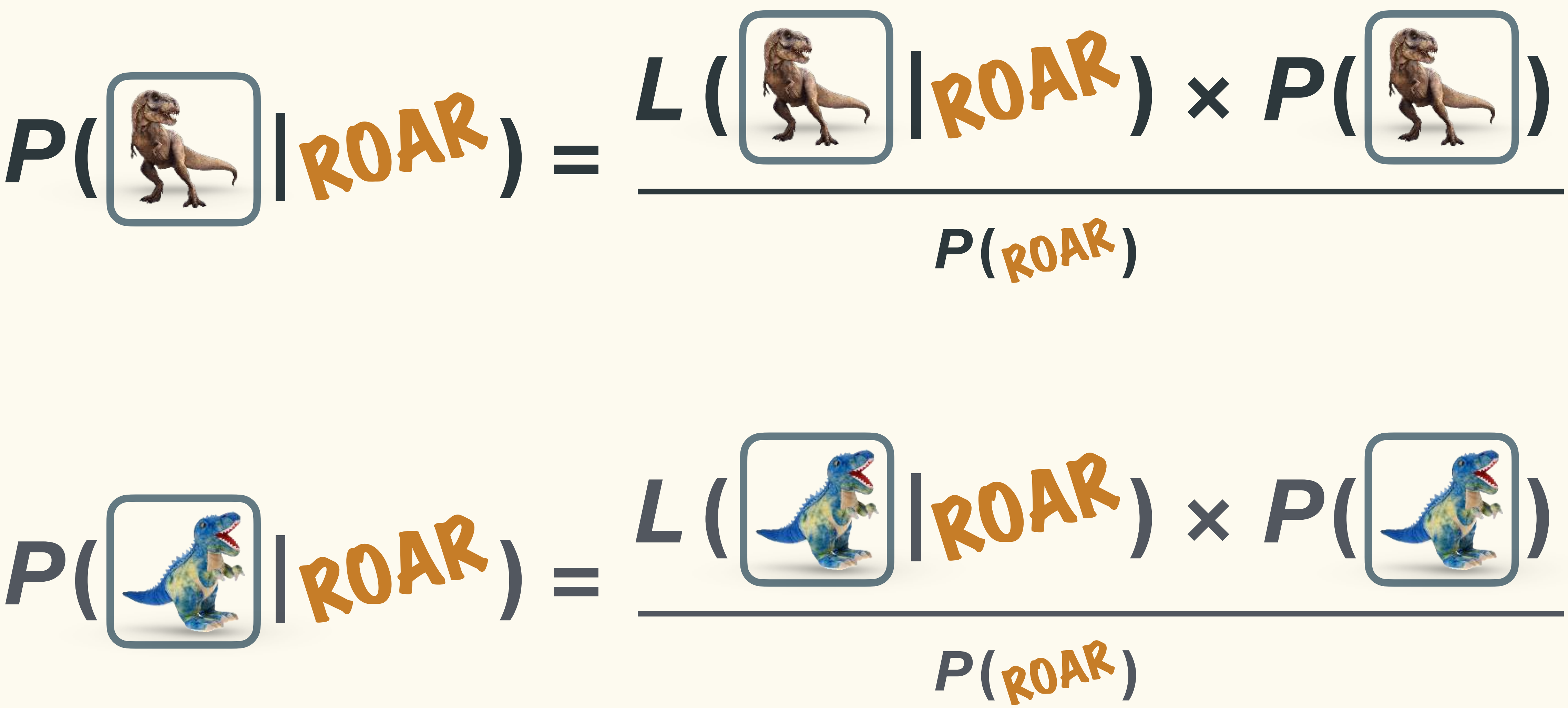
P(ROAR)

Bayes' theorem

Posterior probability **Likelihood** **Prior probability**

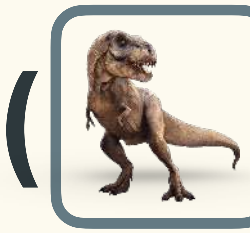


Bayes’ theorem



Bayes' theorem

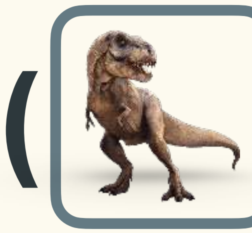
Bayes' theorem

P( | ROAR) = 1 × P() / P(ROAR)

≈ 1

P( | ROAR) = L( | ROAR) × P() / P(ROAR)

Bayes' theorem

P( | ROAR) = 1 × P() / P(ROAR)

P( | ROAR) = 1 × P() / P(ROAR)

Bayes' theorem

P( | ROAR) = P() / P(ROAR)

Prior probabilities

P( | ROAR) = P() / P(ROAR)

Bayes’ theorem

≈ 0.00000001

P( | ROAR) = P() / P(ROAR)

P( | ROAR) = P() / P(ROAR)

Bayes' theorem

P( | ROAR) = P() / P(ROAR) $\approx$ 0.00000001 / P(ROAR)

P( | ROAR) = P() / P(ROAR)

Bayes’ theorem

P( | ROAR) = P() / P(ROAR) $\approx$ 0.00000001 / P(ROAR)

$\approx$ 0.1

P( | ROAR) = P() / P(ROAR)

Bayes' theorem

P( | ROAR) = P() / P(ROAR) ≈ 0.00000001 / P(ROAR)

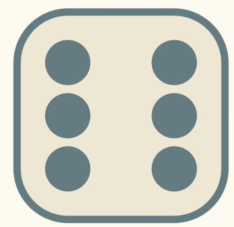
P( | ROAR) = P() / P(ROAR) ≈ 0.1 / P(ROAR)

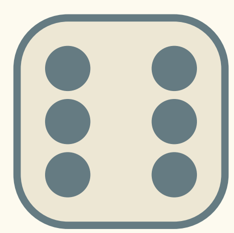
Bayes' theorem

P( | ROAR) ≫ P( | ROAR)

Likelihood

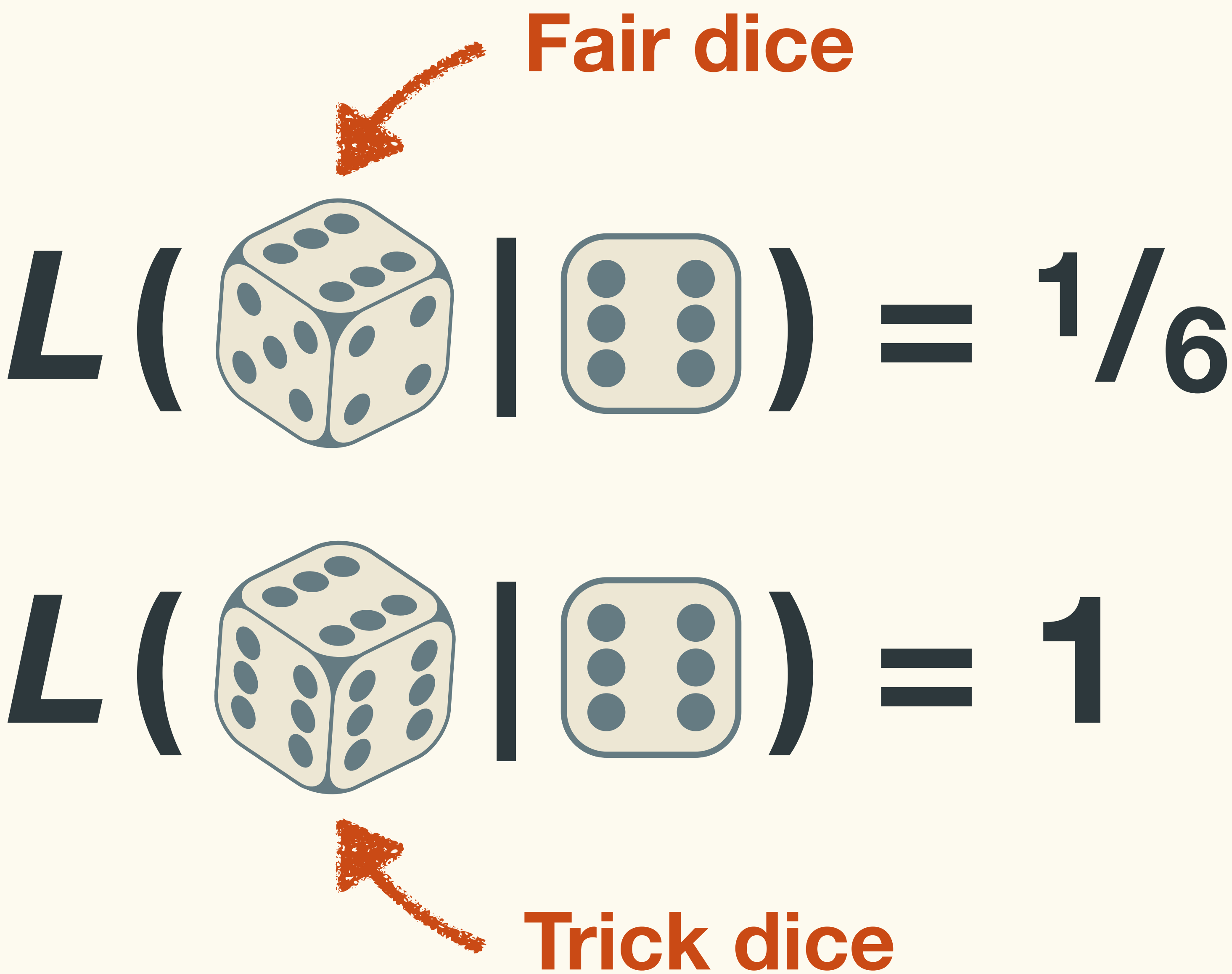
Fair dice

$L($  $|$  $) = 1/6$

$L($  $|$  $) = 1$

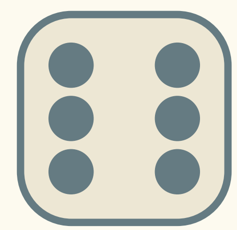
Trick dice

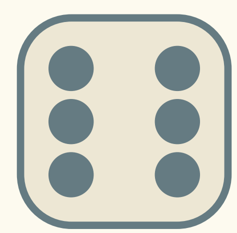
Likelihood



Likelihood

Fair dice

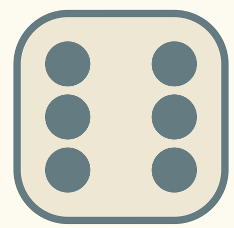
L ( | ) = 1/6

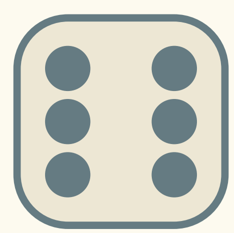
L ( | ) = 1

Trick dice

Likelihood

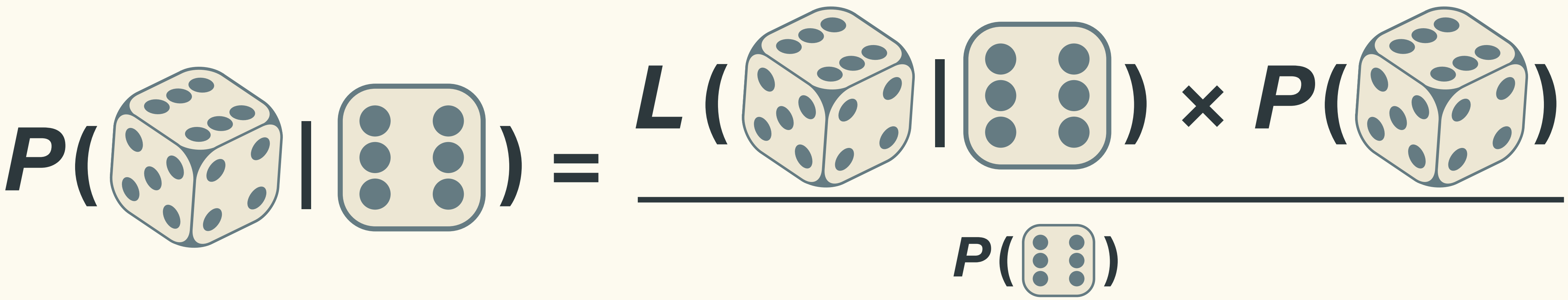
Fair dice

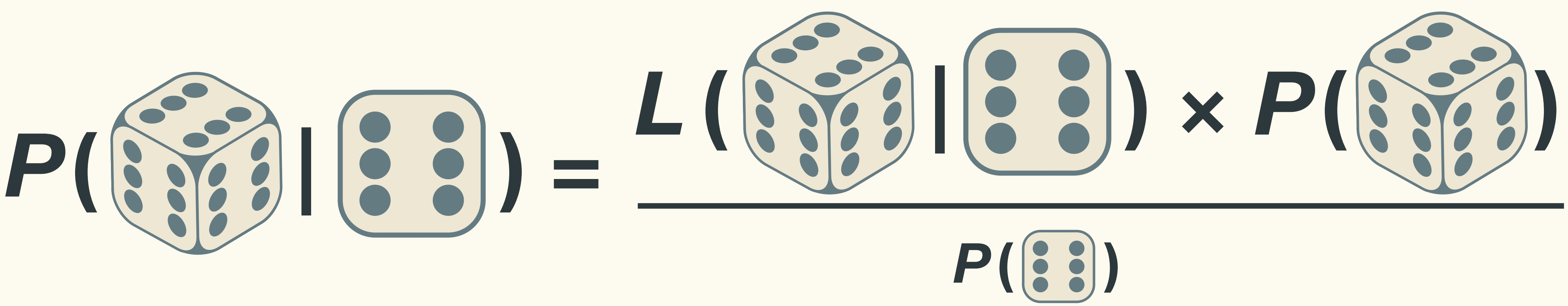
$L($  $|$  $) = 1/6$

$L($  $|$  $) = 1$

Trick dice

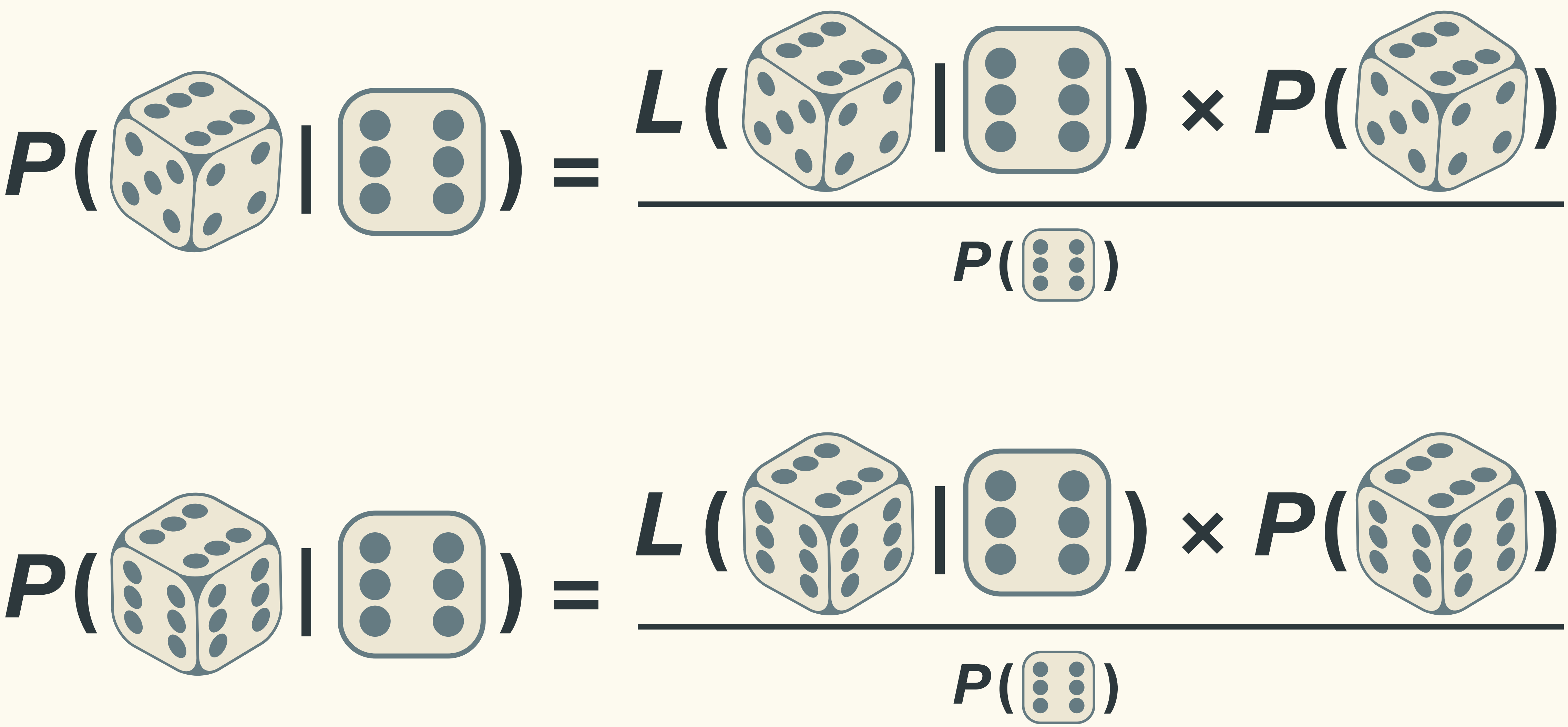
Bayes' theorem





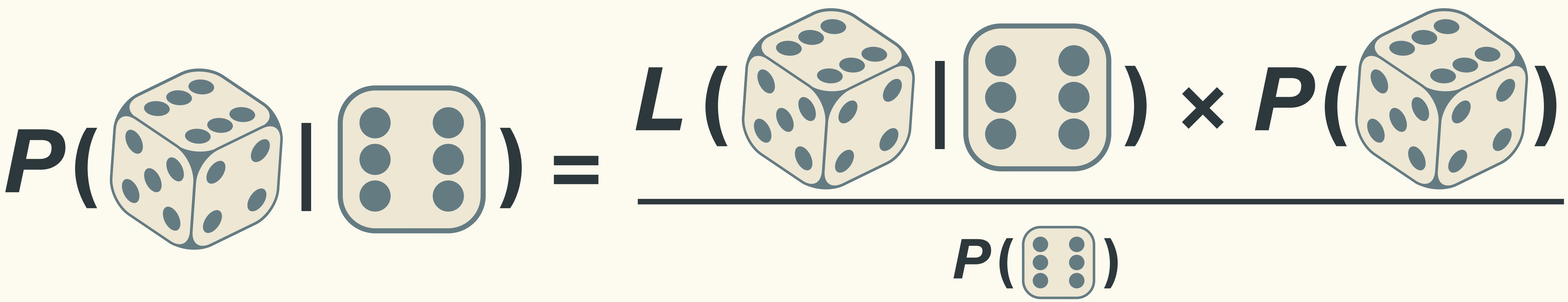
Bayes’ theorem

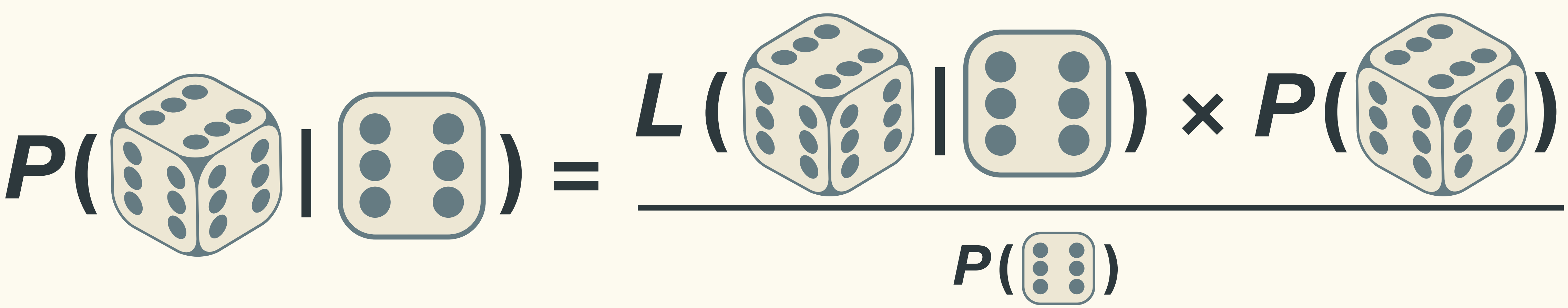
Posterior probability **Likelihood** **Prior probability**



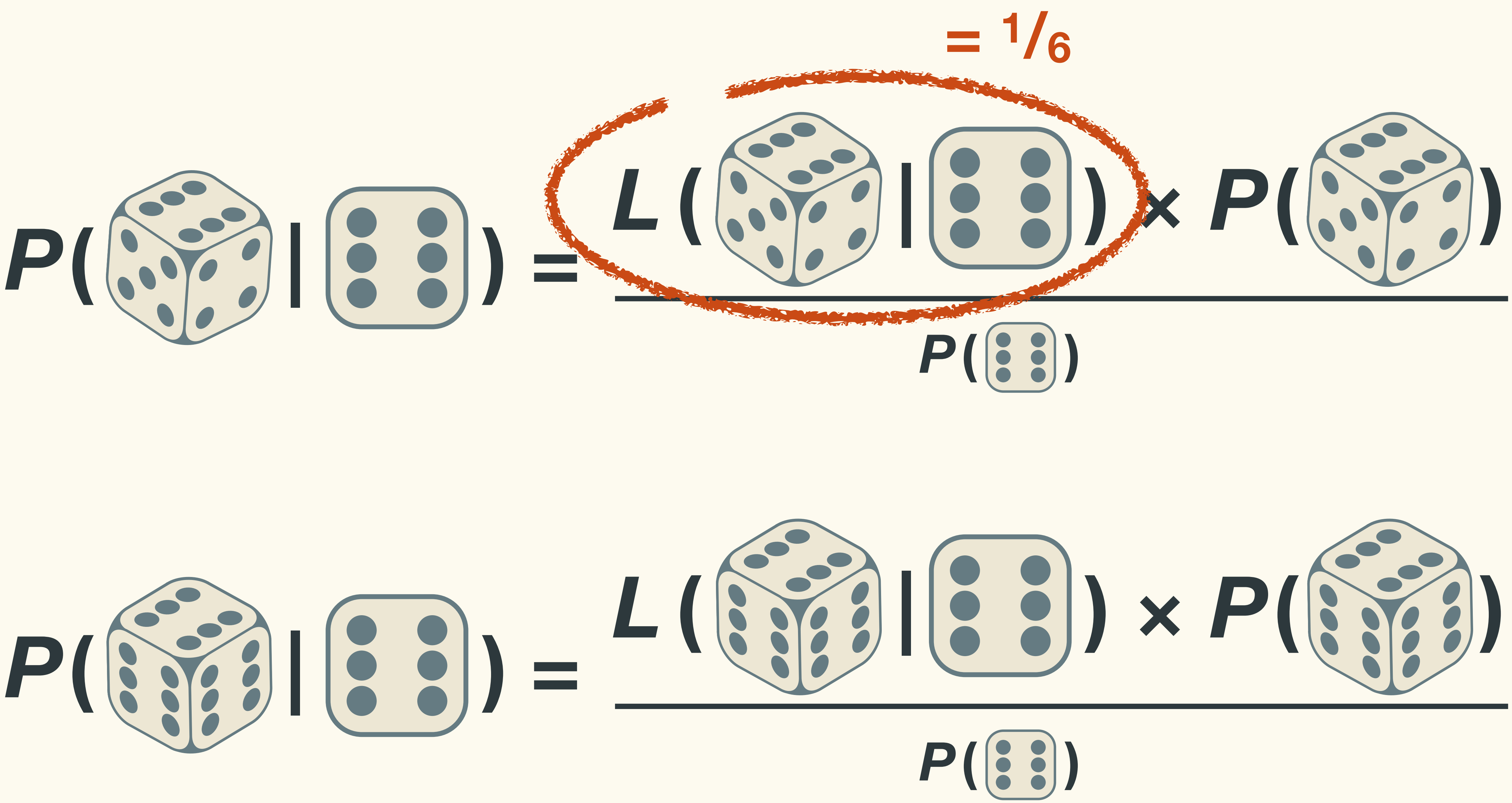
Bayes’ theorem

Posterior probability **Likelihood** **Prior probability**

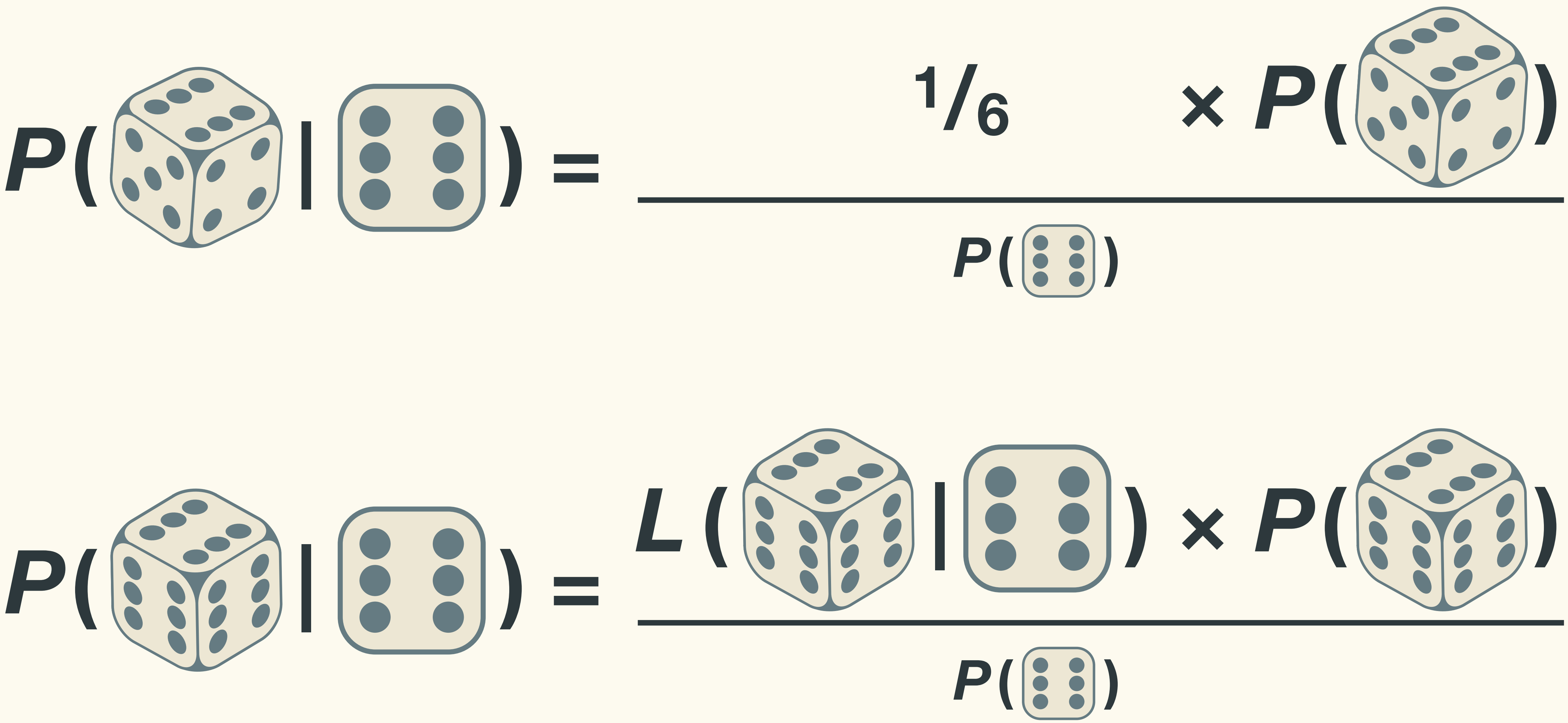




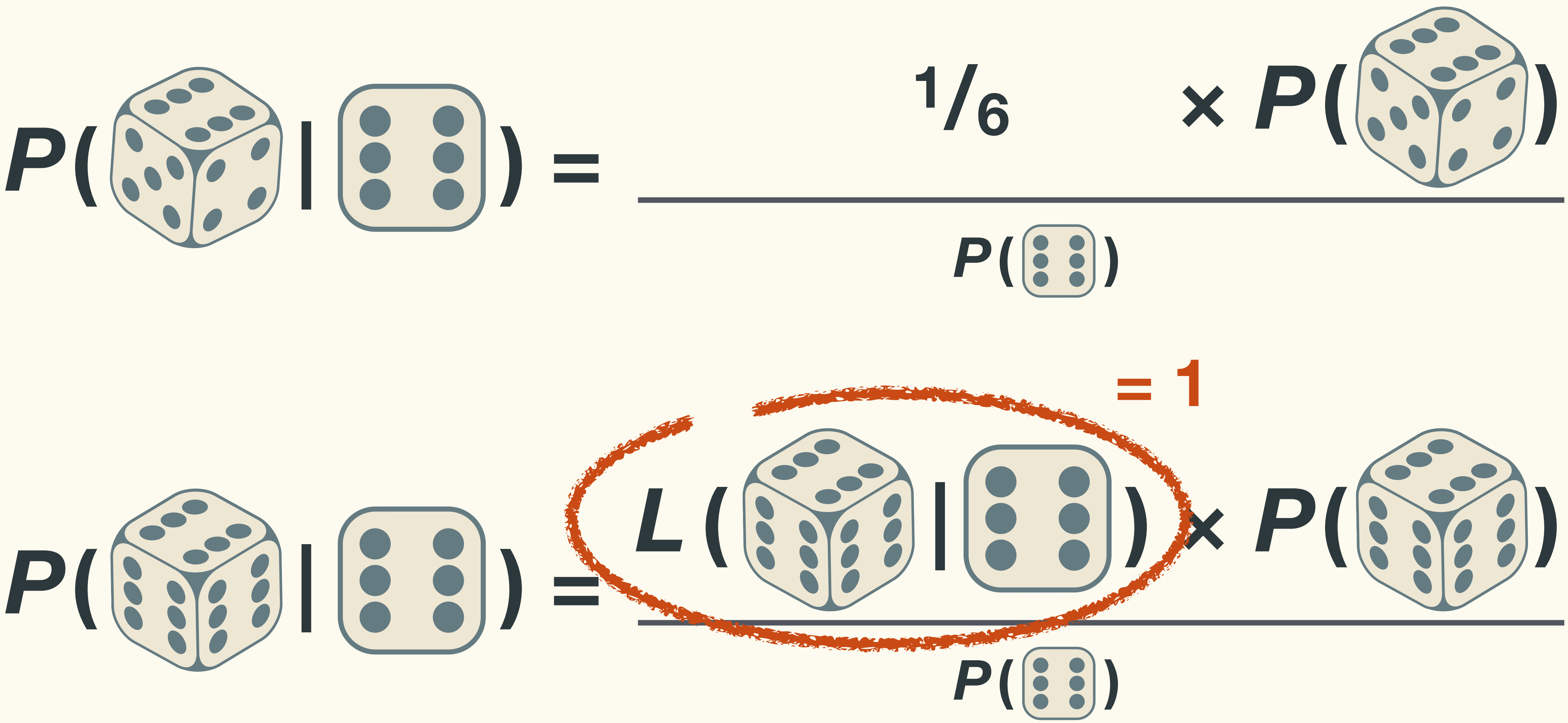
Bayes' theorem



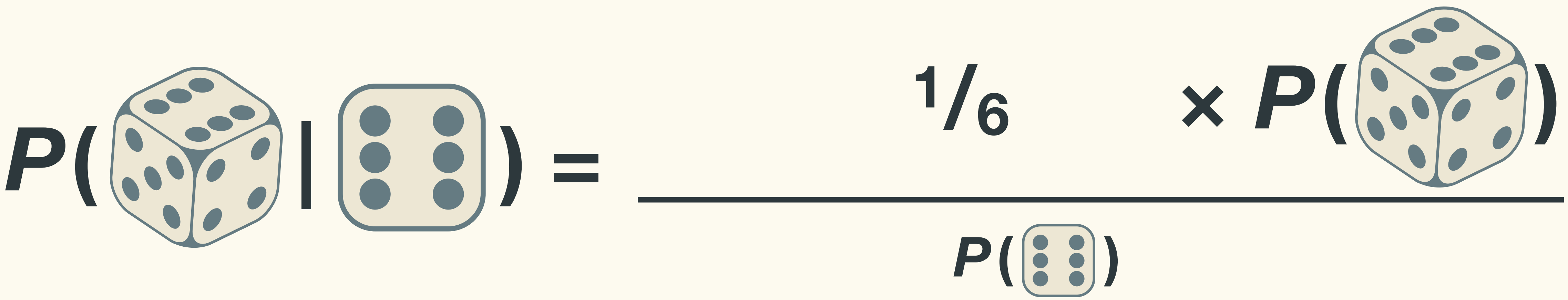
Bayes’ theorem



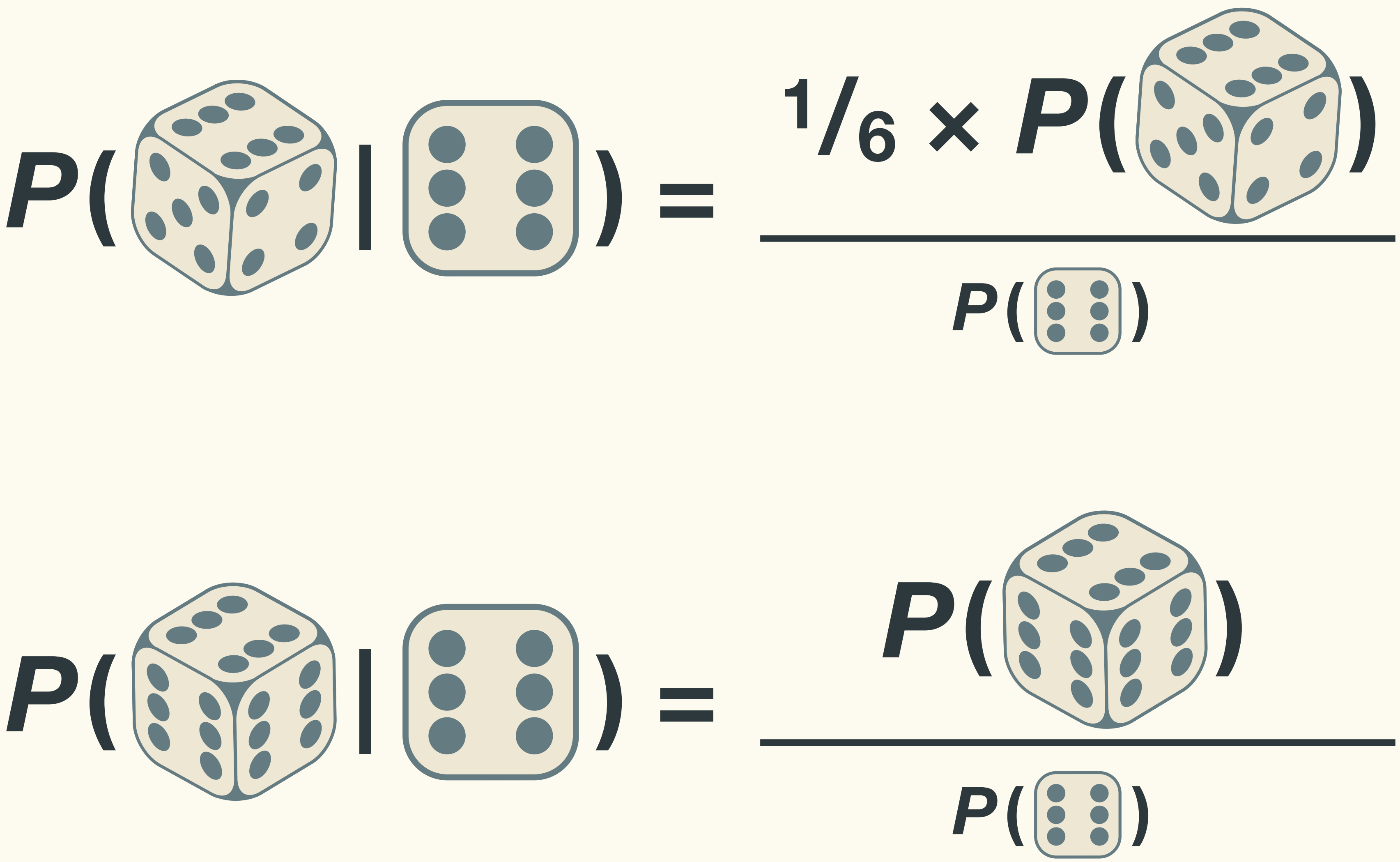
Bayes' theorem



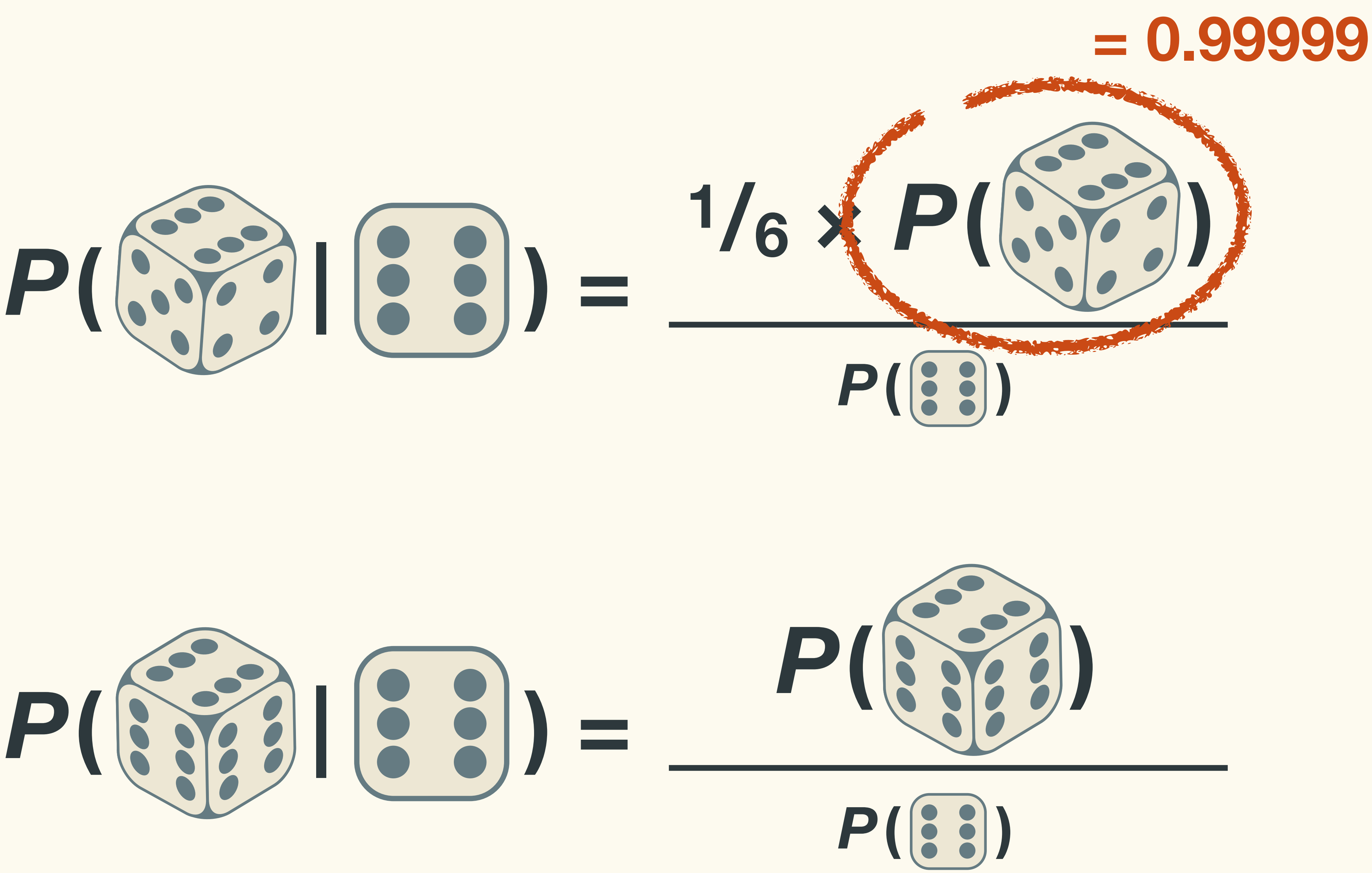
Bayes’ theorem



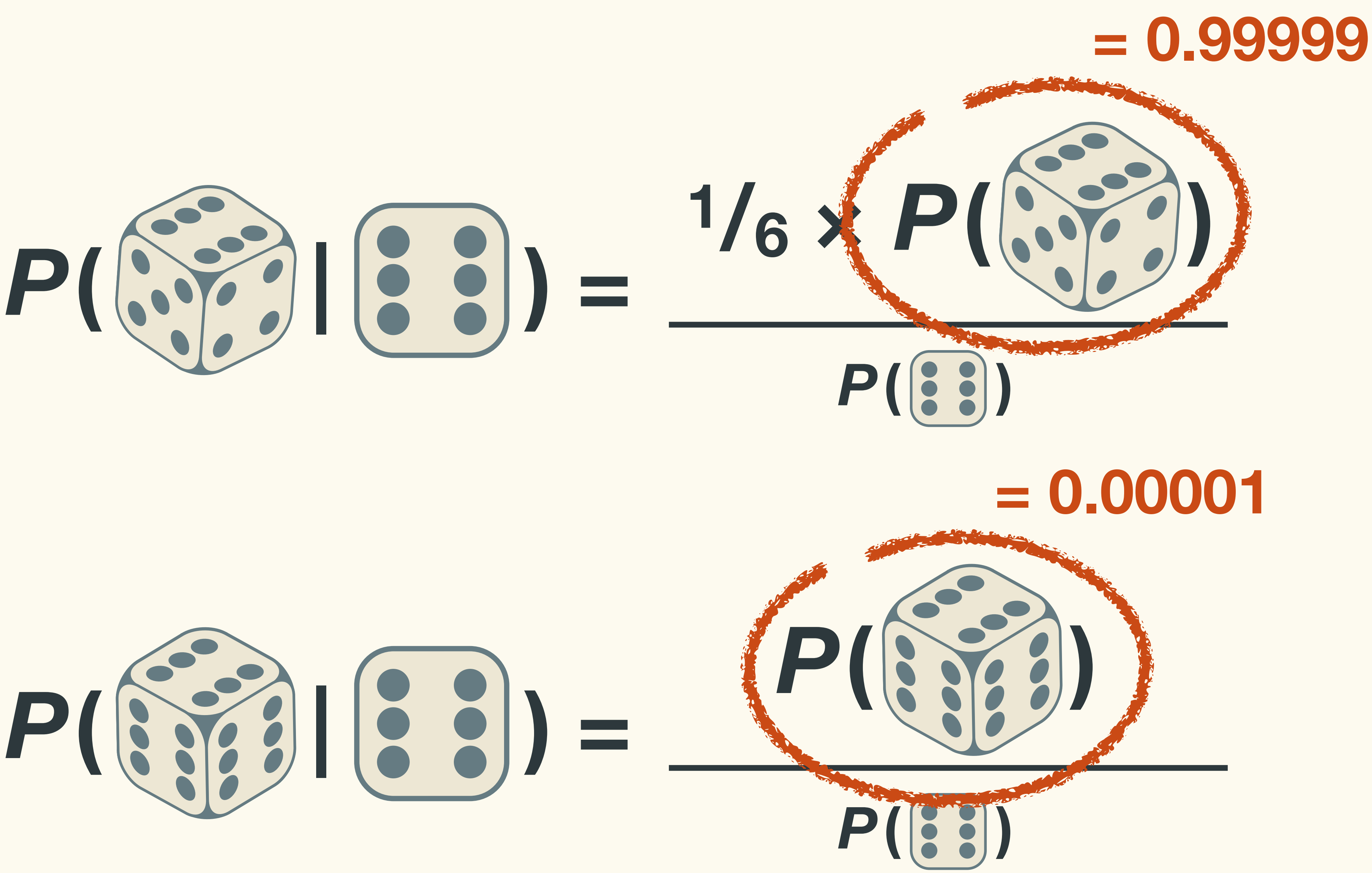
Bayes' theorem



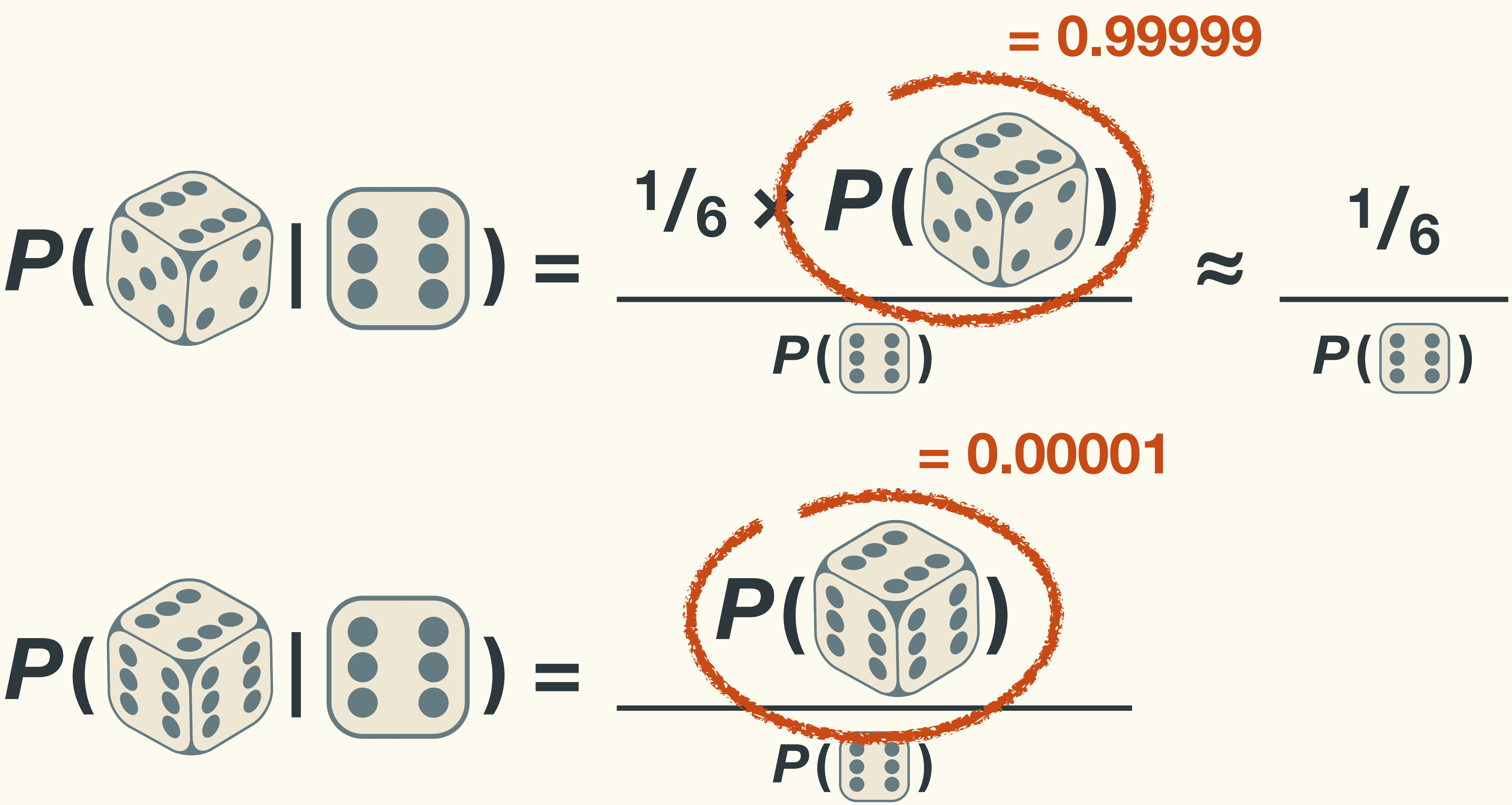
Bayes' theorem



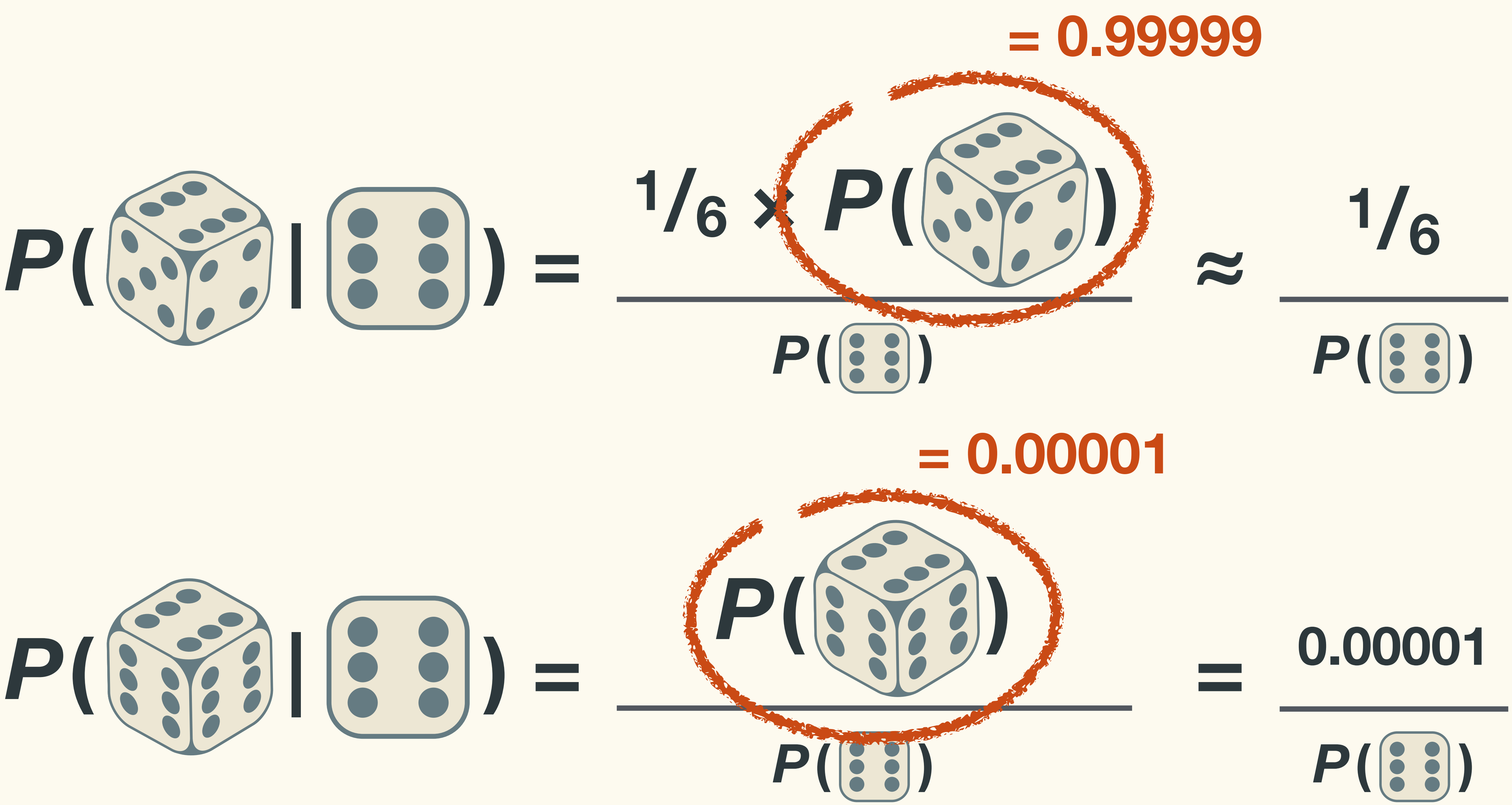
Bayes' theorem



Bayes' theorem



Bayes' theorem



Estimating model parameters using Bayesian inference

MCMC

Markov-chain Monte Carlo

Monte Carlo methods

Monte Carlo methods

Stanisław Ulam
(1909–1984)

Monte Carlo methods

Stanisław Ulam
(1909–1984)

Monte Carlo methods

Stanisław Ulam
(1909–1984)

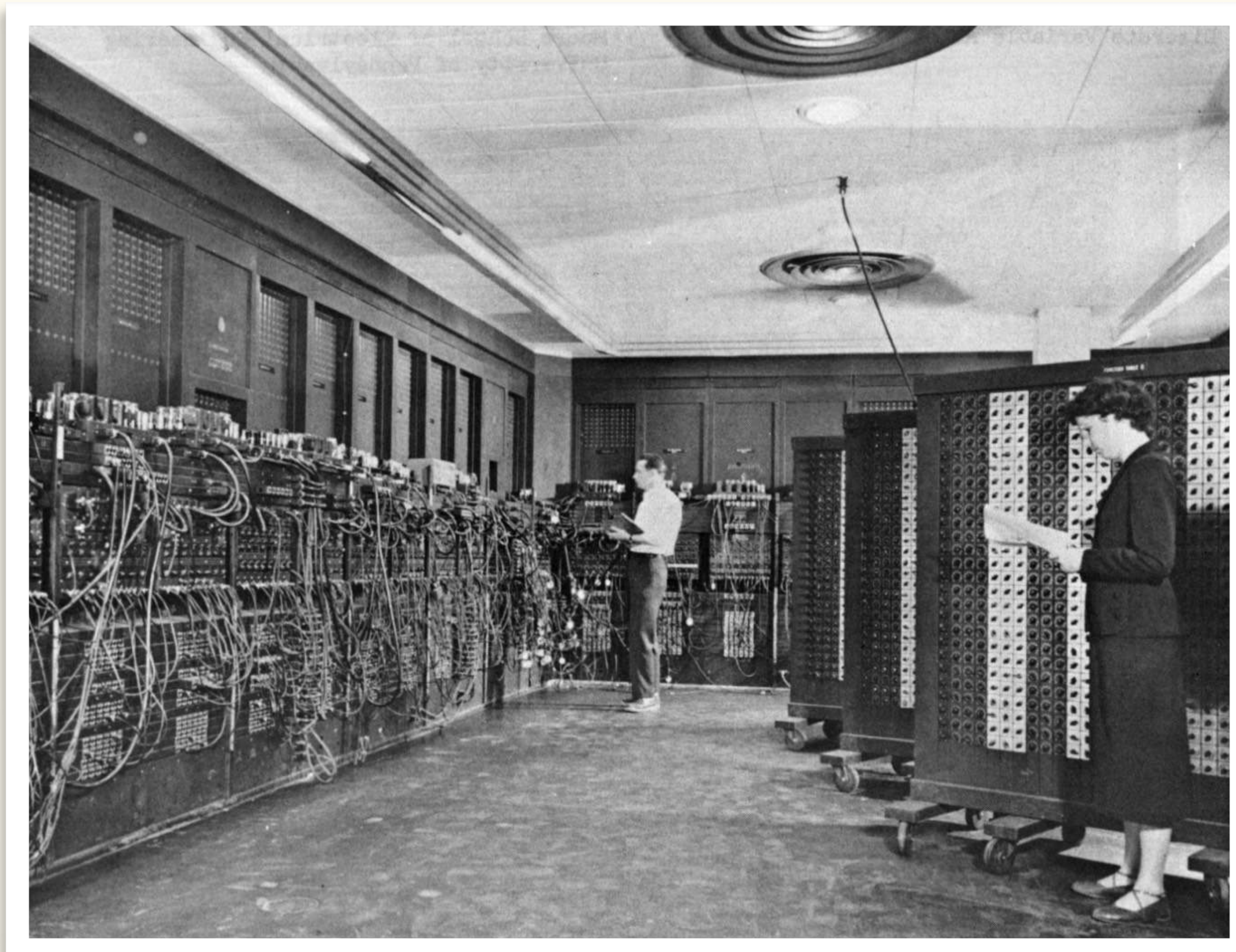
"What are the chances that a Canfield solitaire laid out with 52 cards will come out successfully?"

Stanisław Ulam, 1946

Eckhardt (1987) *Los Alamos Science* Special Issue, 131–143

Monte Carlo methods

ENIAC
1946



Monte Carlo methods

Stanisław Ulam
(1909–1984)

“What are the chances that a Canfield solitaire laid out with 52 cards will come out successfully?”

Stanisław Ulam, 1946

Eckhardt (1987) *Los Alamos Science* Special Issue, 131–143

Monte Carlo methods

Stanisław Ulam
(1909–1984)

Eckhardt (1987) *Los Alamos Science* Special Issue, 131–143

Monte Carlo methods

Stanisław Ulam
(1909–1984)

"Stan had an uncle who would borrow money from relatives because he 'just had to go to Monte Carlo'."

Nicholas Metropolis

Eckhardt (1987) *Los Alamos Science* Special Issue, 131–143

Monte Carlo methods

Markov chains

Andrey Markov
(1856–1922)

Markov chains

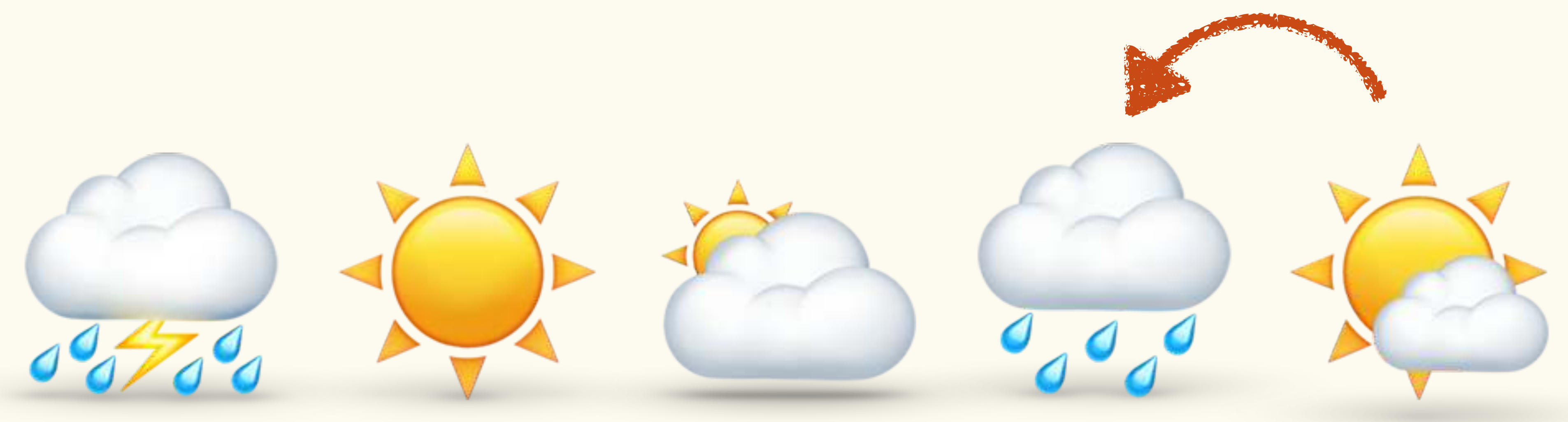
Markov chains

Markov chains

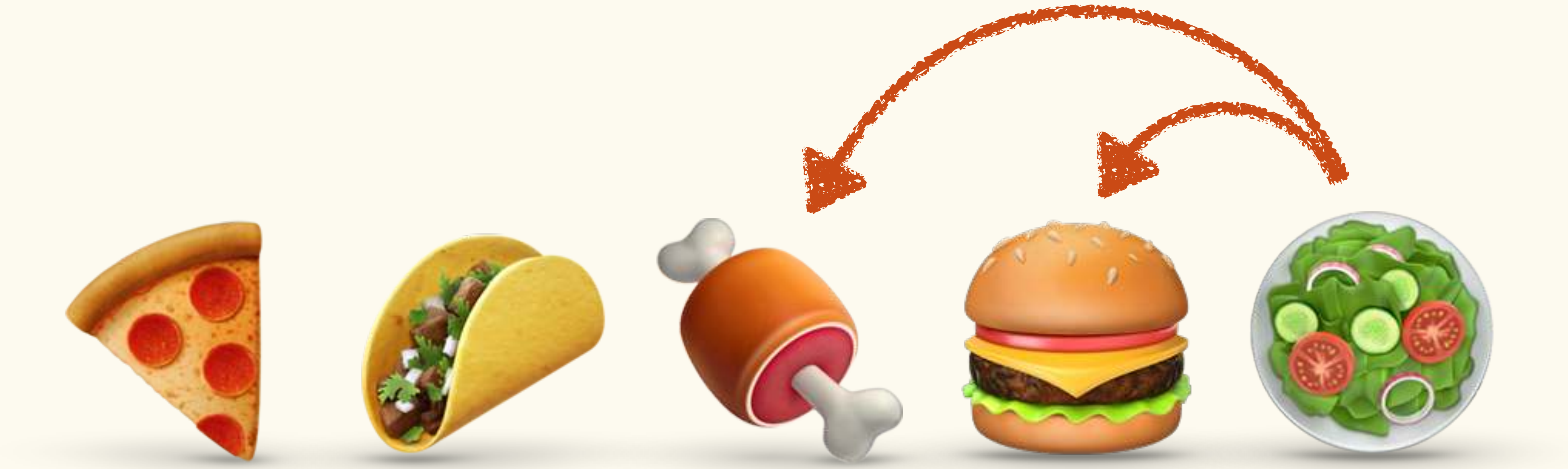
Markov chains

Markov chains

Markov chains



Markov chains



MCMC

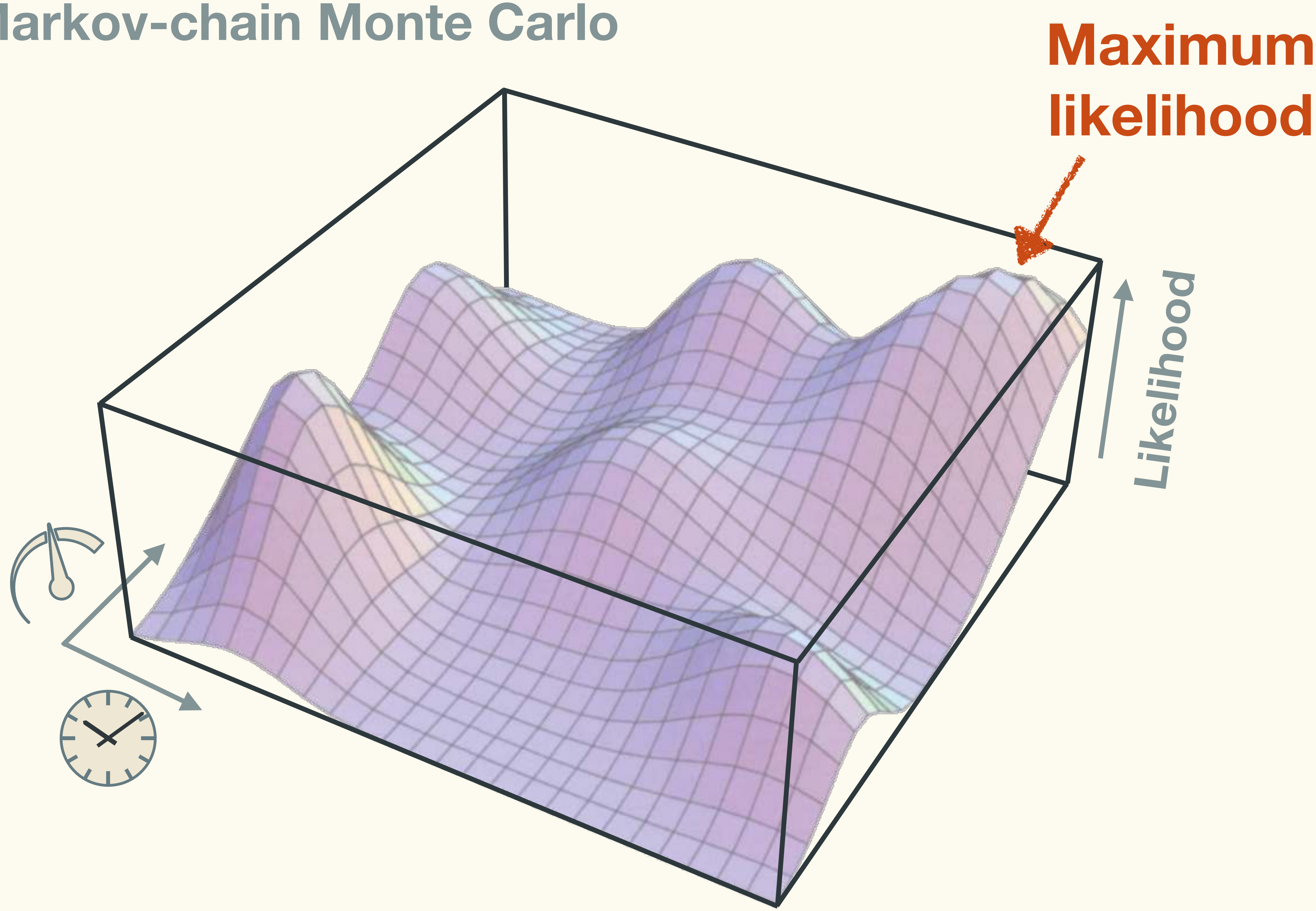
Markov-chain

Monte Carlo

Hastings (1970) *Biometrika* 57, 97–109; Metropolis *et al*. (1953) *J Chem Phys* 21, 1087–1092

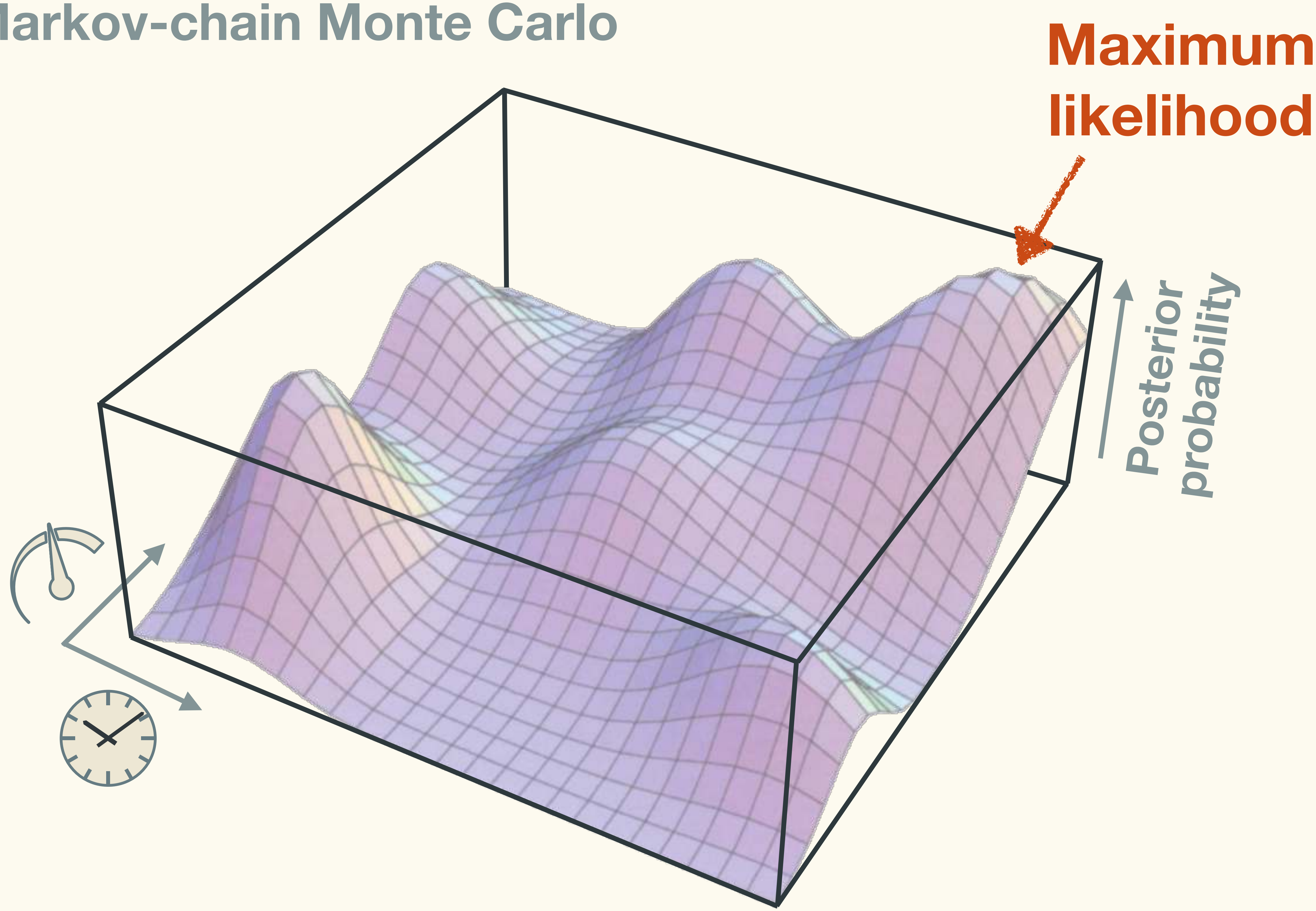
MCMC

Markov-chain Monte Carlo



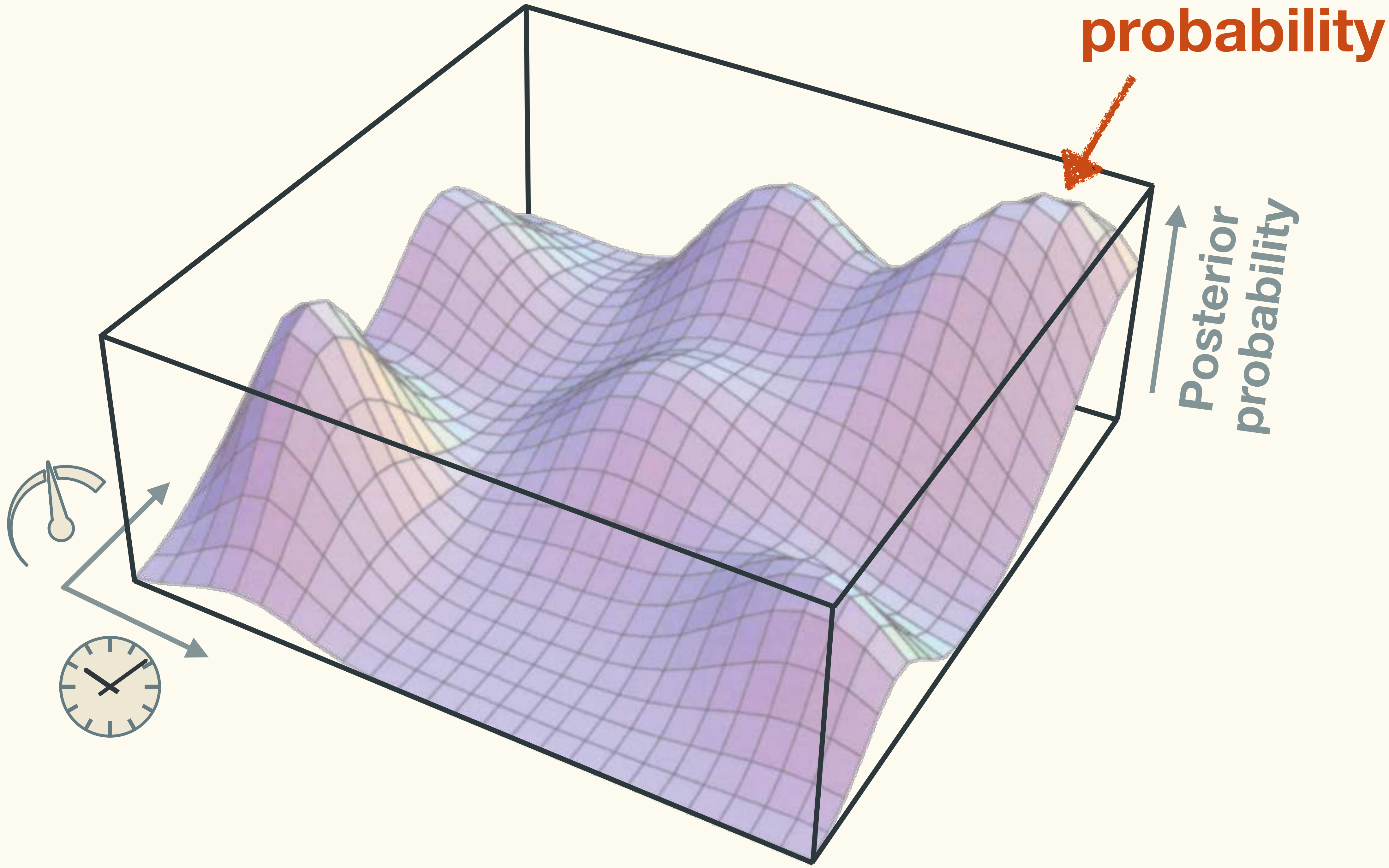
MCMC

Markov-chain Monte Carlo



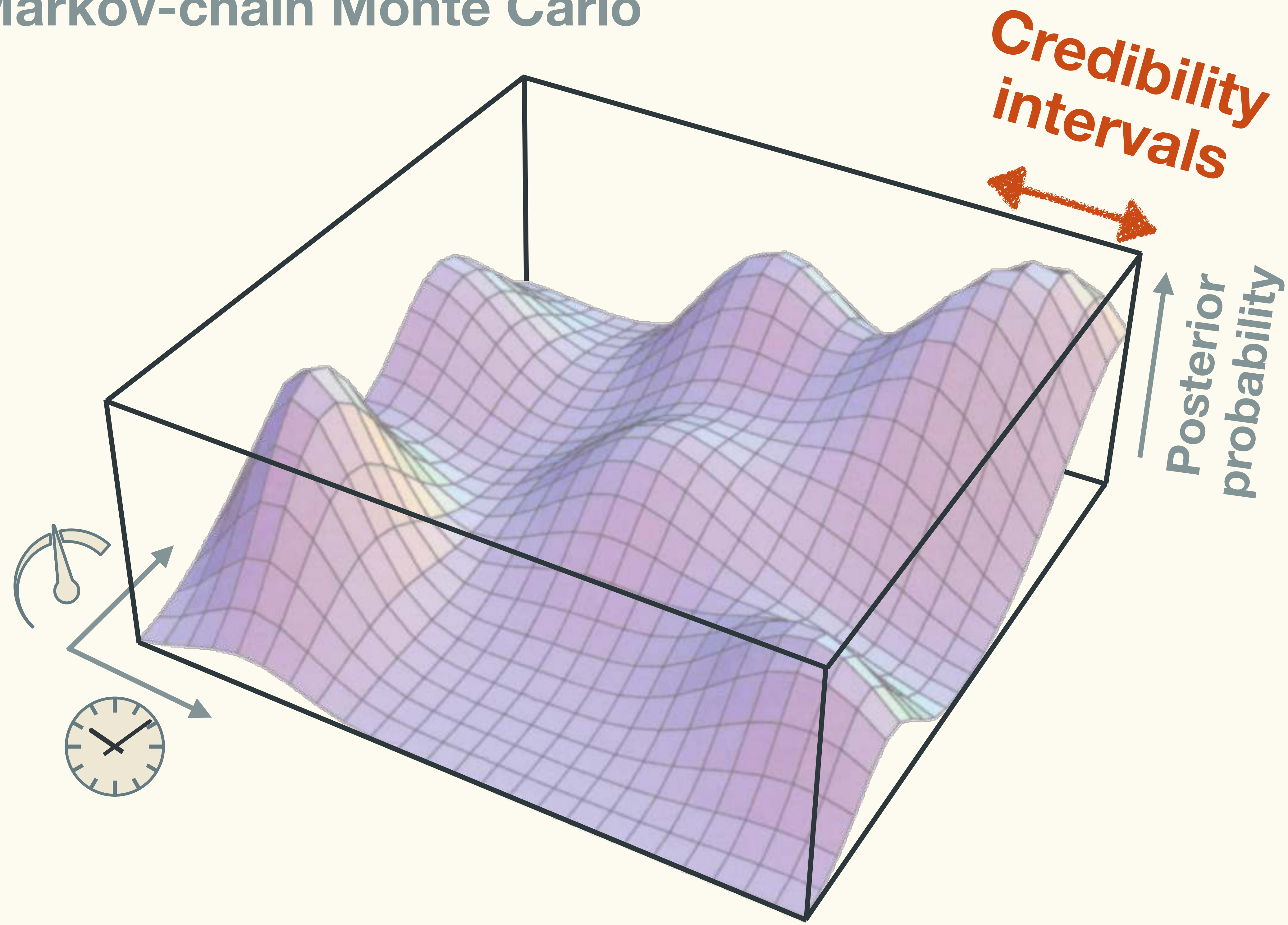
MCMC

Markov-chain Monte Carlo



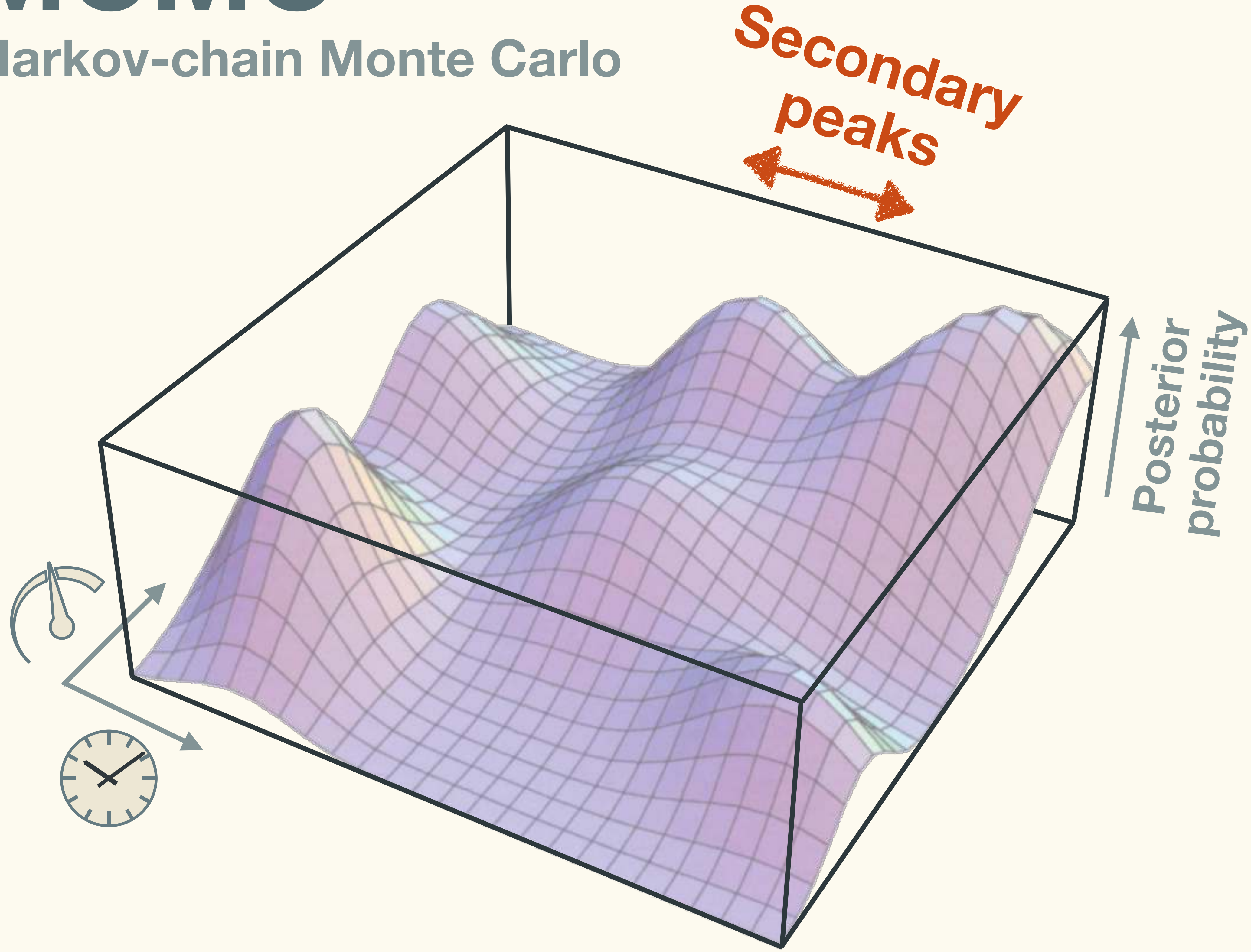
MCMC

Markov-chain Monte Carlo



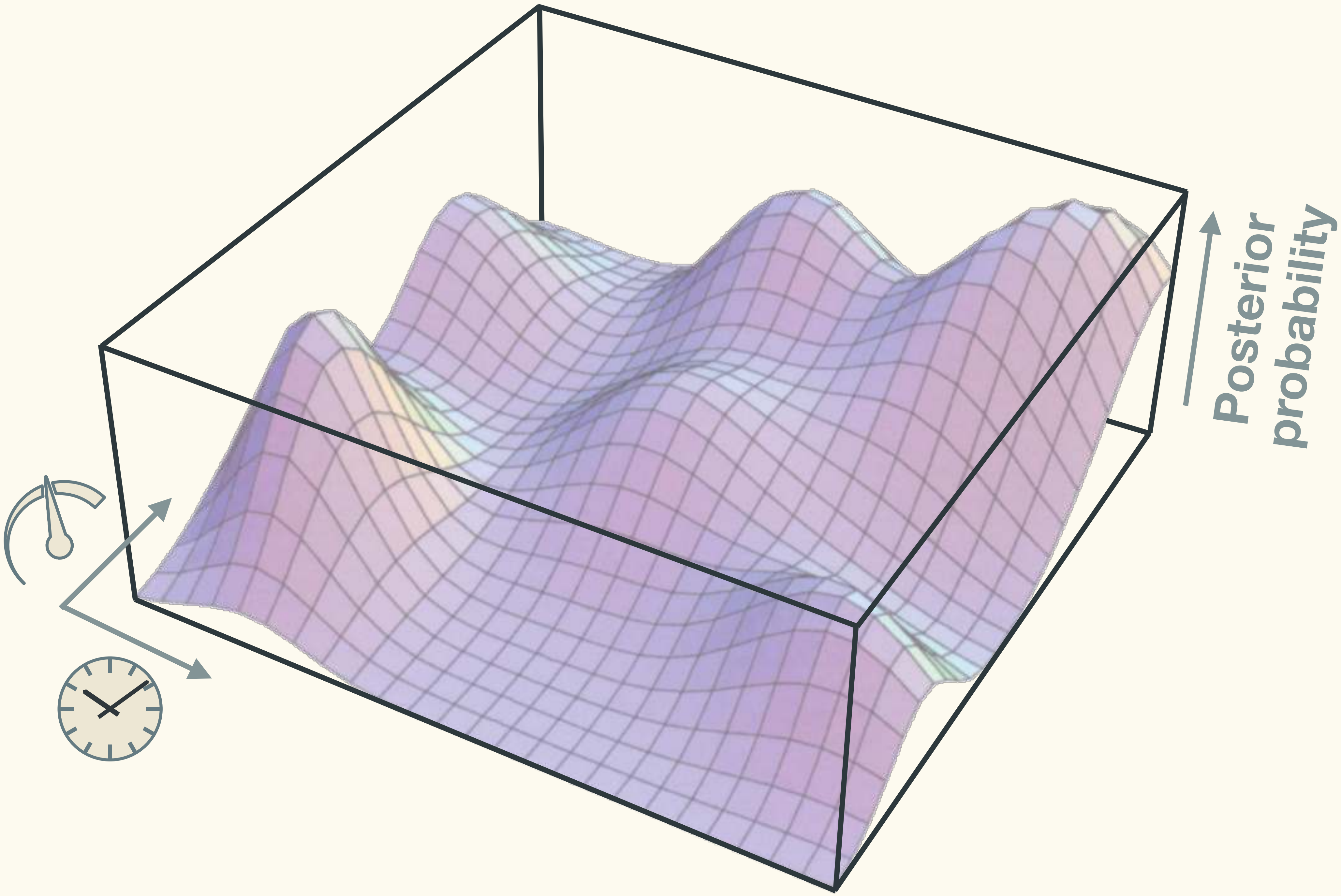
MCMC

Markov-chain Monte Carlo



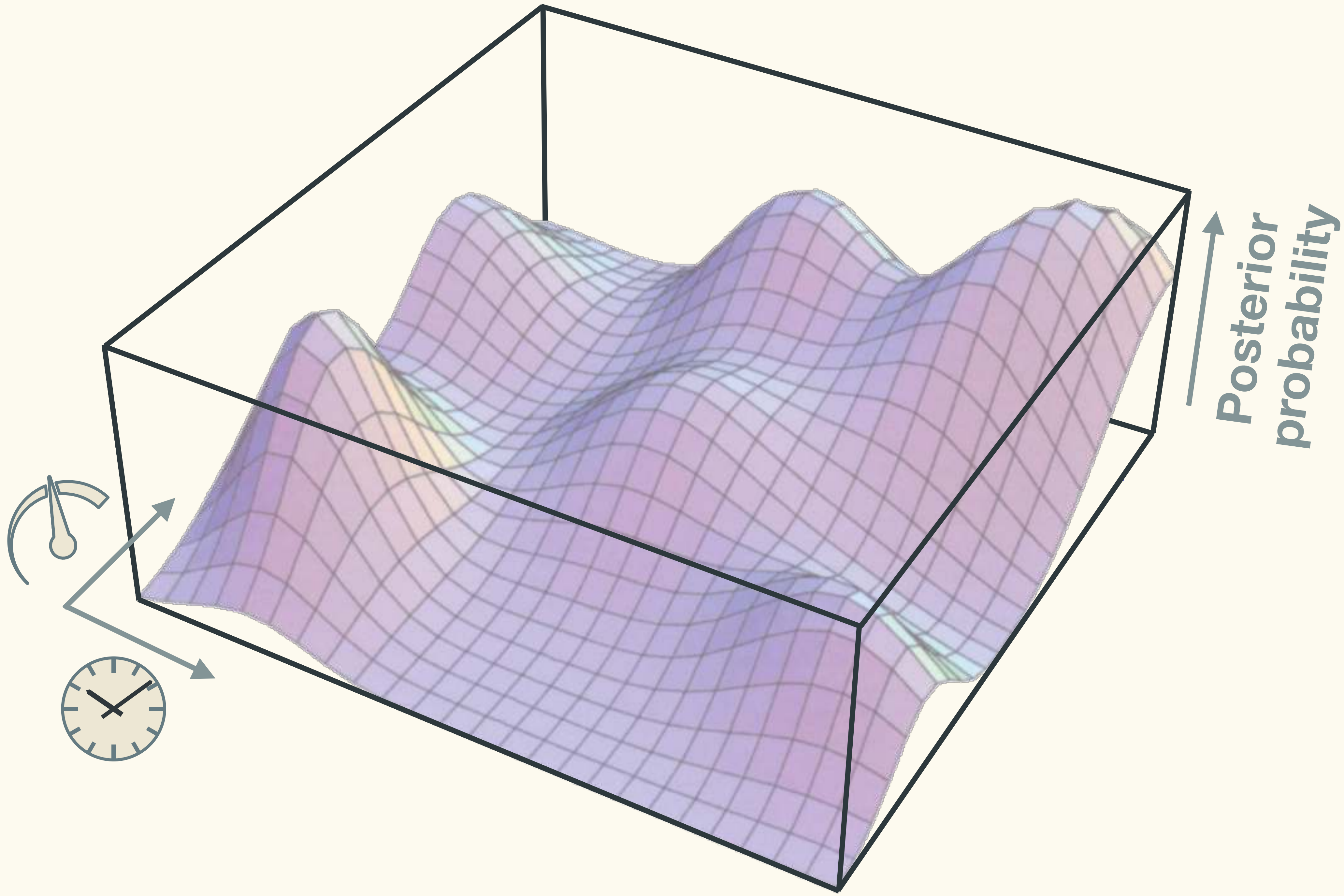
MCMC

Markov-chain Monte Carlo



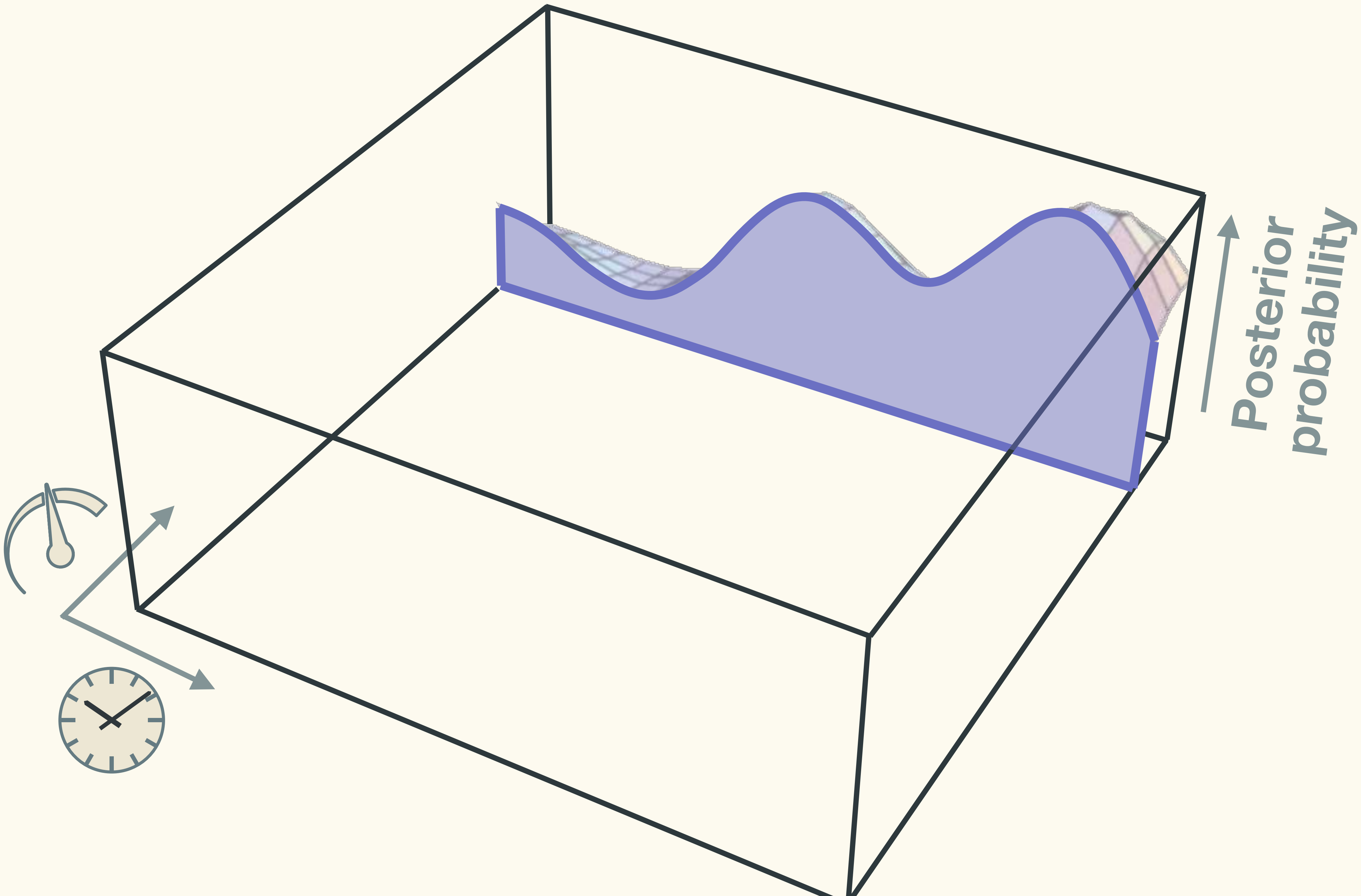
MCMC

Markov-chain Monte Carlo



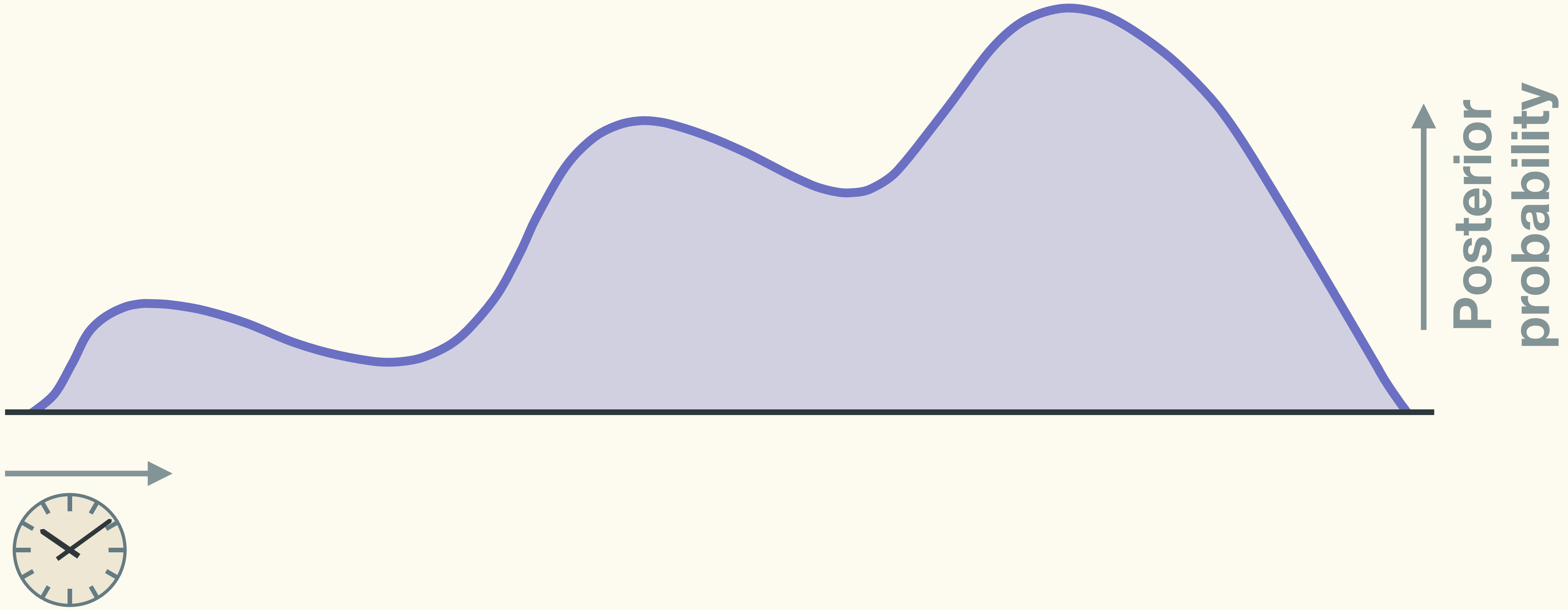
MCMC

Markov-chain Monte Carlo



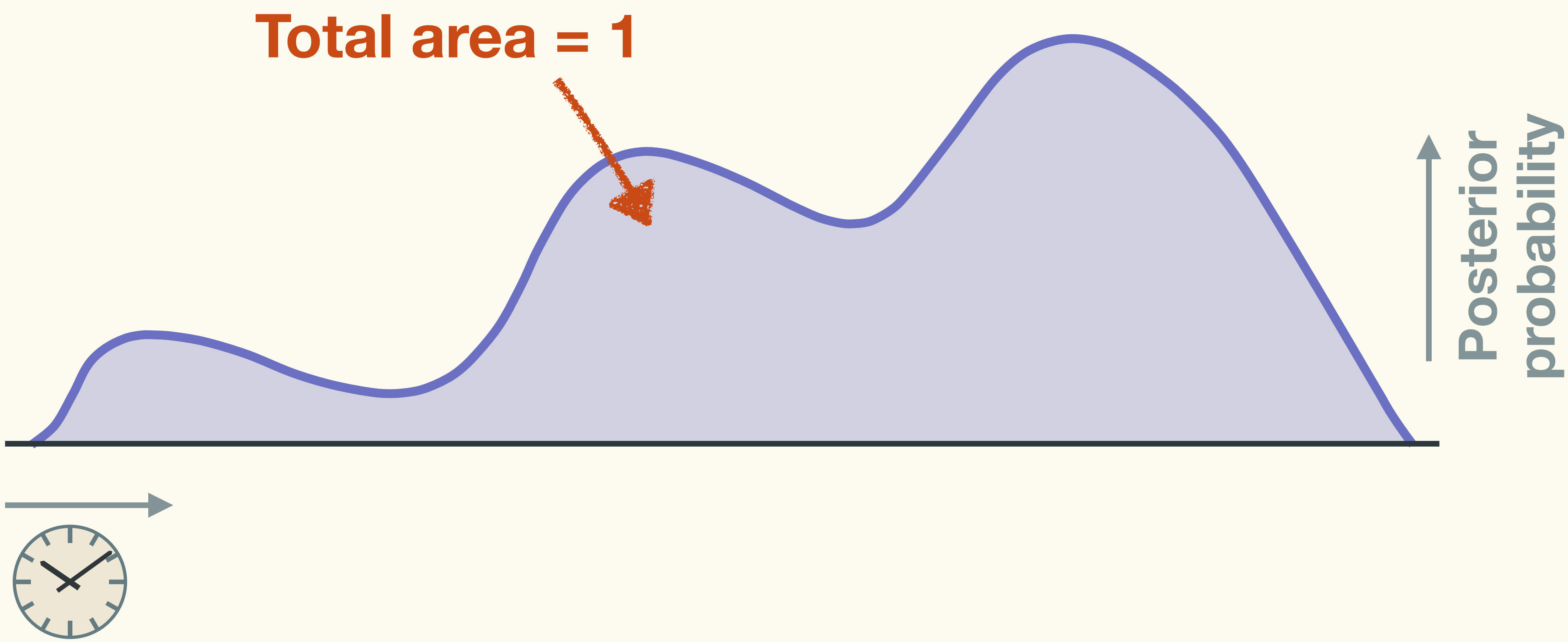
MCMC

Markov-chain Monte Carlo



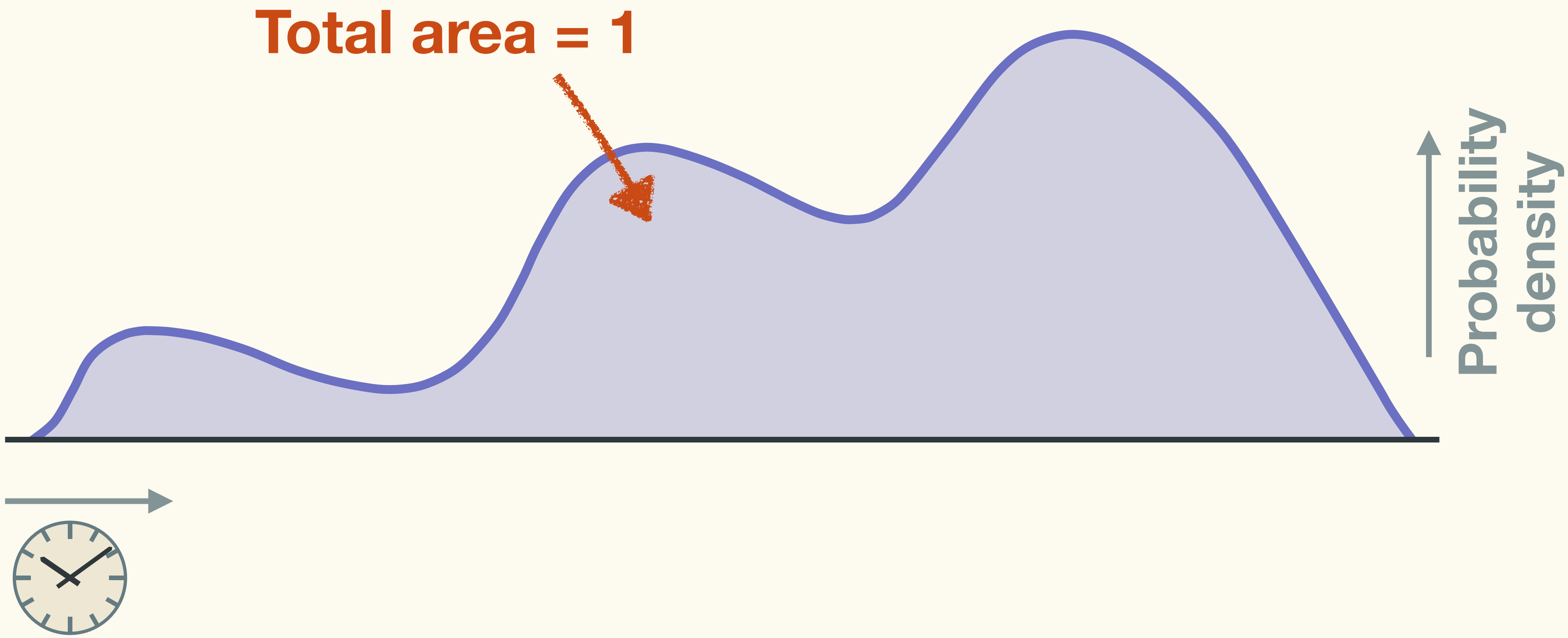
MCMC

Markov-chain Monte Carlo



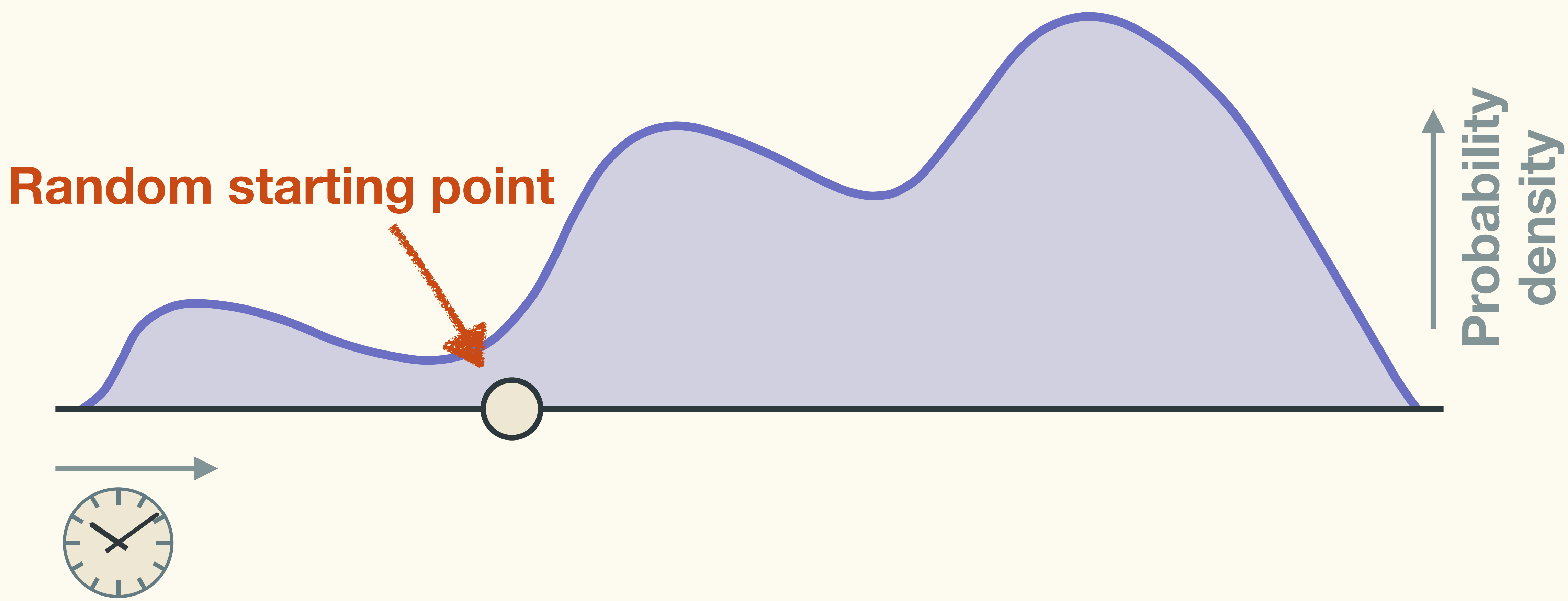
MCMC

Markov-chain Monte Carlo



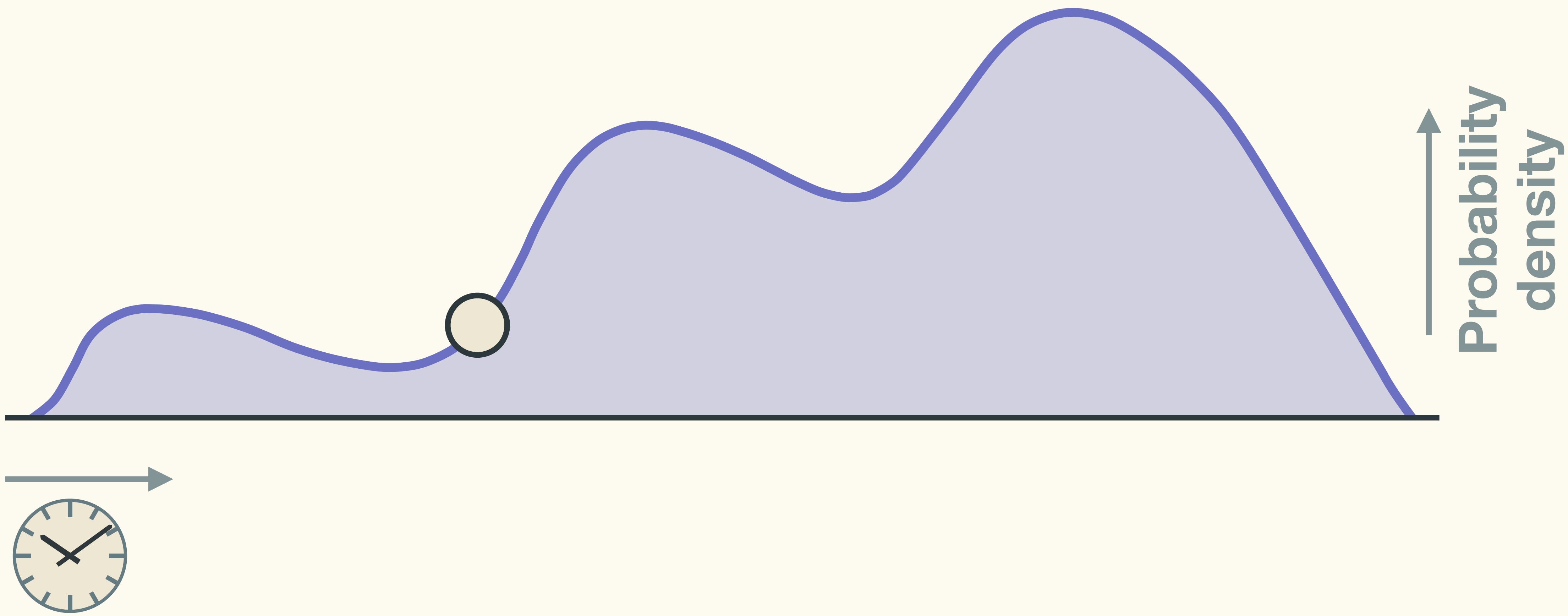
MCMC

Markov-chain Monte Carlo



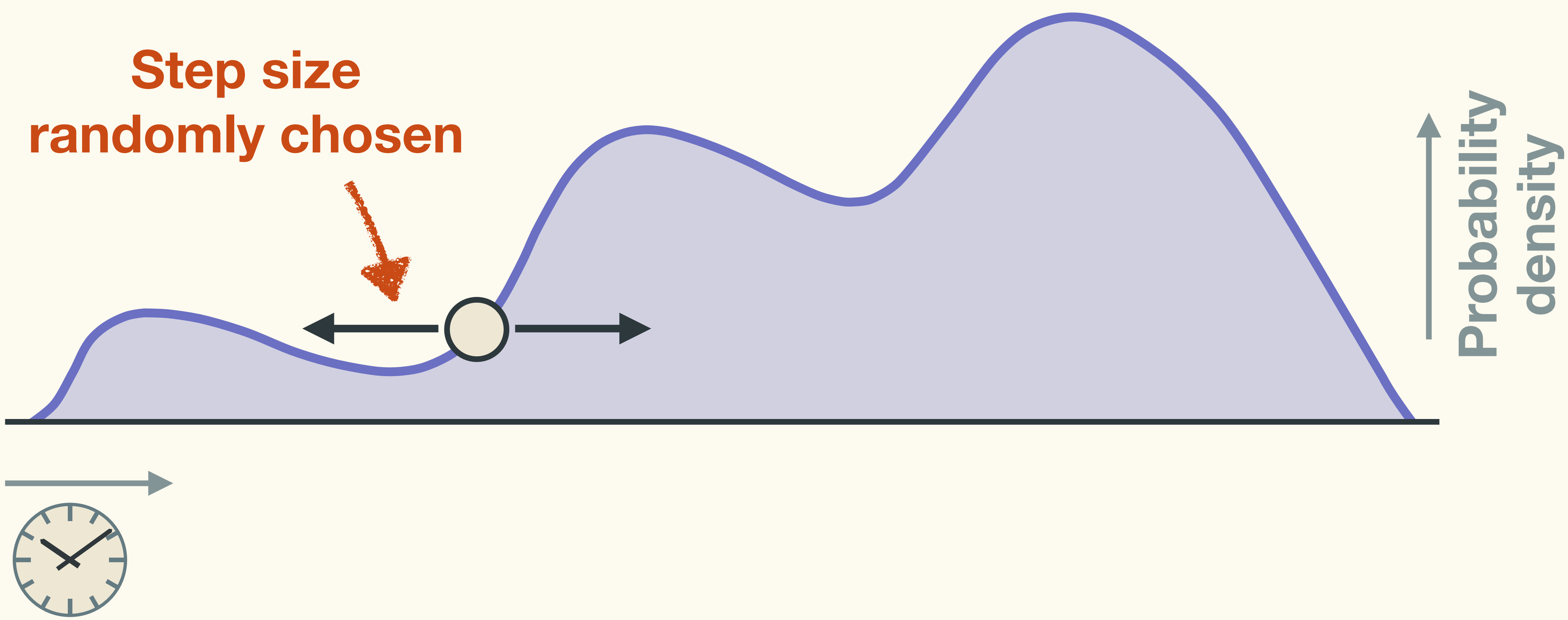
MCMC

Markov-chain Monte Carlo



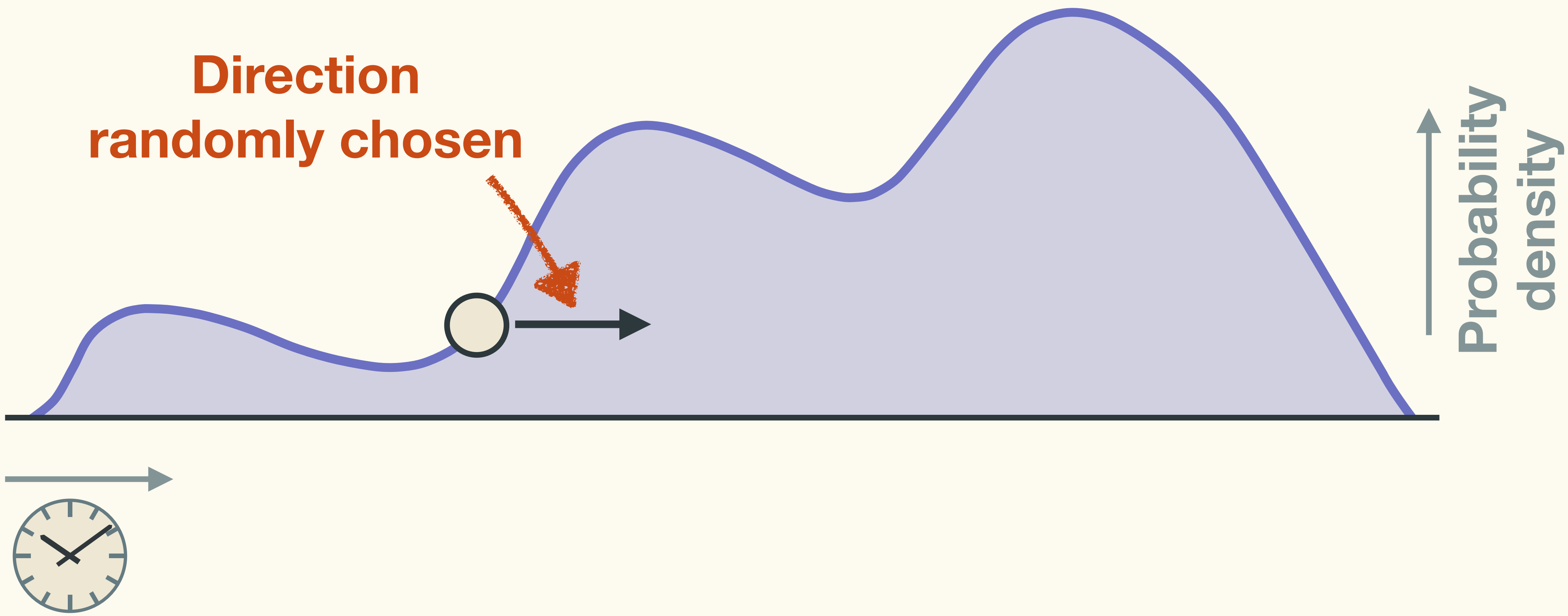
MCMC

Markov-chain Monte Carlo



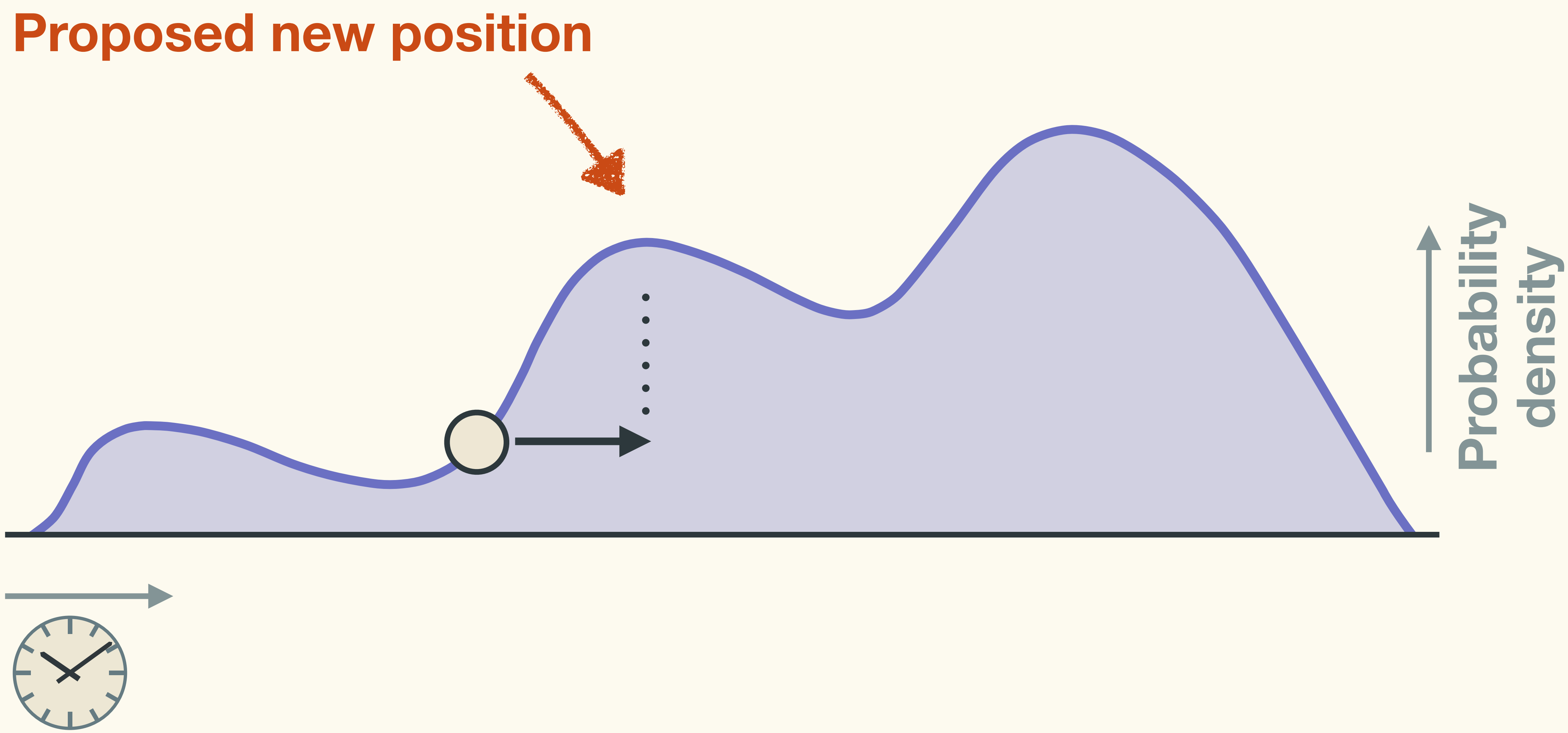
MCMC

Markov-chain Monte Carlo



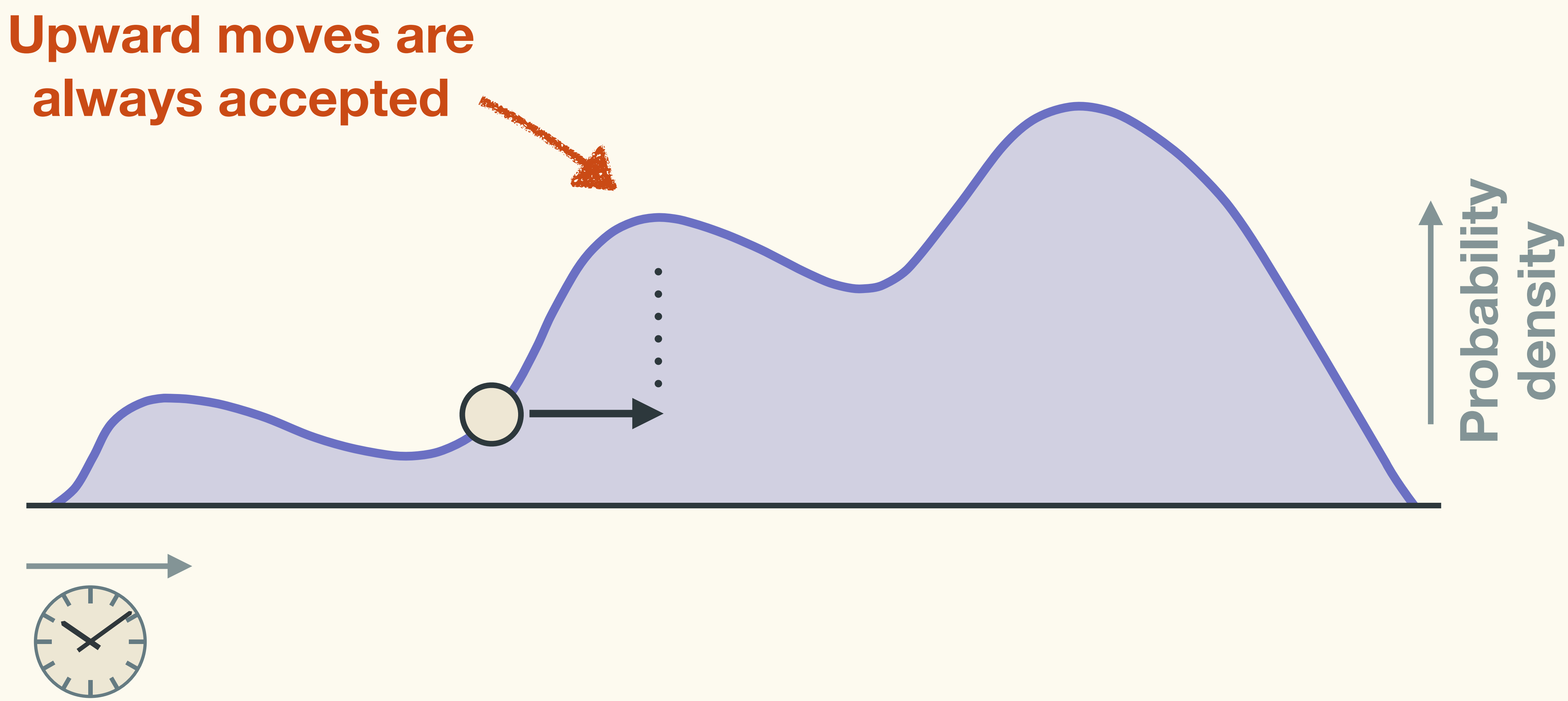
MCMC

Markov-chain Monte Carlo



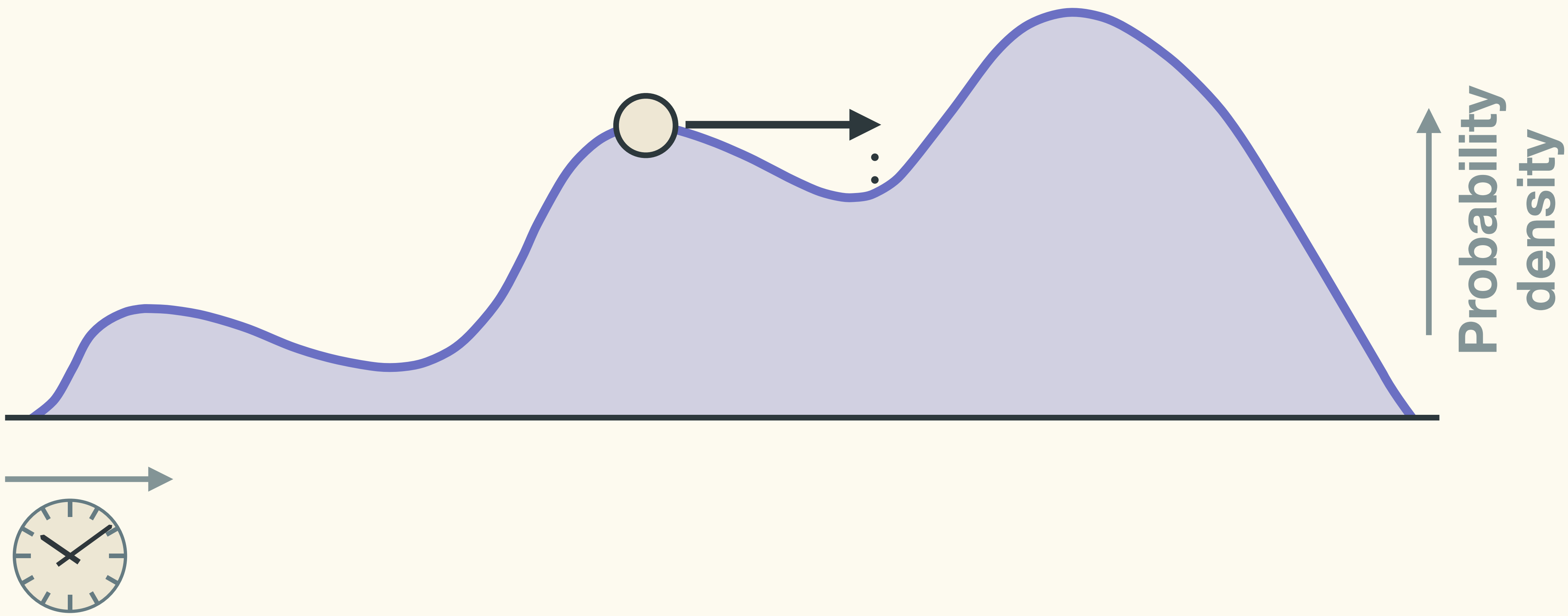
MCMC

Markov-chain Monte Carlo



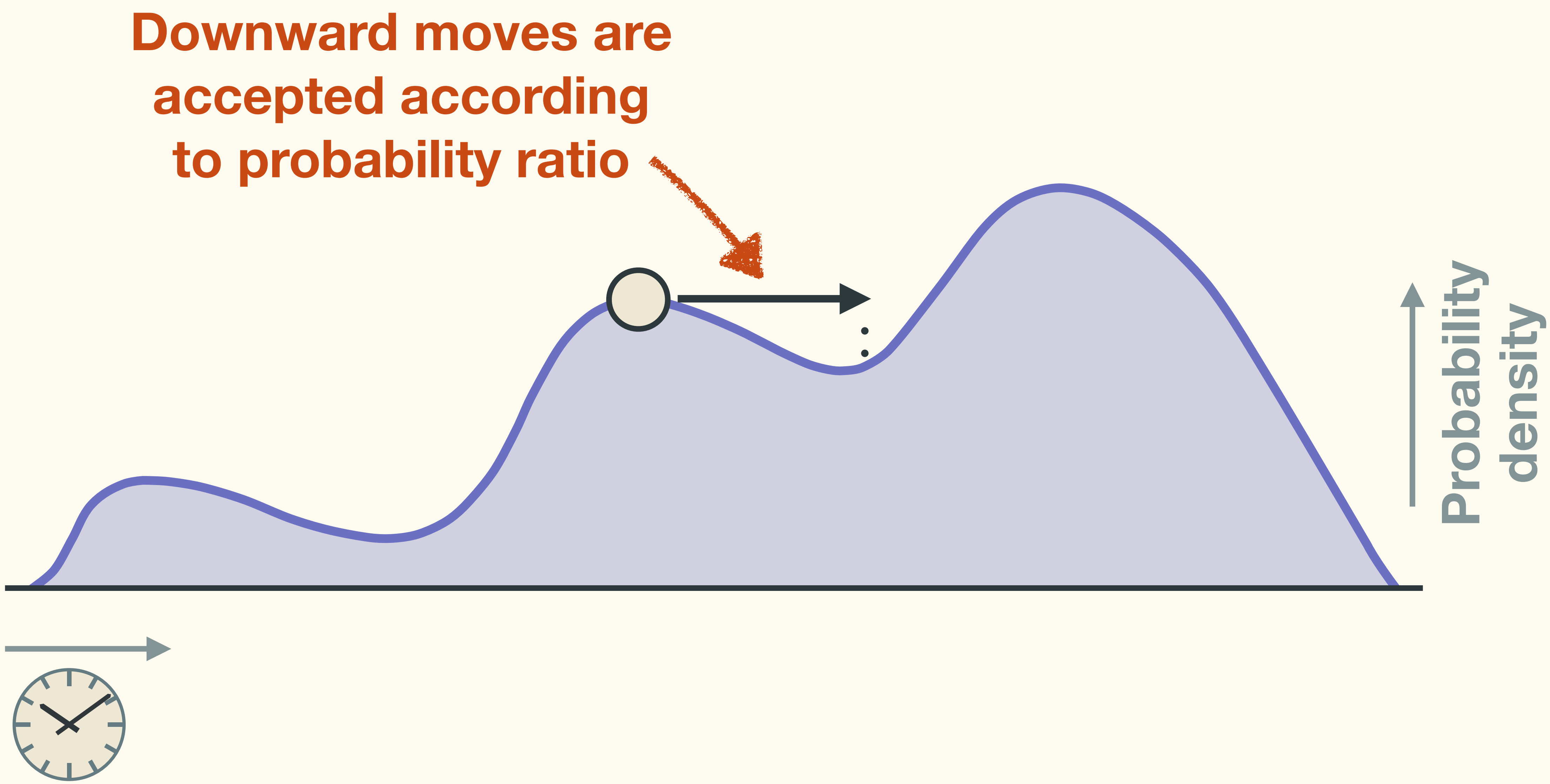
MCMC

Markov-chain Monte Carlo



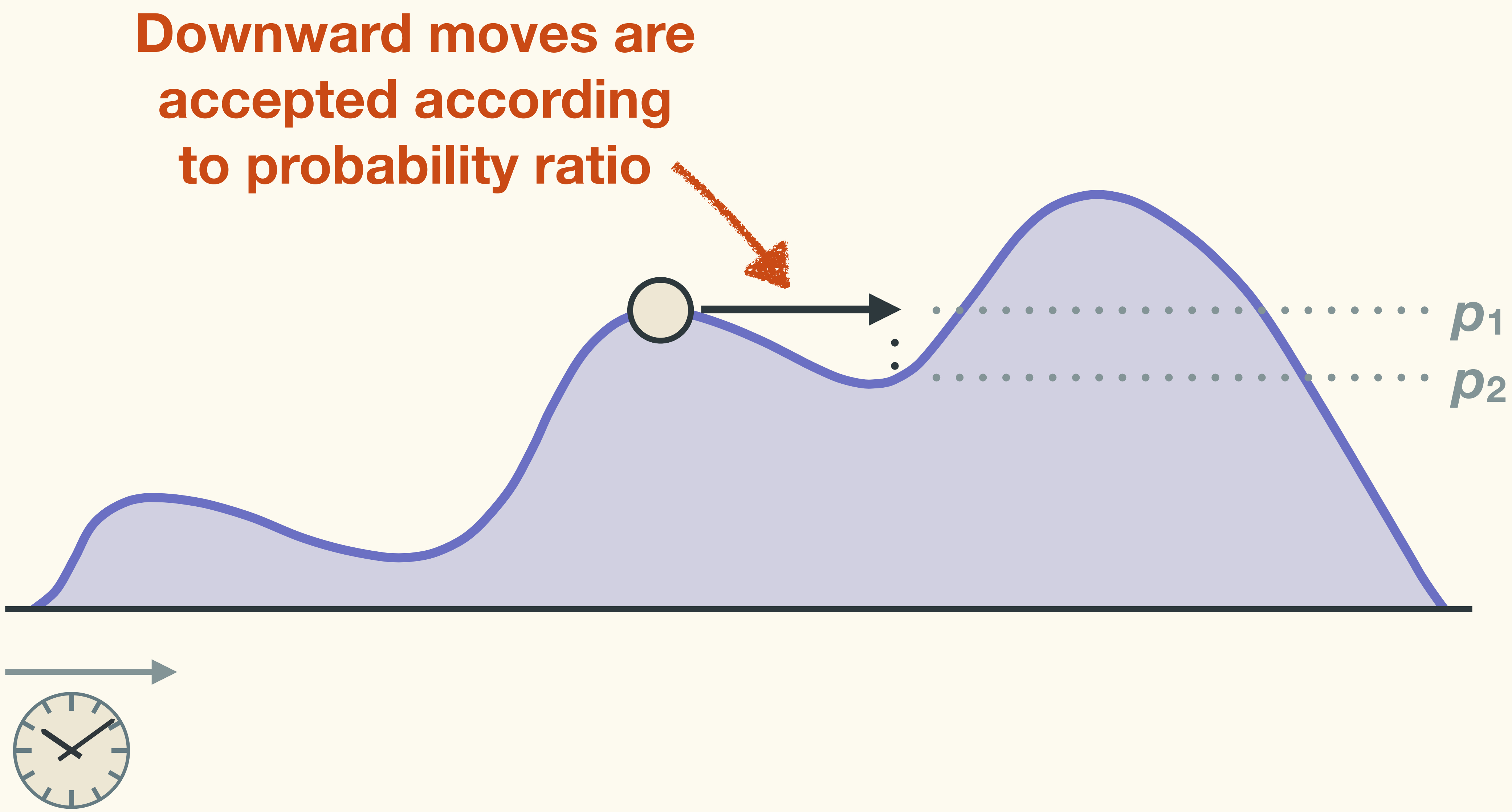
MCMC

Markov-chain Monte Carlo



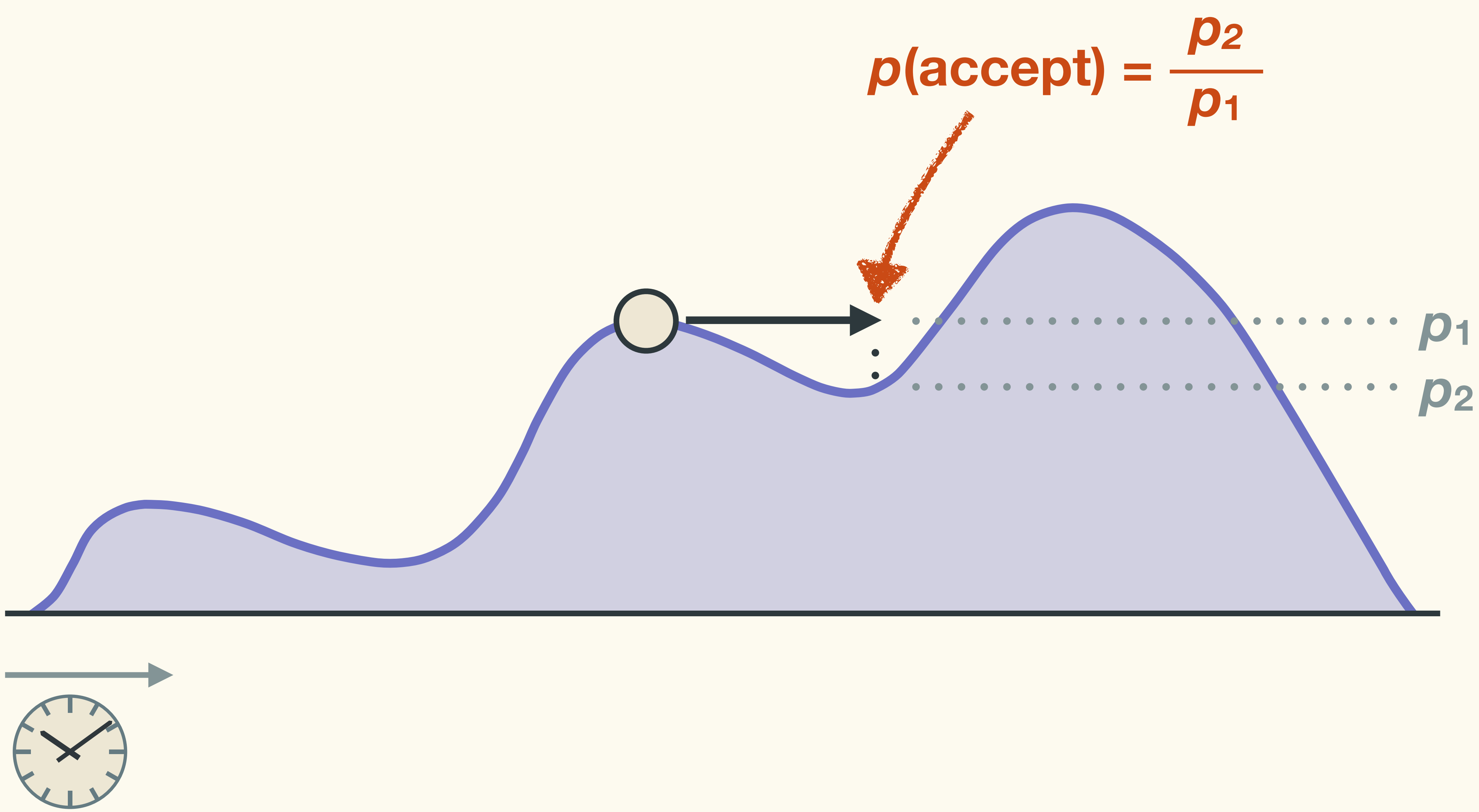
MCMC

Markov-chain Monte Carlo



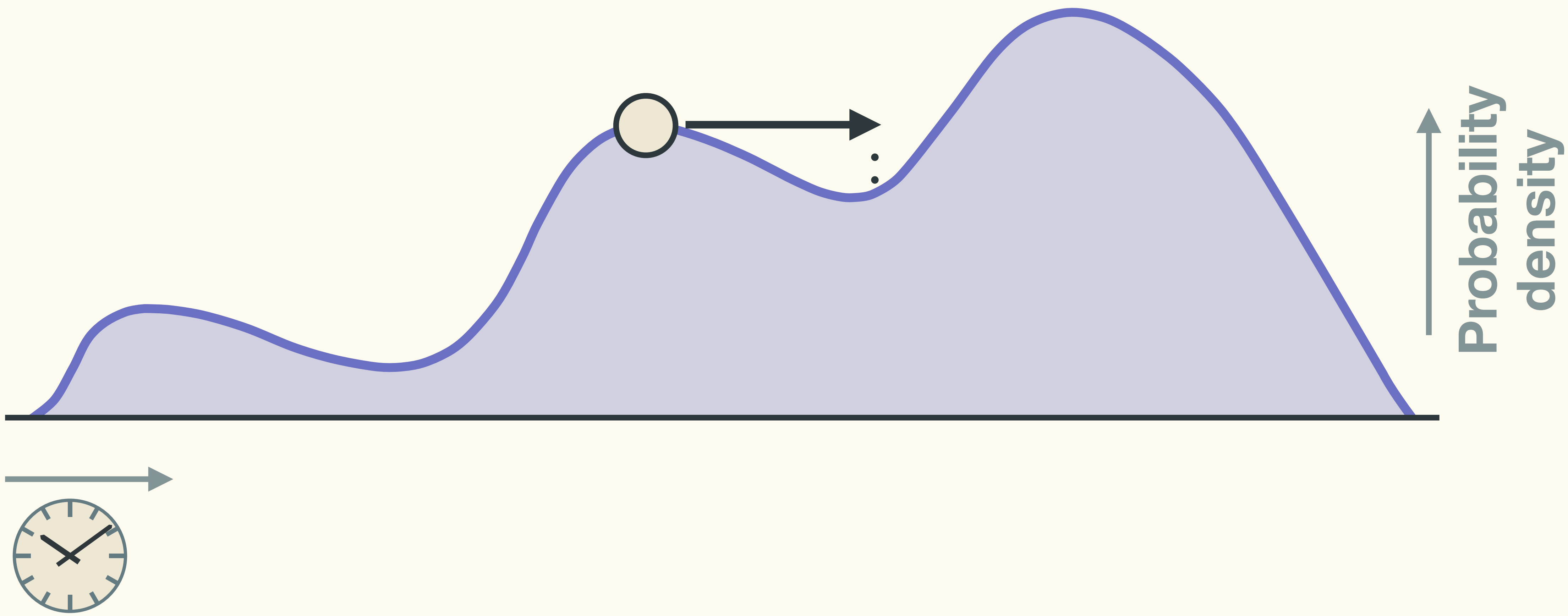
MCMC

Markov-chain Monte Carlo



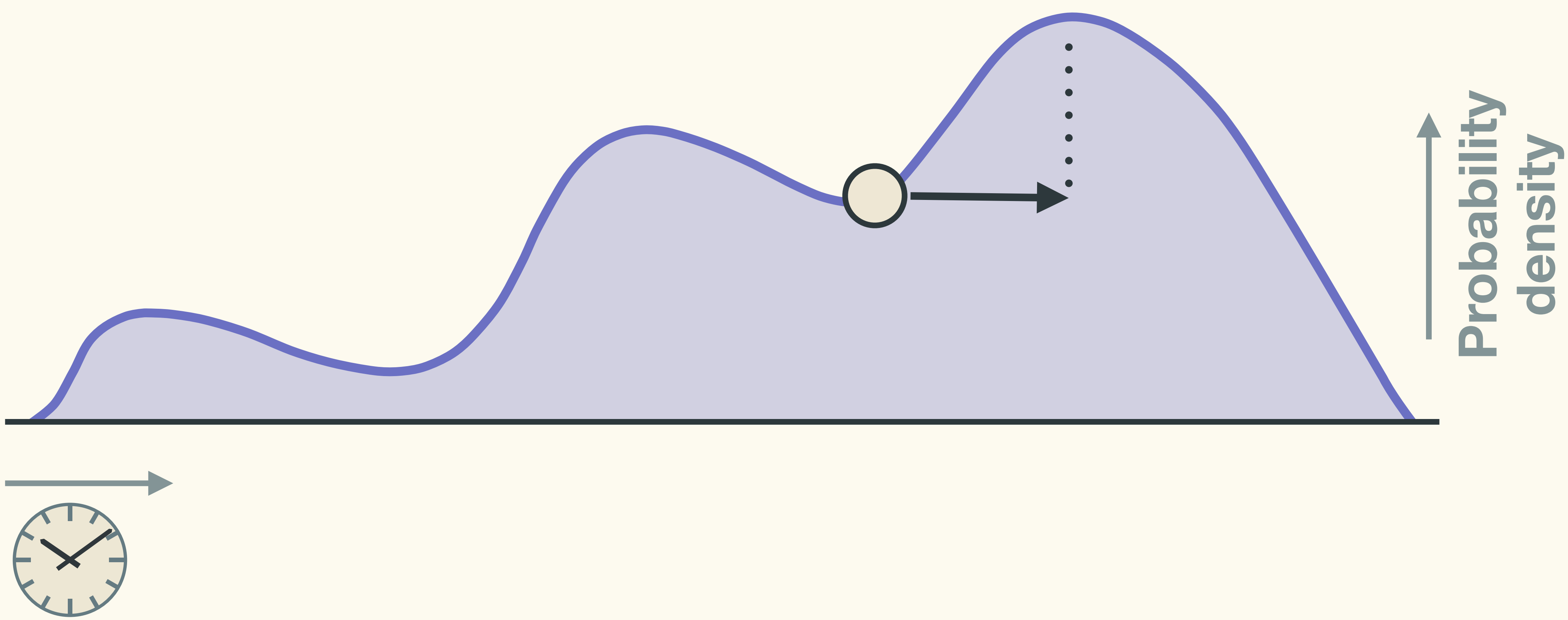
MCMC

Markov-chain Monte Carlo



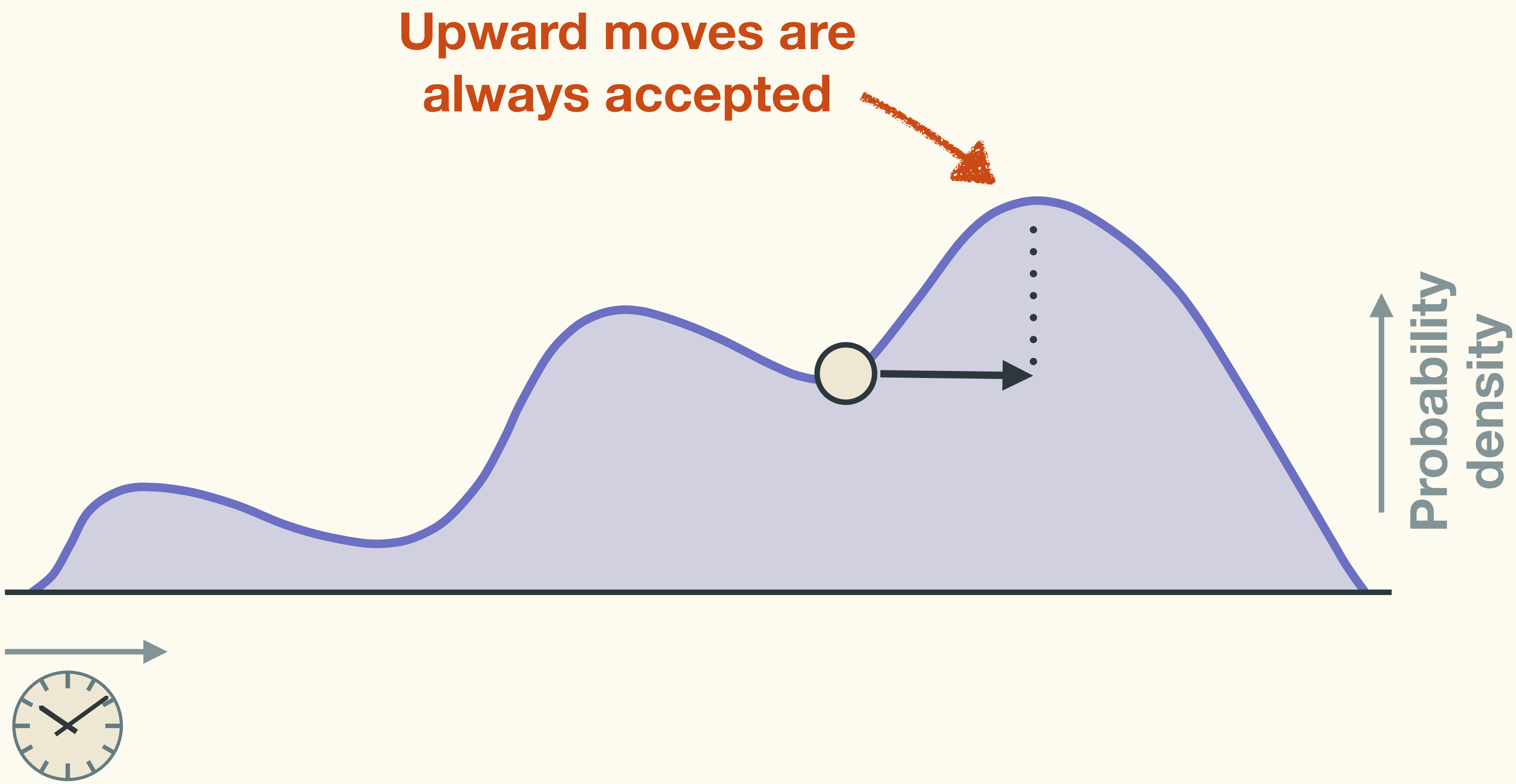
MCMC

Markov-chain Monte Carlo



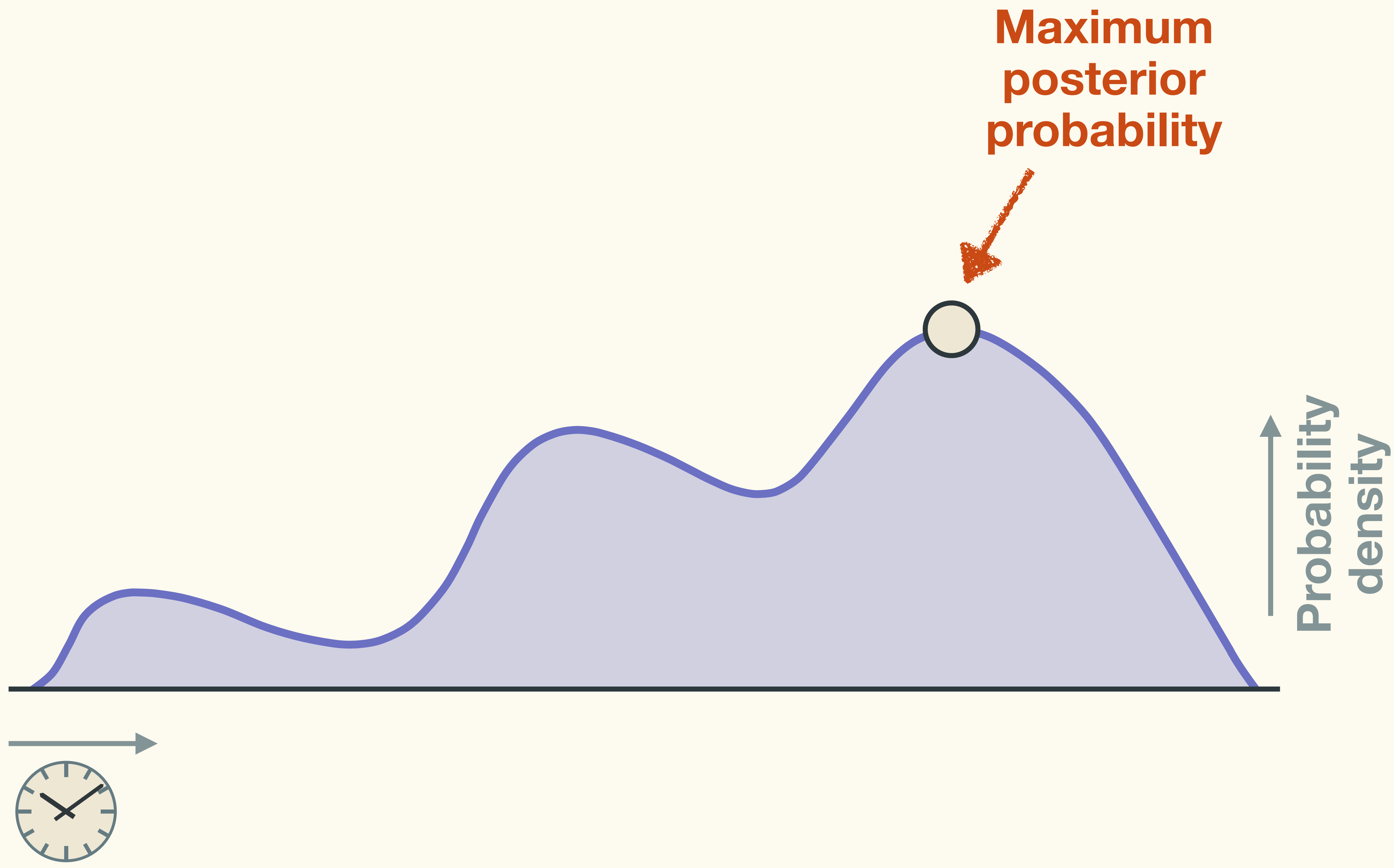
MCMC

Markov-chain Monte Carlo



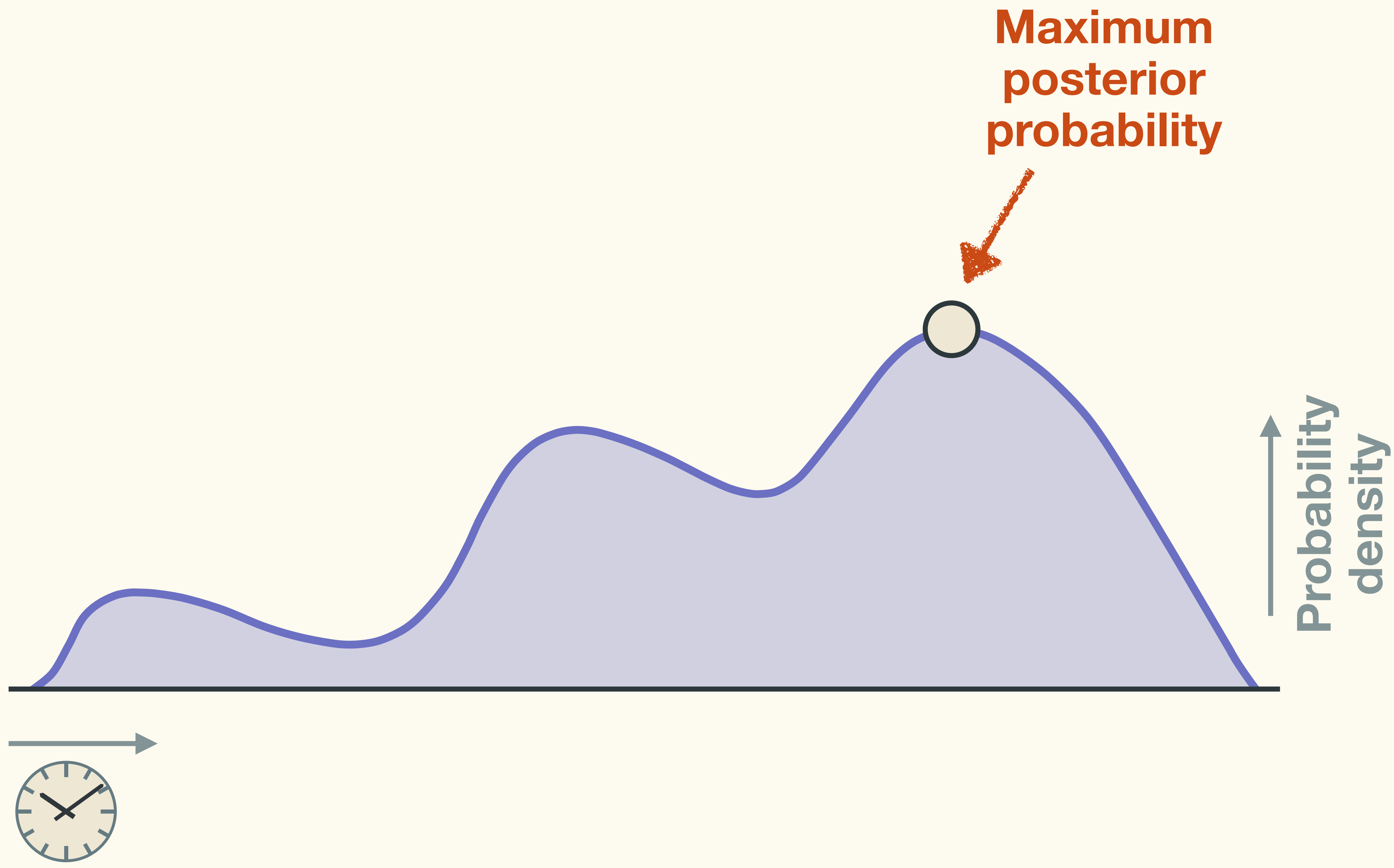
MCMC

Markov-chain Monte Carlo



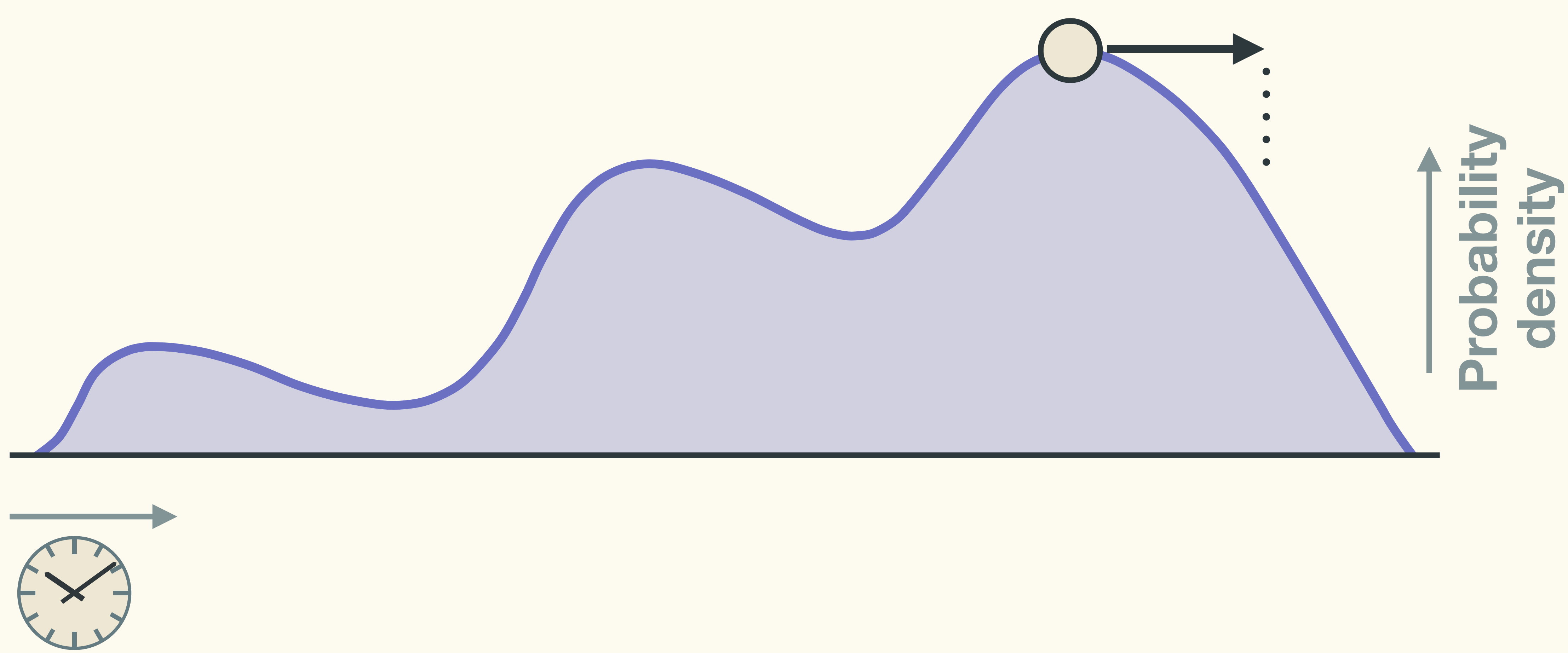
MCMC

Markov-chain Monte Carlo



MCMC

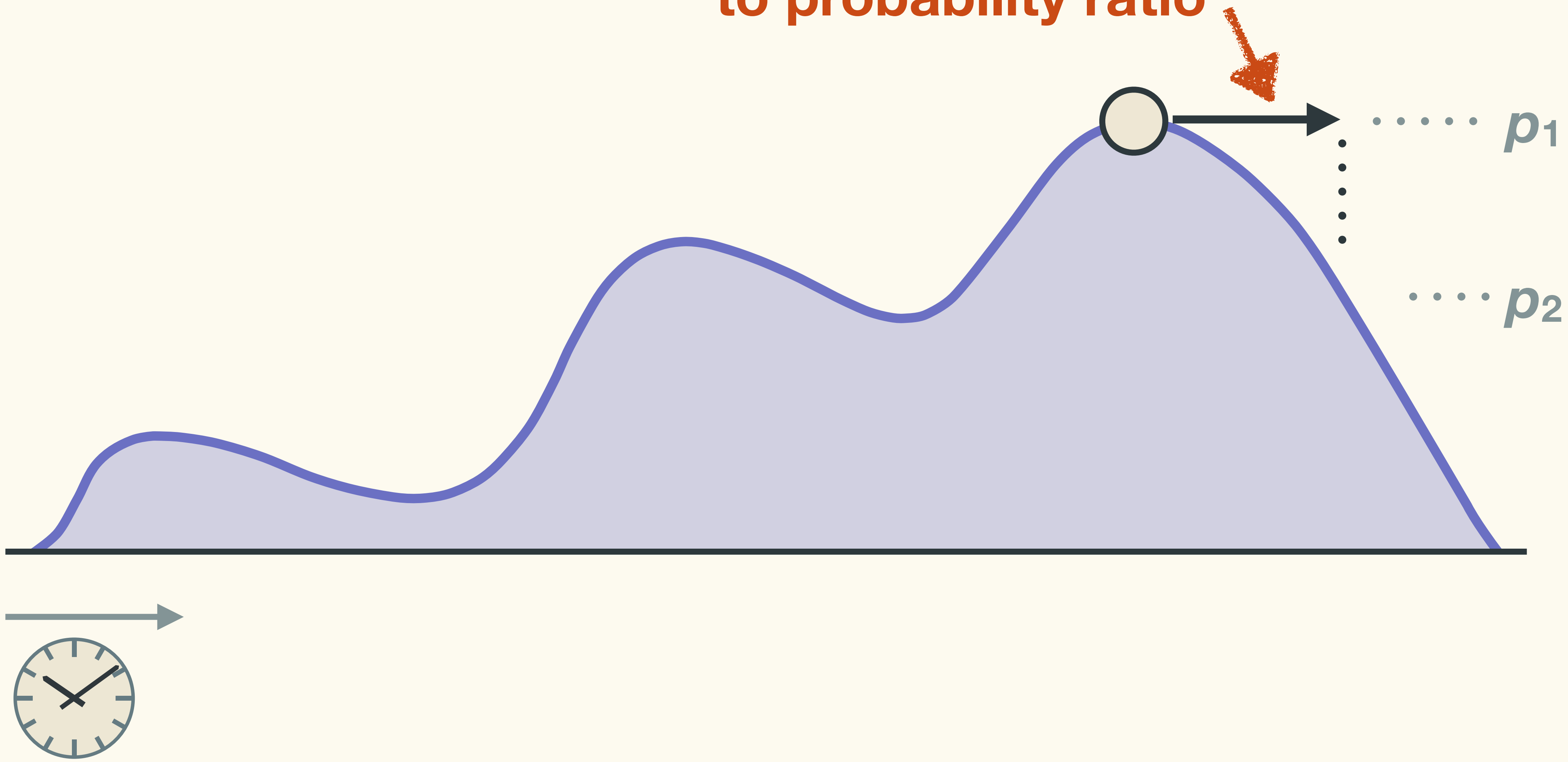
Markov-chain Monte Carlo



MCMC

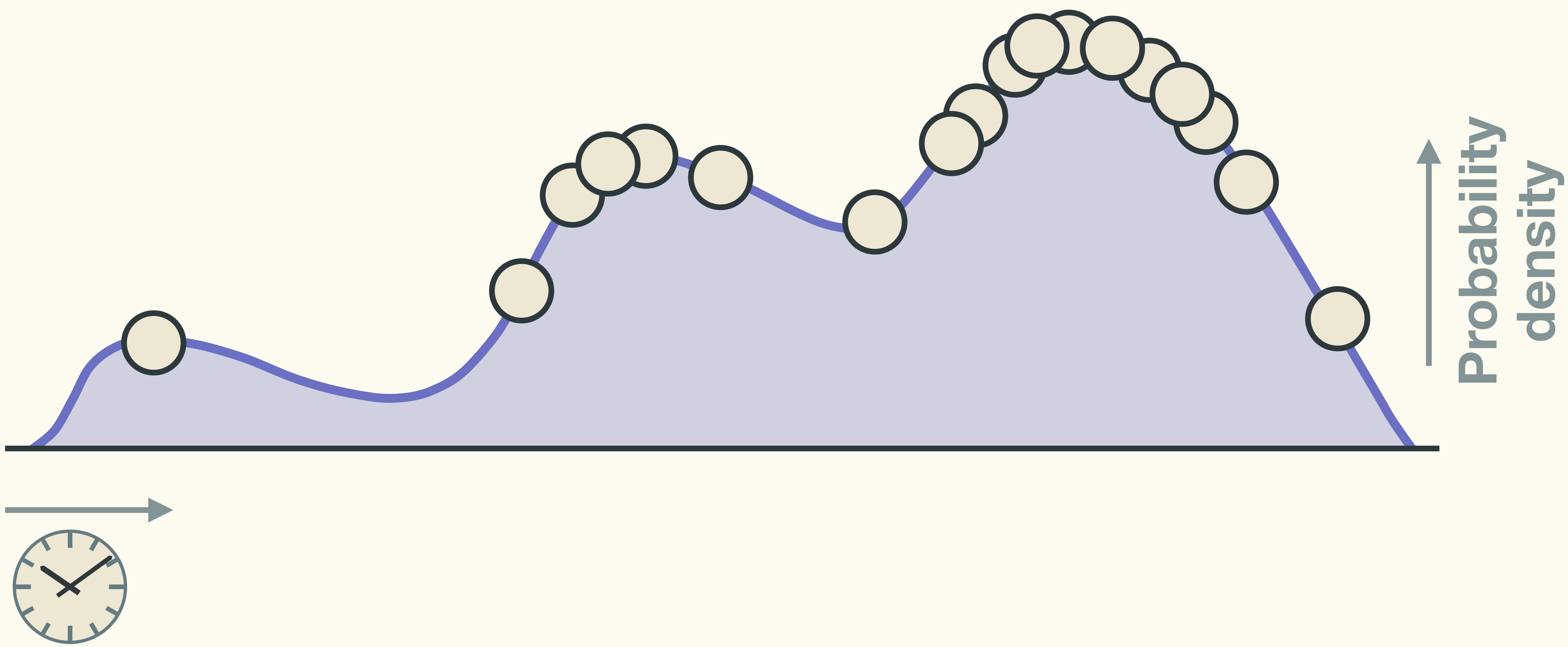
Markov-chain Monte Carlo

Downward moves are accepted according to probability ratio



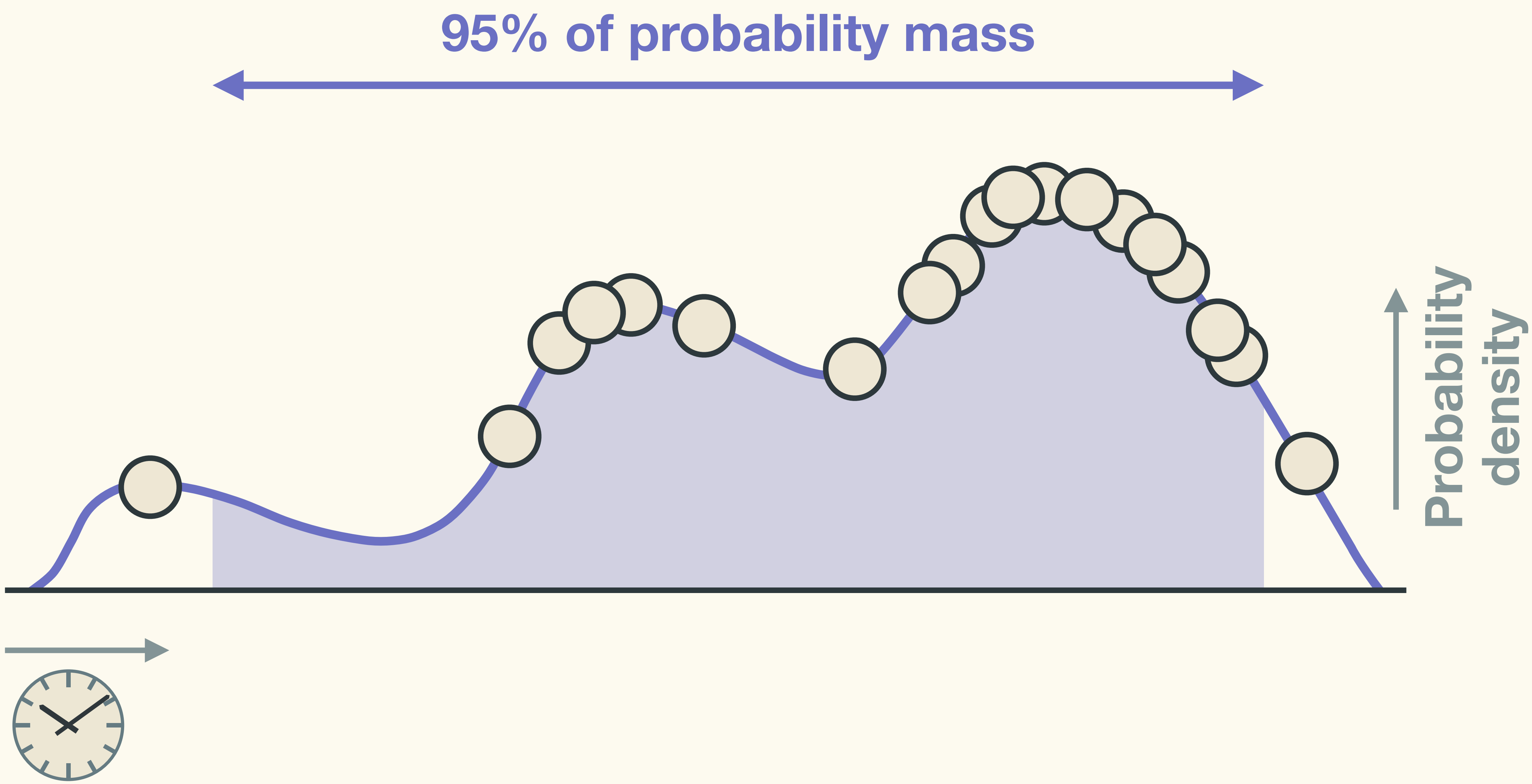
MCMC

Markov-chain Monte Carlo



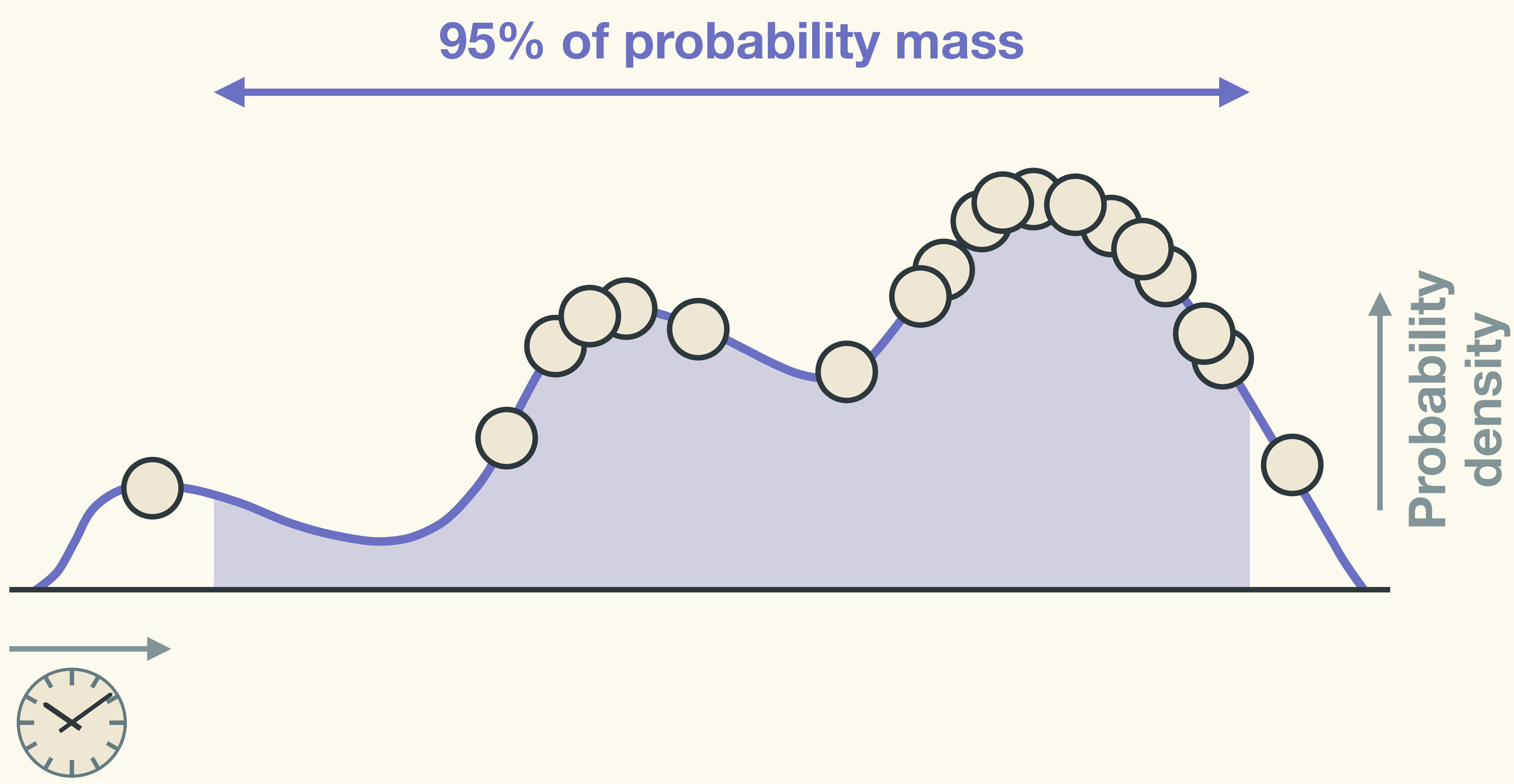
MCMC

Markov-chain Monte Carlo



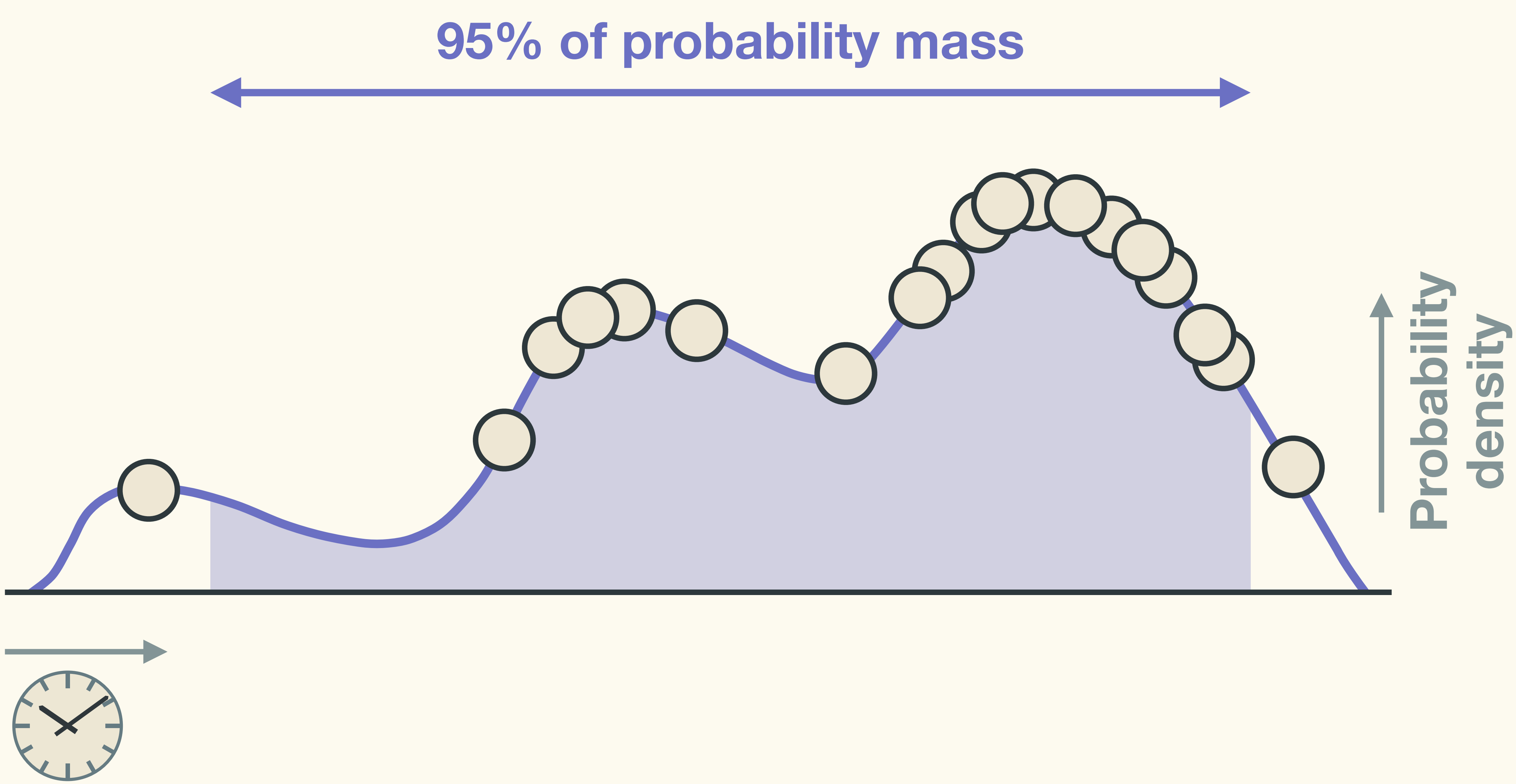
MCMC

Markov-chain Monte Carlo



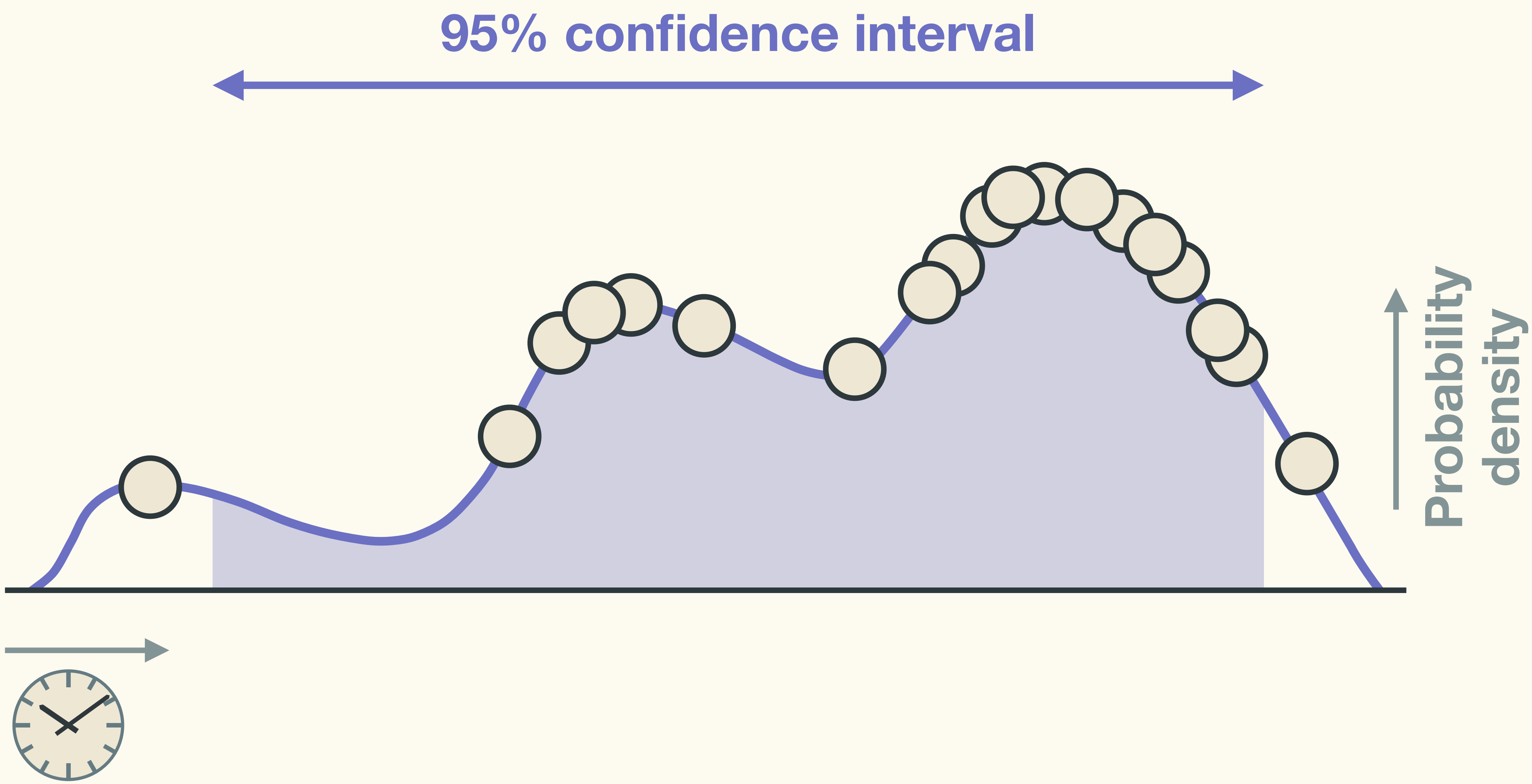
MCMC

Markov-chain Monte Carlo



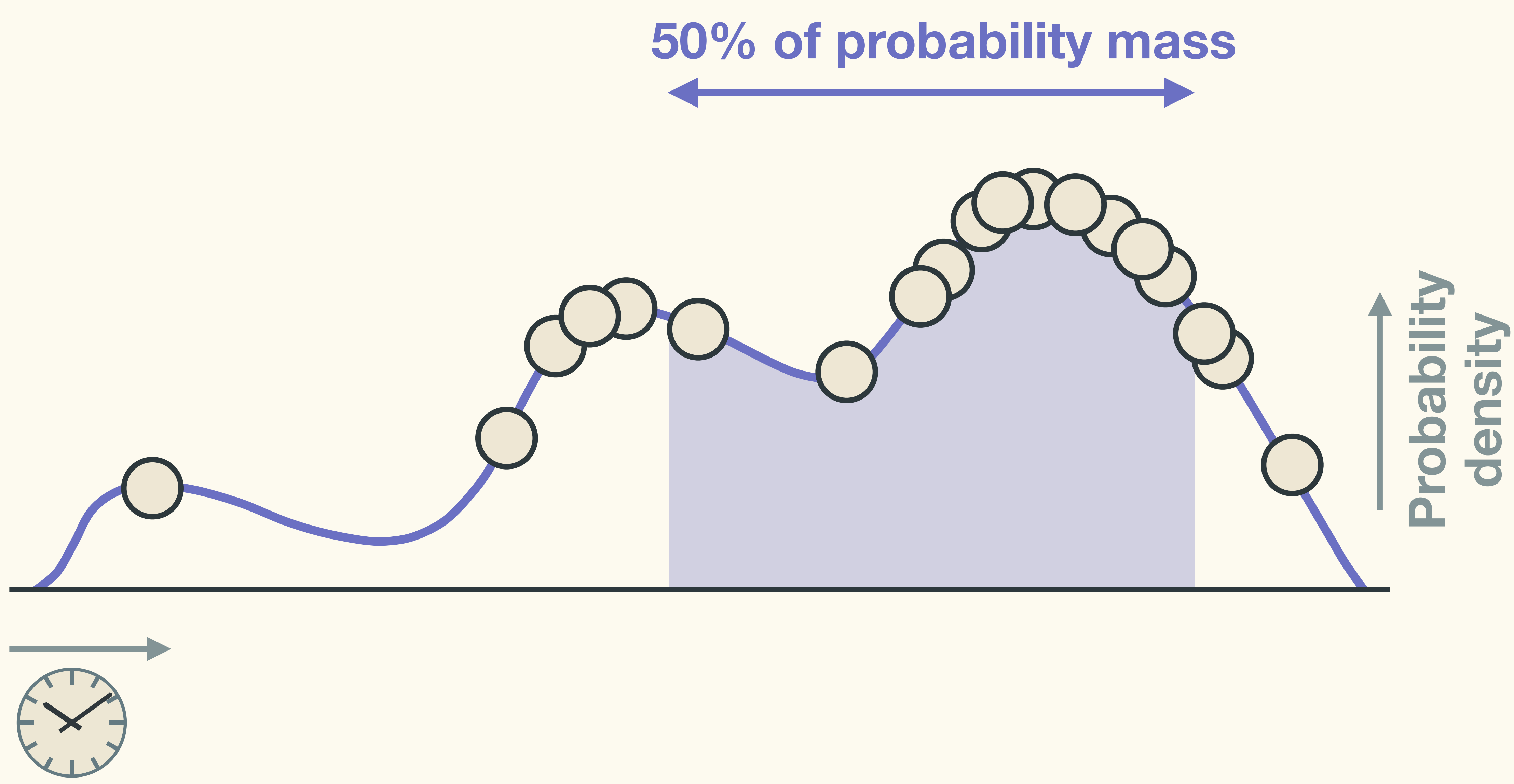
MCMC

Markov-chain Monte Carlo



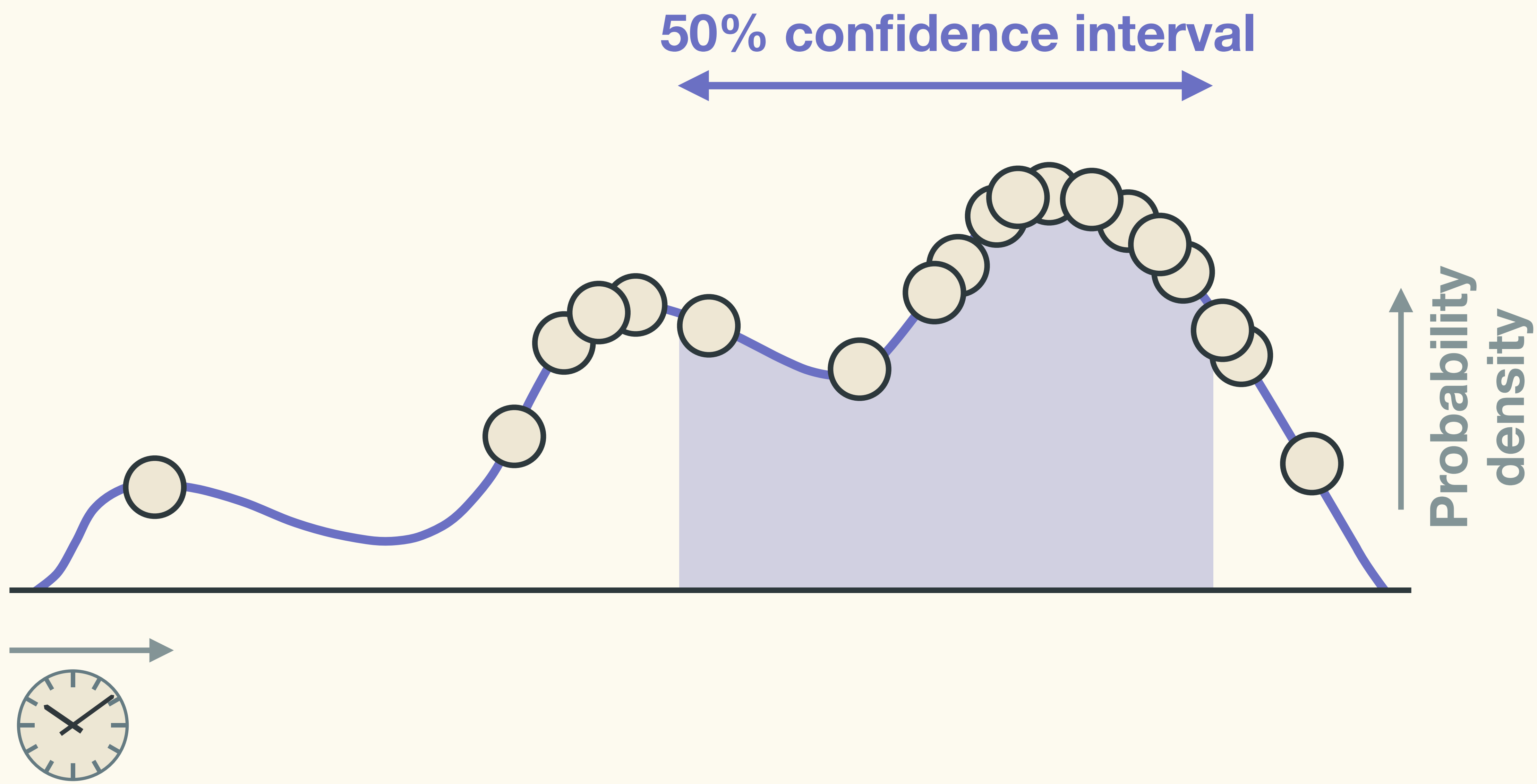
MCMC

Markov-chain Monte Carlo



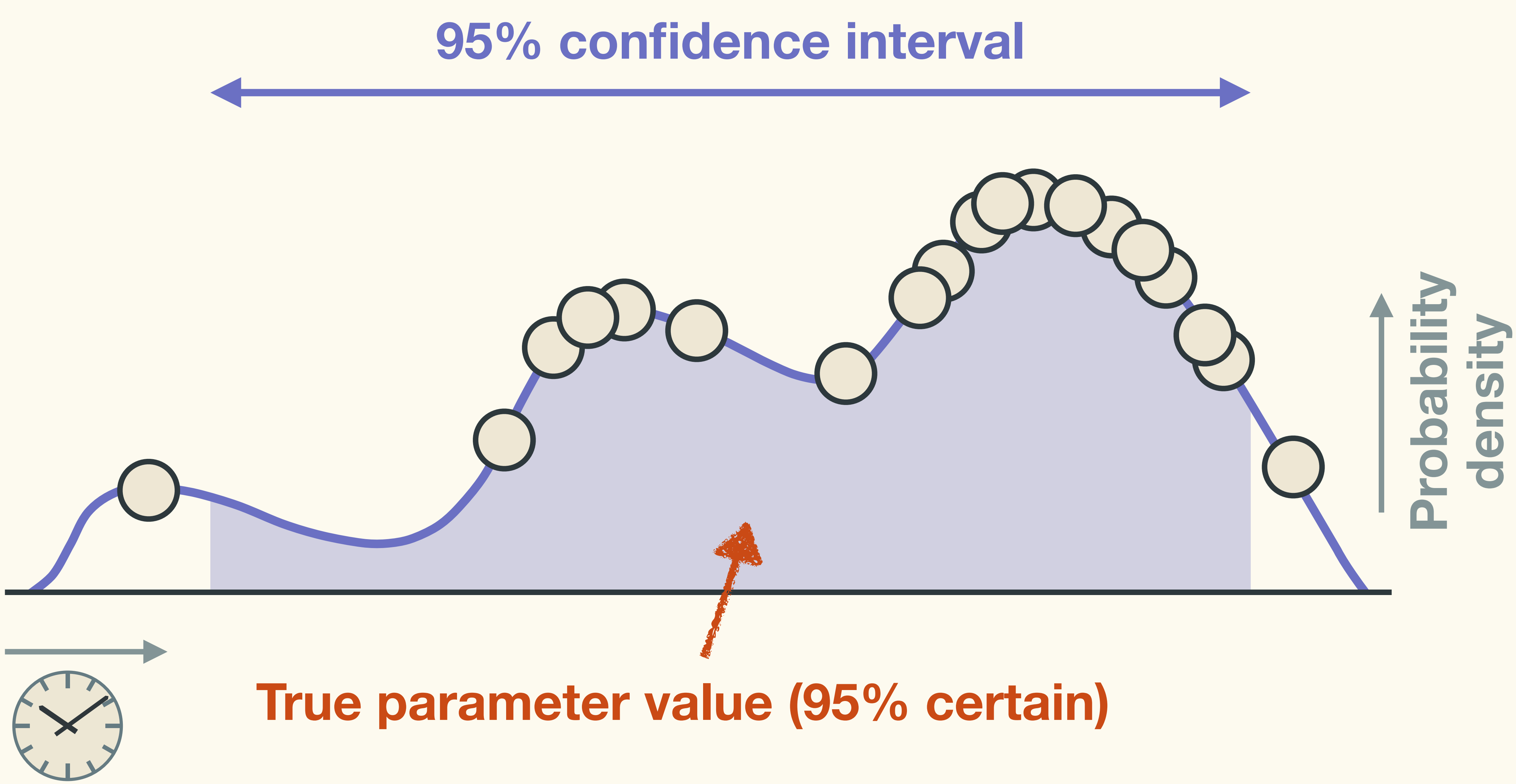
MCMC

Markov-chain Monte Carlo



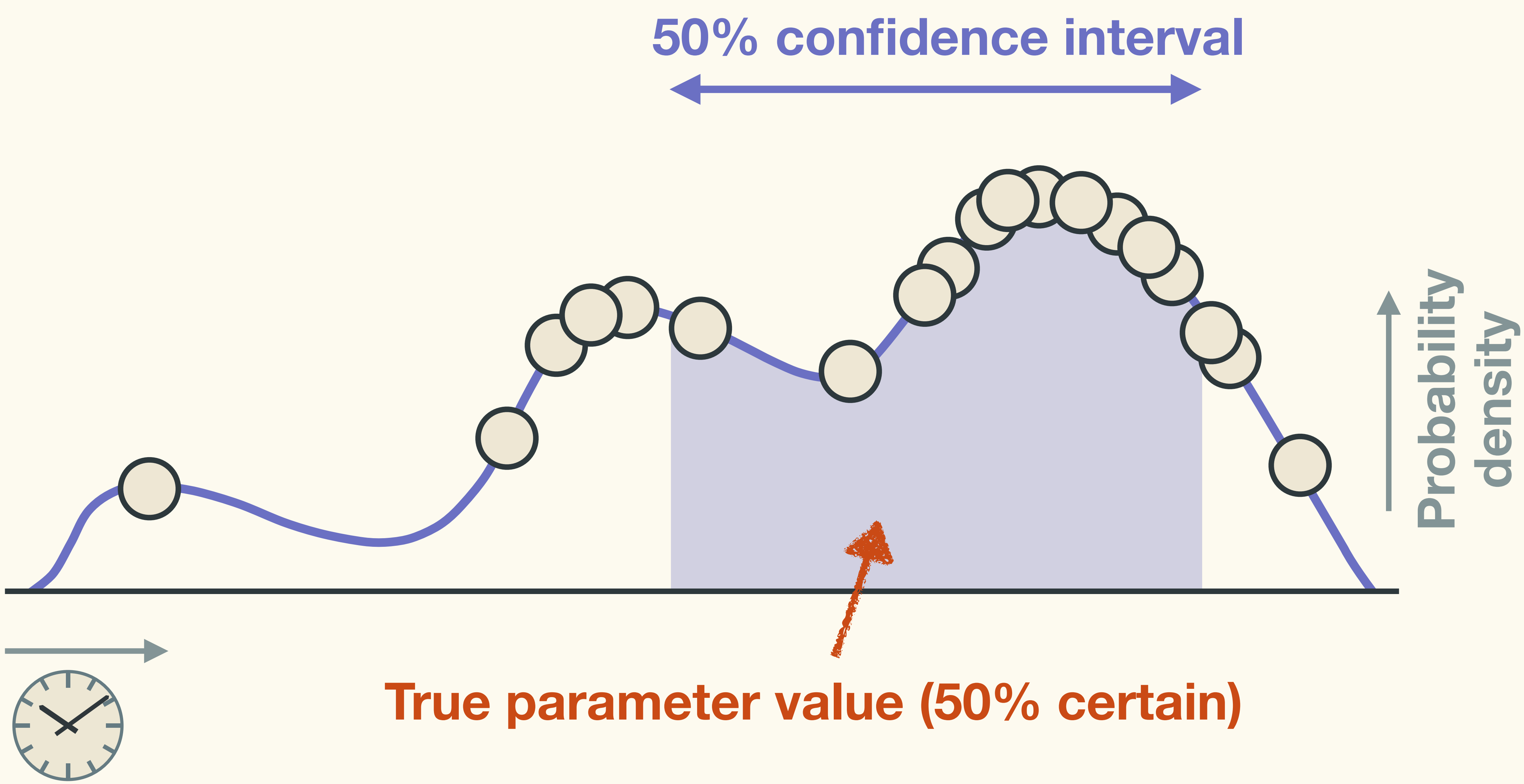
MCMC

Markov-chain Monte Carlo



MCMC

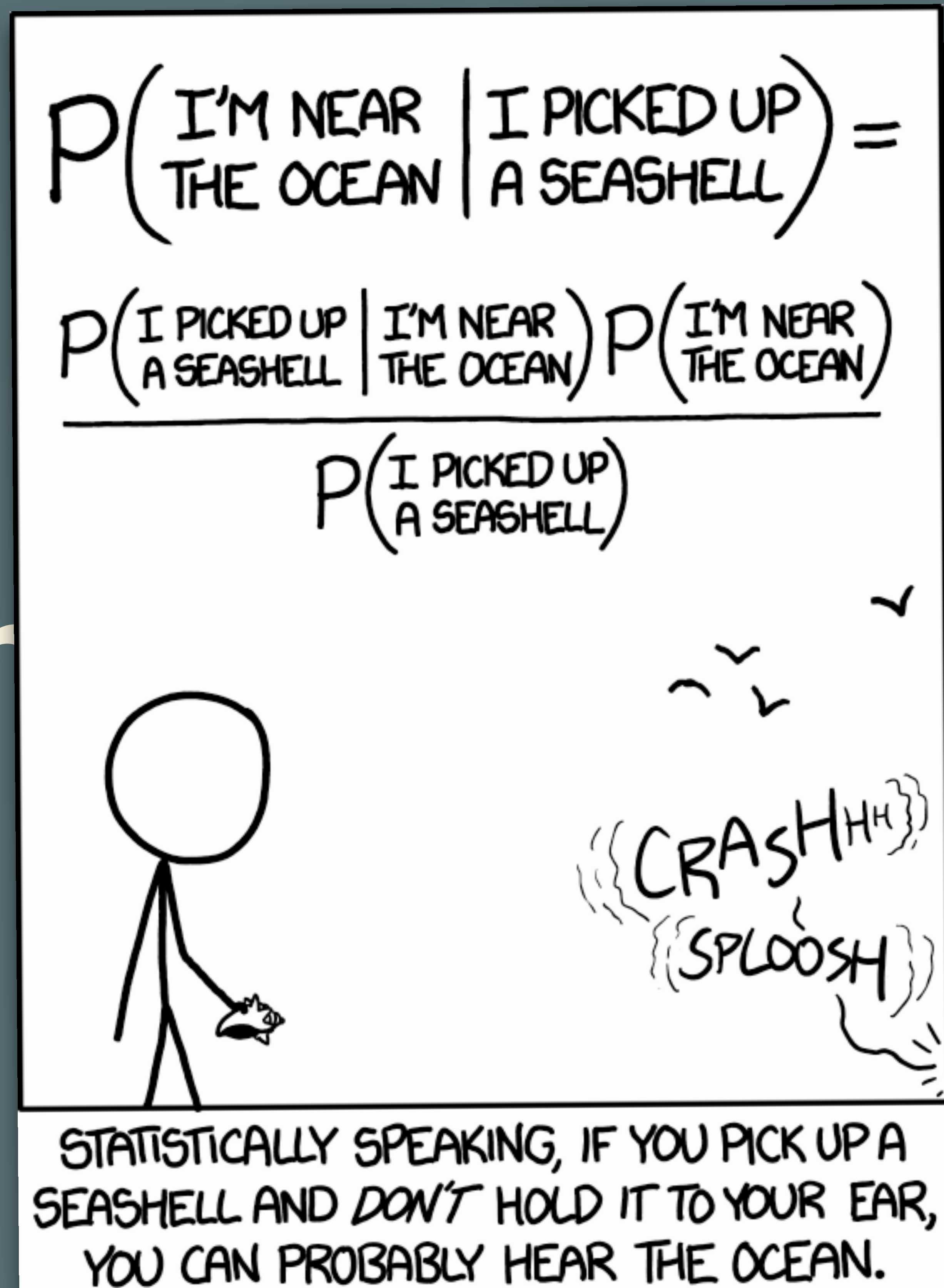
Markov-chain Monte Carlo



MCMC Robot

Thanks

Thar


P(I'M NEAR THE OCEAN | I PICKED UP A SEASHELL) =
P(I PICKED UP A SEASHELL | I'M NEAR THE OCEAN) P(I'M NEAR THE OCEAN)
P(I PICKED UP A SEASHELL)
CRASHHHH
SPLOOSH
STATISTICALLY SPEAKING, IF YOU PICK UP A SEASHELL AND DON'T HOLD IT TO YOUR EAR, YOU CAN PROBABLY HEAR THE OCEAN.