Stefano Sarao Mannelli
in collaboration with
Giulio Biroli, Chiara Cammarota,
Florent Krzakala and Lenka Zdeborová
NeurIPS 2019

Who Is Afraid of Big Bad Minima?

IPhT
SACLAY

Main Message

Question:
when does gradient flow dynamics find “good” minima
in high-dimensional non-convex problems?

Existing works:
assume that spurious minima have to disappear.

We show gradient flow can work even when spurious minima are present!

Ingredients

spiked matrix-tensor model
with gradient flow (GF)

$$\arg\min_x ||xx^T - Y||_2^2 + ||x^{\otimes p} - T||_2^2$$

with

$$x, x^* \in \mathbb{S}^{N-1} \qquad Y = x^*(x^*)^T + \text{noise} \qquad T = (x^*)^{\otimes p} + \text{noise}$$

and $N \gg 1$

Ingredients

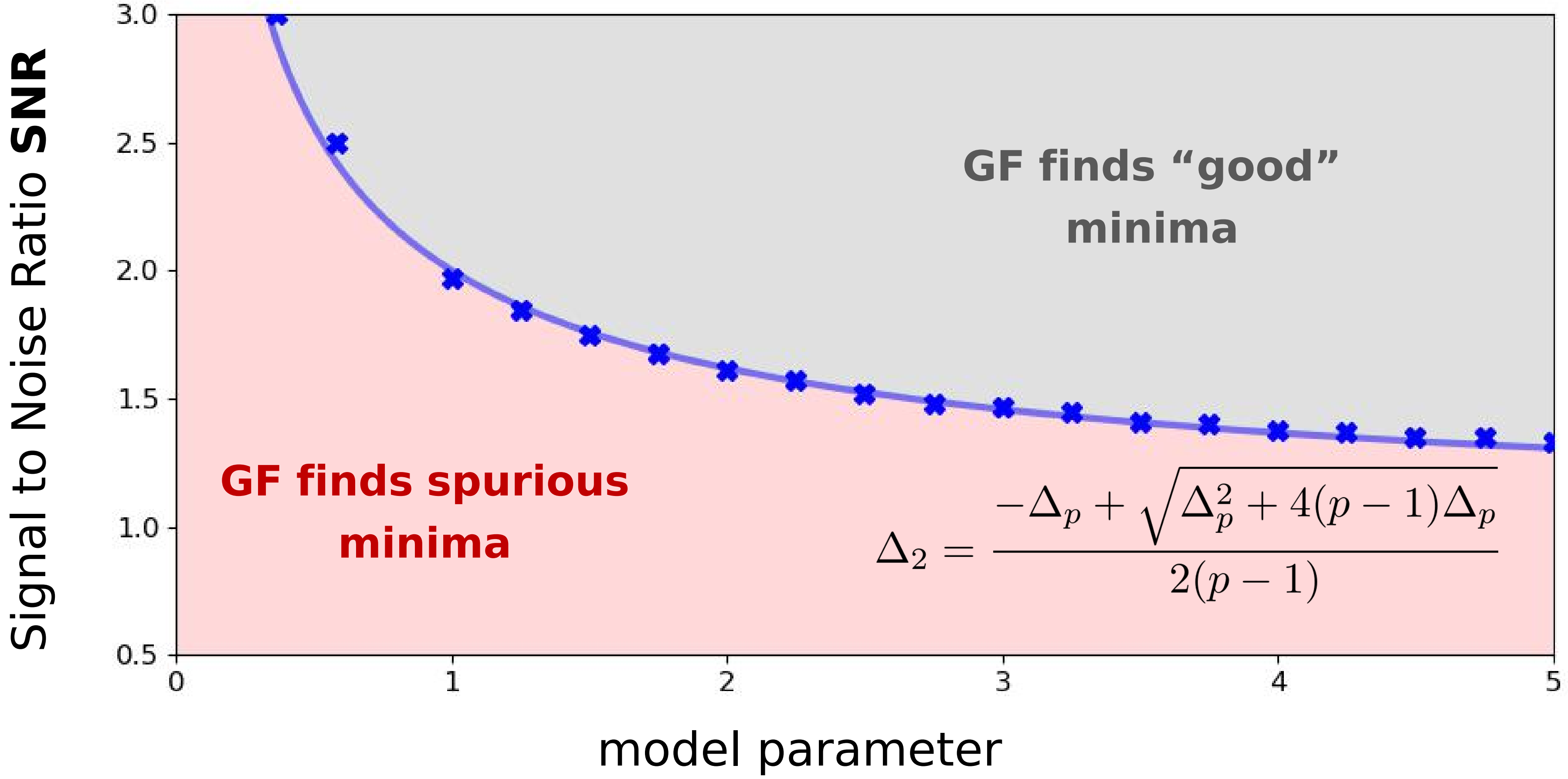
spiked matrix-tensor model
with gradient flow (GF)

Kac-Rice formula
to mathematically
describe the geometry
[Kac 1943, Rice 1944]

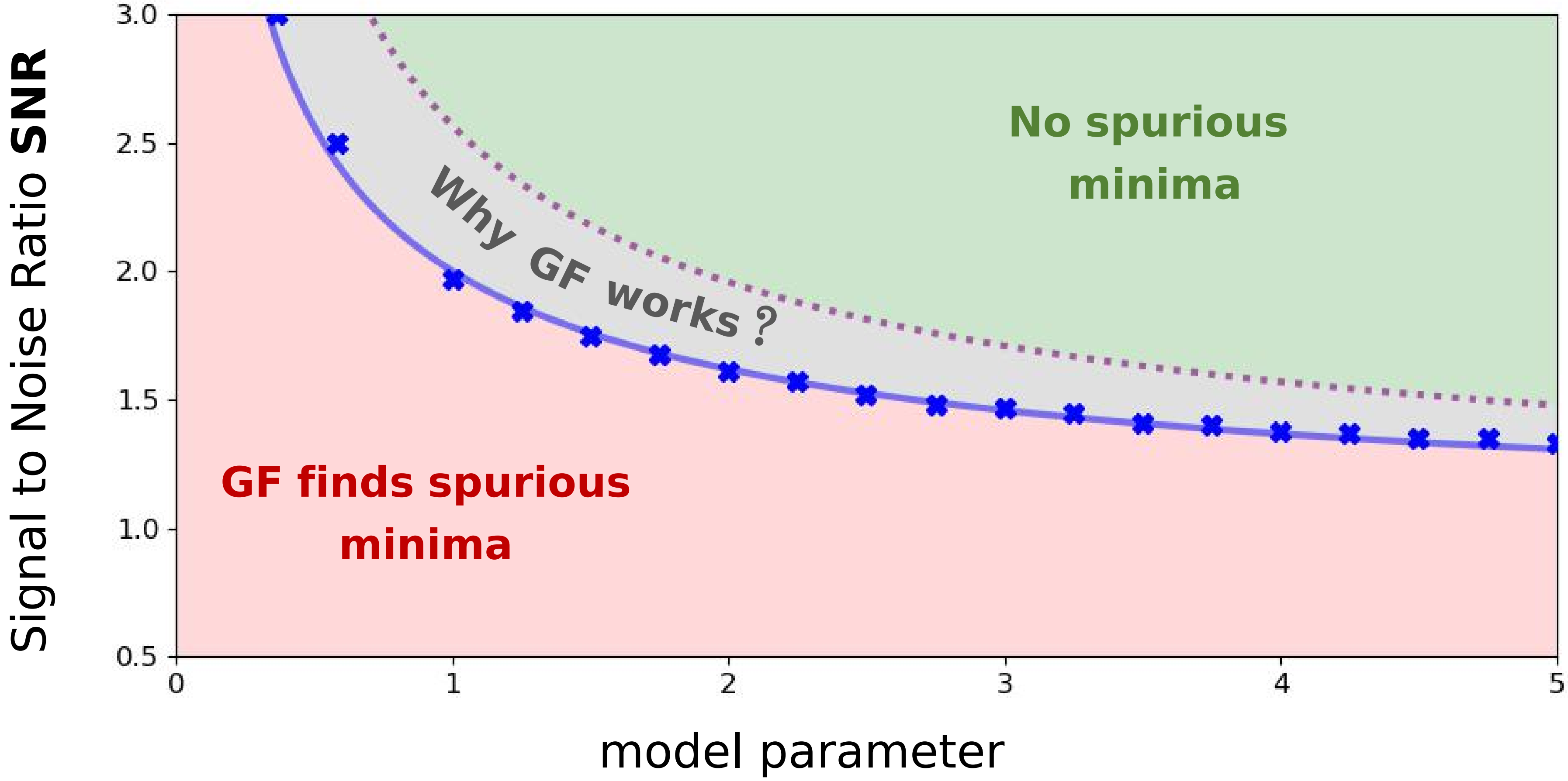
Martin-Siggia-Rose
formalism to characterize
GF dynamics
[Crisanti, Horner,
Sommers 1993]

Gradient flow

1st down-to-the-constant analysis of gradient flow performance

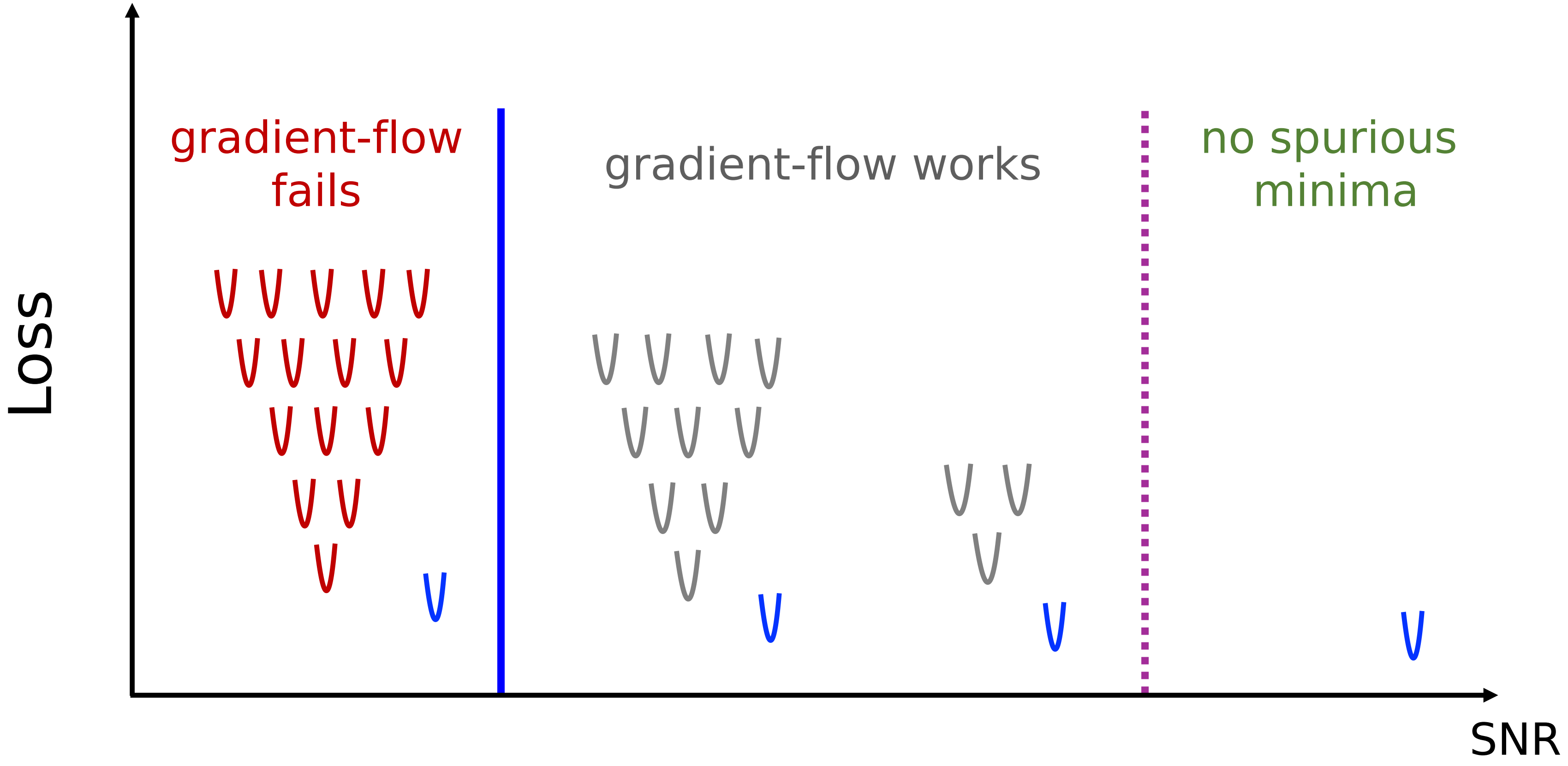


Geometry

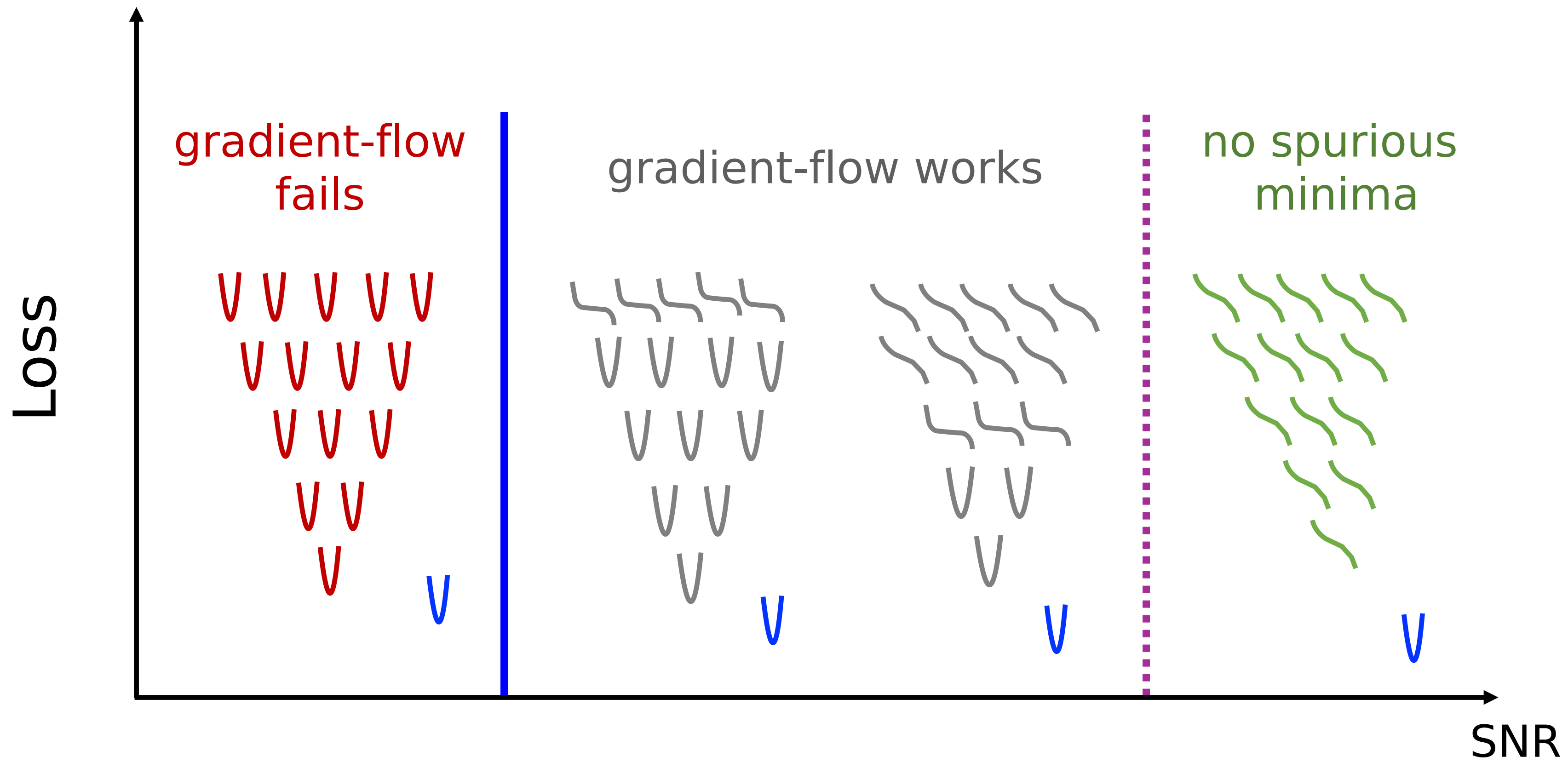


No spurious minima
Why GF works ?
GF finds spurious minima
3.0
2.5
2.0
1.5
1.0
0.5
0
1
2
3
4
5
Signal to Noise Ratio **SNR**
model parameter

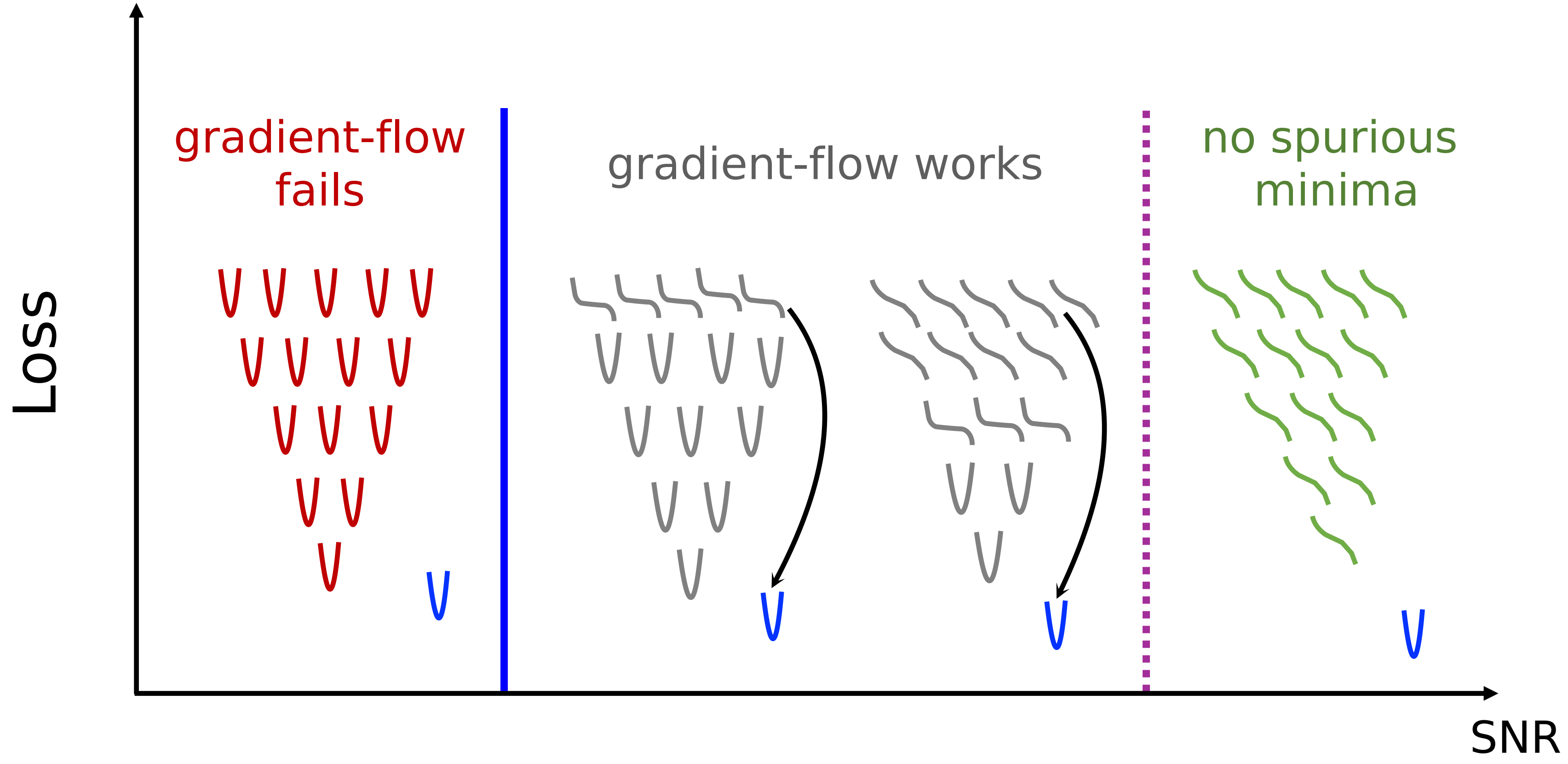
What's going on?!



What's going on?!



What's going on?!



Thank you! :D

come see my poster #127