

Réduire le coût énergétique de l'IA grâce aux nanodispositifs

Alice Mizrahi

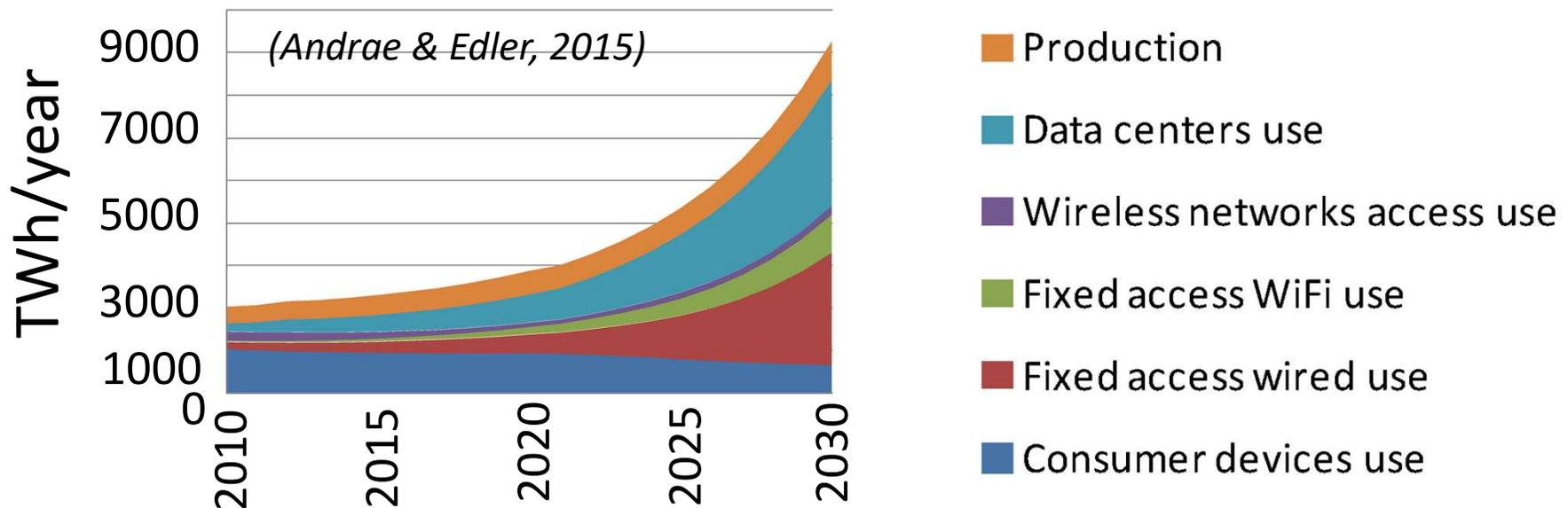
Unité Mixte de Physique CNRS/Thales



THALES



Les avancées d'intelligence artificielle : un danger pour l'environnement ?



Consommation du secteur des technologies de l'information
↔ consommation électrique de l'Inde

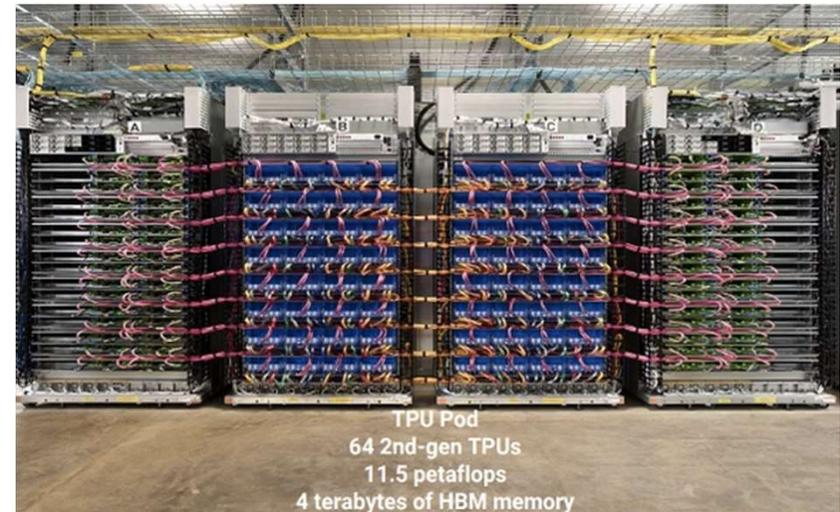
Fort coût énergétique: impact objets intelligents



Données

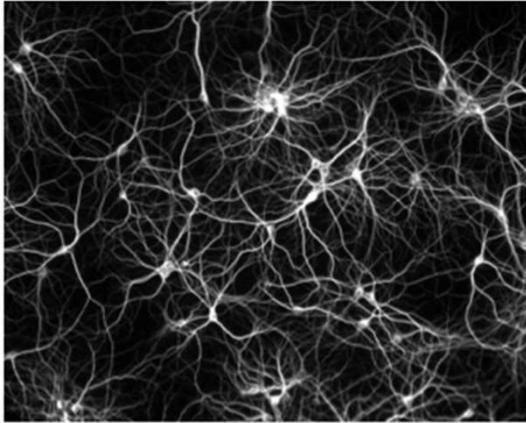


Résultats



- x** Pas d'apprentissage dans l'objet
- x** Objet pas autonome
- x** Problèmes sécurité, vie privée

Pistes : différences cerveau/ordinateurs



20 W

Mémoire (synapses) et calcul
(neurones) interconnectés

Utilise la dynamique des neurones

Imprécis, aléatoire



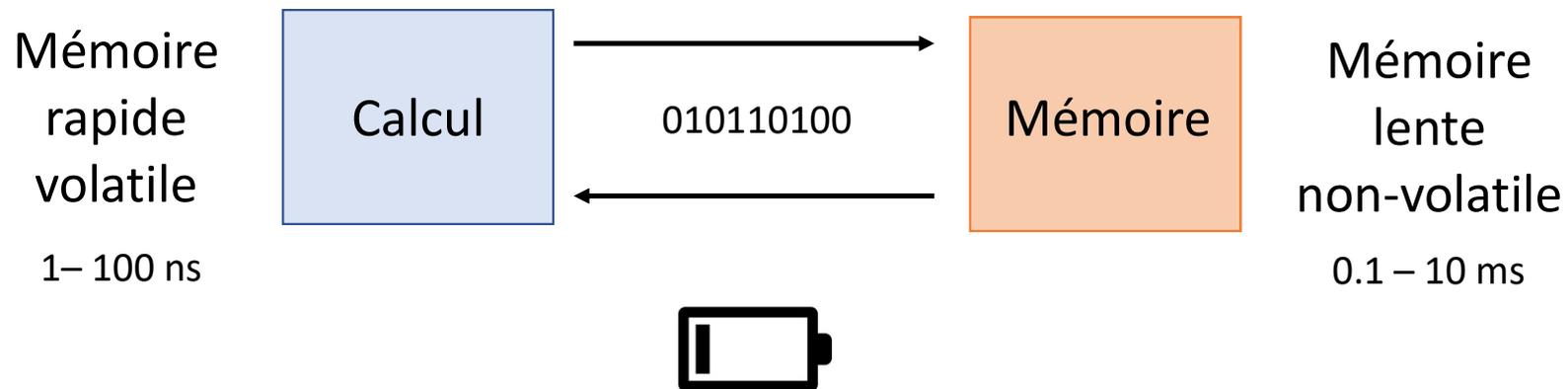
MW

Mémoire et calcul séparés

Logique Booléenne, états statiques

Très précis, déterministe

Communiquer entre mémoire et calcul consomme énormément



Energie communication mémoire:
60x calcul (sur puce)
3500x calcul (hors puce)

Pedram et al., IEEE Design & Test, 2016

Des nano-mémoires non volatiles et rapides

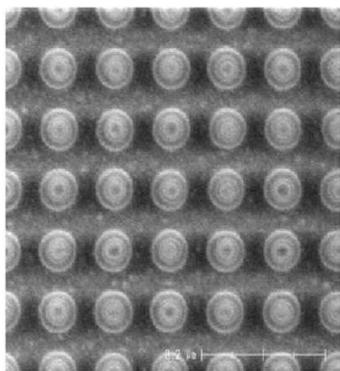
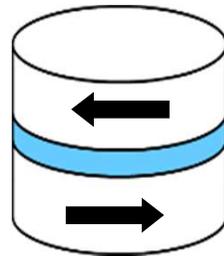
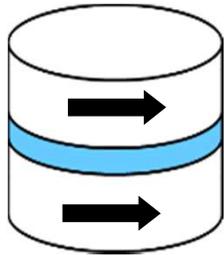
« Jonction tunnel magnétique »

Résistance
basse (« 0 »)

Résistance
haute (« 1 »)

Aimant

Aimant



Chung et al., IEDM 2016

« Memristor »

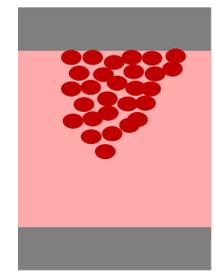
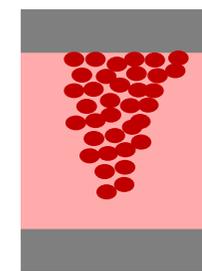
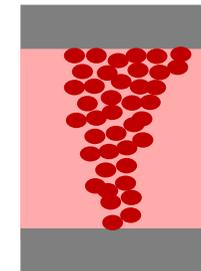
Résistance
basse

Résistance
haute

Métal

Isolant

Métal



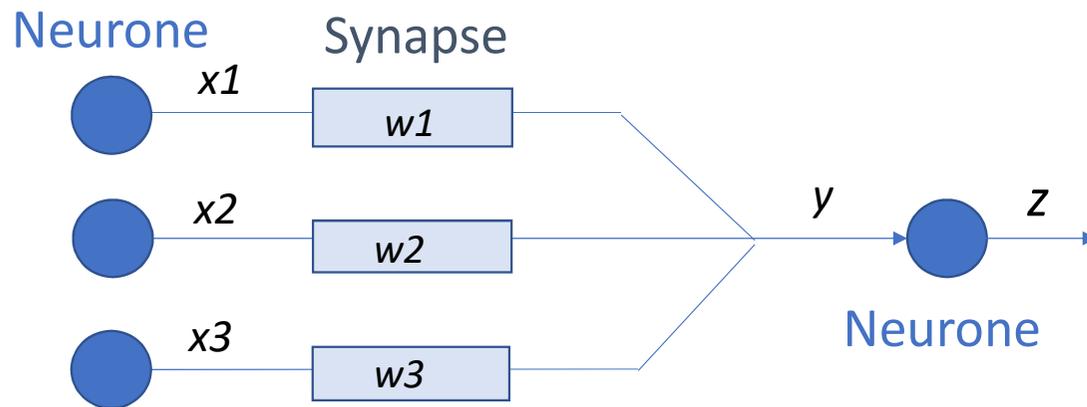
Etat analogique

Lecture et écriture électrique

Taille 10 – 100 nm

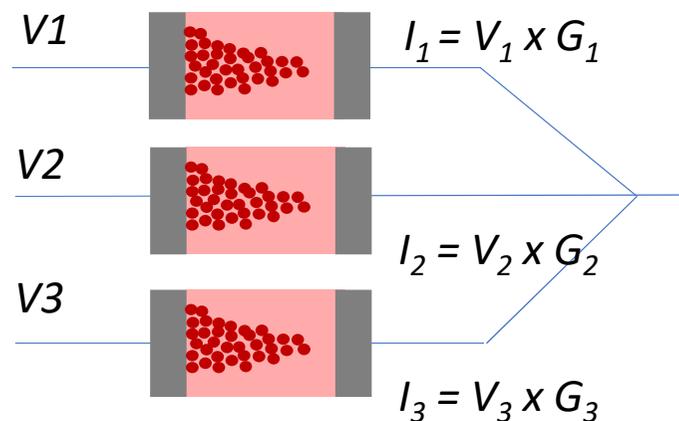
Compatible avec électronique conventionnelle

Réseau de neurones avec nano-mémoires comme synapses



$$y = \sum_i x_i w_i = x_1 w_1 + x_2 w_1 + x_3 w_3$$

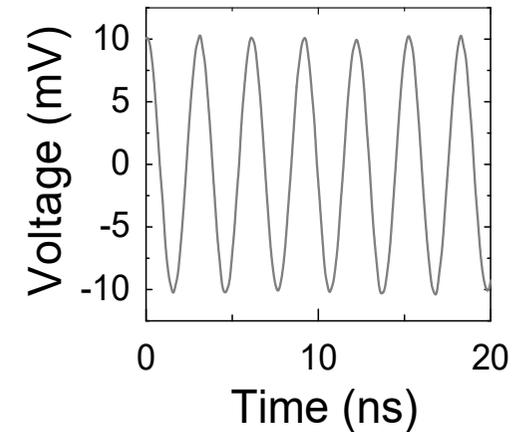
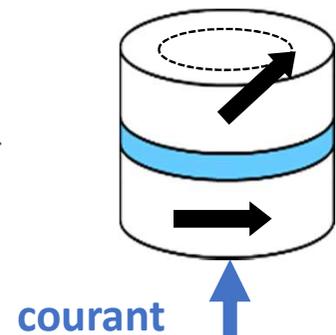
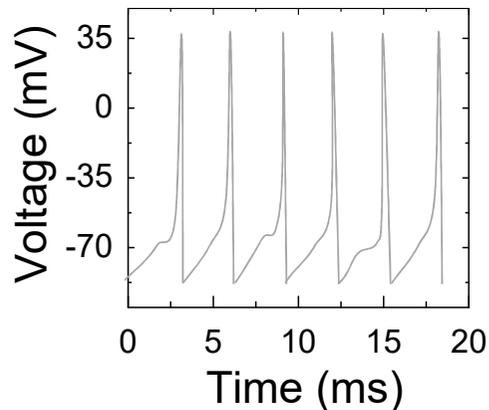
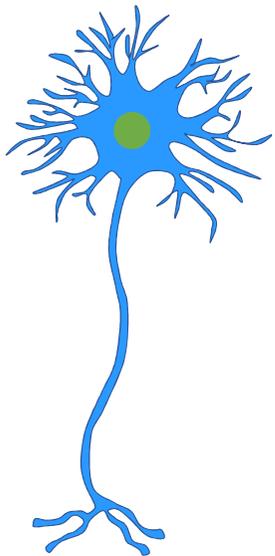
$$z = f(y) \quad \text{Non-linéaire}$$



$$I = I_1 + I_2 + I_3 = \sum_i V_i G_i$$

I courant
V tension
G conductance

Utiliser la dynamique non-linéaire des nano-oscillateurs pour faire des neurones



- Reconnaissance de voyelles, chiffres parlés
- Consomme moins qu'un oscillateur en électronique conventionnelle



*Equipe de Julie Grollier à l'UMR CNRS/Thales
Torrejon et al., Nature 2017 ; Romera et al., Nature 2018*

Accepter les erreurs et l'aléatoire

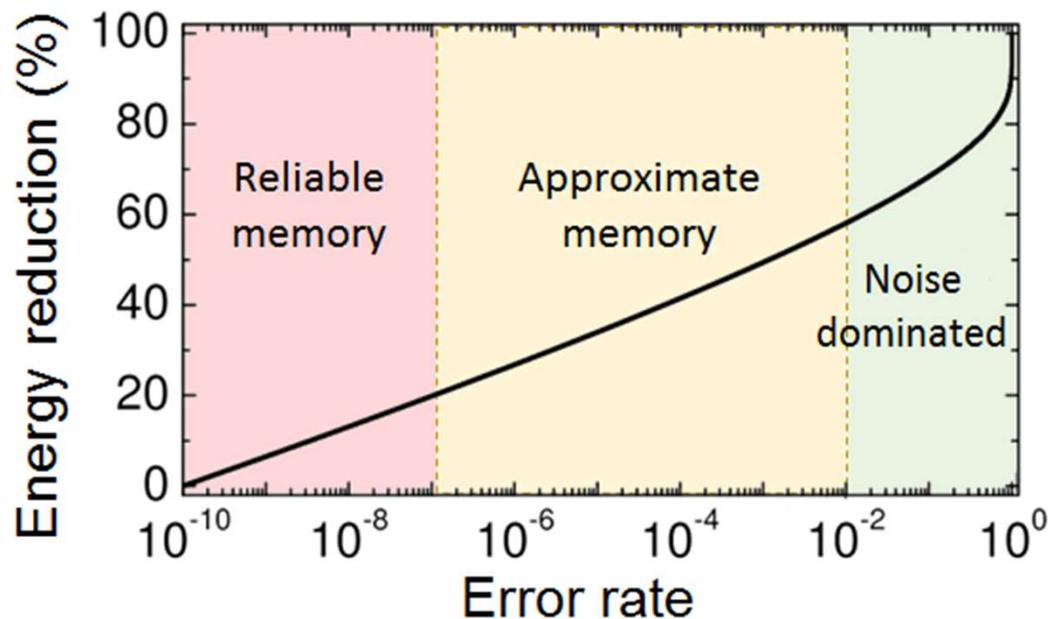


Signaux bruités,
Environnement biologique
Composants avec variabilité
Comportement aléatoire

...

Pourtant calcule et consomme peu !

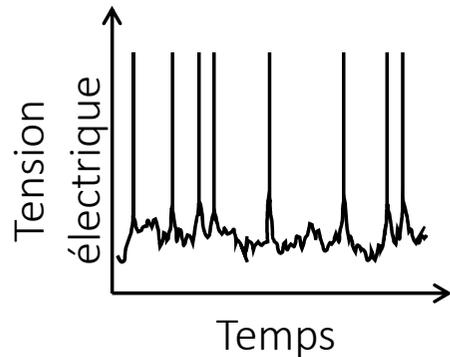
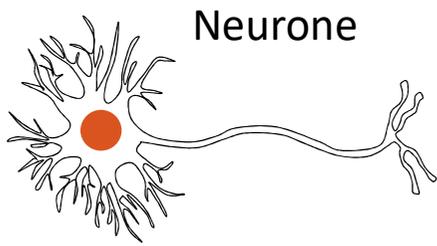
Compromis précision/énergie avec le calcul approximatif



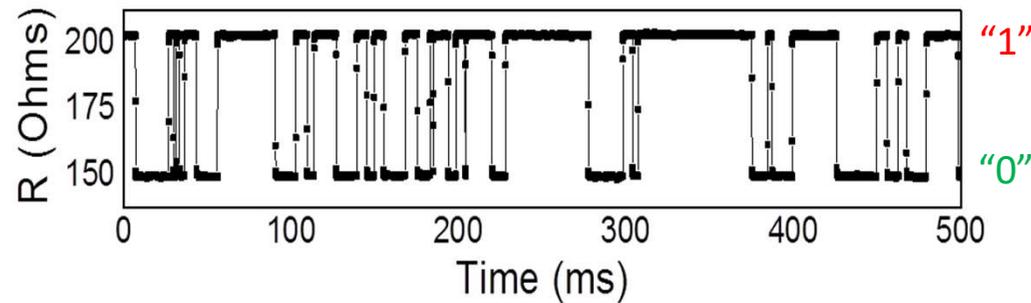
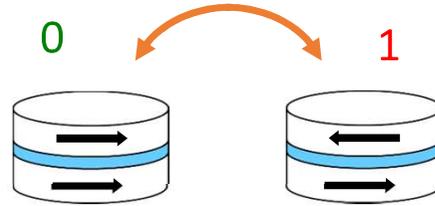
Nano-mémoire magnétique :
Fort courant de programmation -> précis
Faible courant -> moins précis

Après apprentissage, un réseau de neurones peut avoir des synapses peu précises et calculer avec succès

Un nano-oscillateur aléatoire comme neurone artificiel



Température



- Simulation d'un circuit de calcul pour le contrôle d'un robot 
- Consomme moins qu'un neurone artificiel sans nano-dispositif! 

L'énergie du nano-oscillateur est fournie par la température ambiante ! 😊

Utiliser les nano-dispositifs pour l'IA : où en est on ?



Expériences

Preuves de concept avec quelques nano-dispositifs



Simulations

Circuits avec électronique conventionnelle et certaines des nano-dispositifs

Ce que l'on va faire :

- Intégration des dispositifs dans des circuits électroniques de calcul conventionnels
- Démonstrateurs avec de nombreux nano-dispositifs
 - Taches de calcul plus complexes

Utiliser les nano-dispositifs pour une IA moins énergivore



Besoin d'une IA qui consomme moins
Requiert de nouveaux types d'ordinateurs
Inspiration pour cela : le **cerveau** !
Éléments clefs pour construire : **nano-dispositifs**



Calculer dans la mémoire
(« mémoire intelligence »)



Nano-mémoires rapides et
non-volatiles

Utiliser propriétés
dynamiques non-linéaires



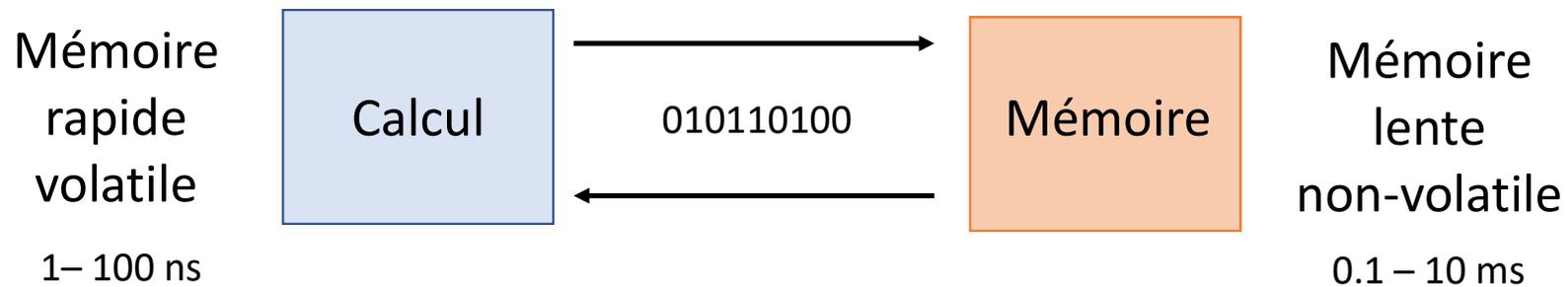
Nano-oscillateurs non-linéaires

Accepter l'imprécision
et l'aléatoire



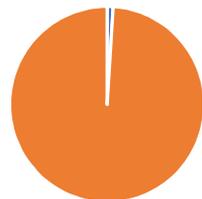
Programmation basse consommation
Nano-oscillateurs aléatoires

Communiquer entre mémoire et calcul consomme énormément



Temps d'exécution
réseau de neurones

AlexNet



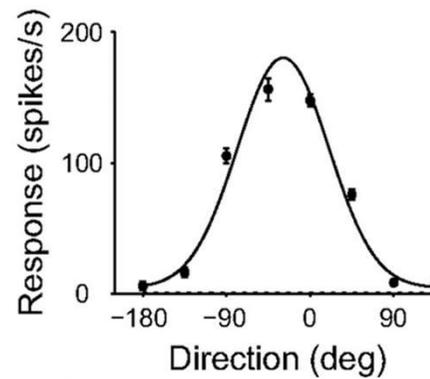
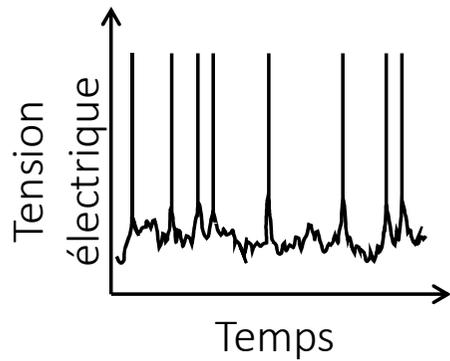
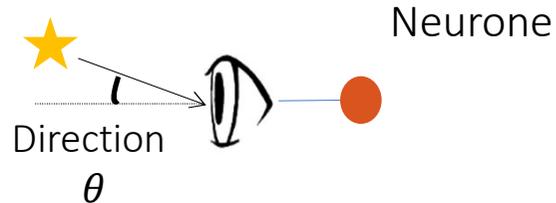
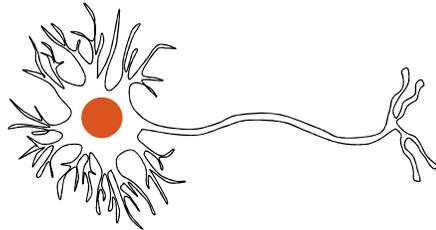
LSTM

■ Calcul ■ Mémoire

Energie communication mémoire:
60x calcul (sur puce)
3500x calcul (hors puce)



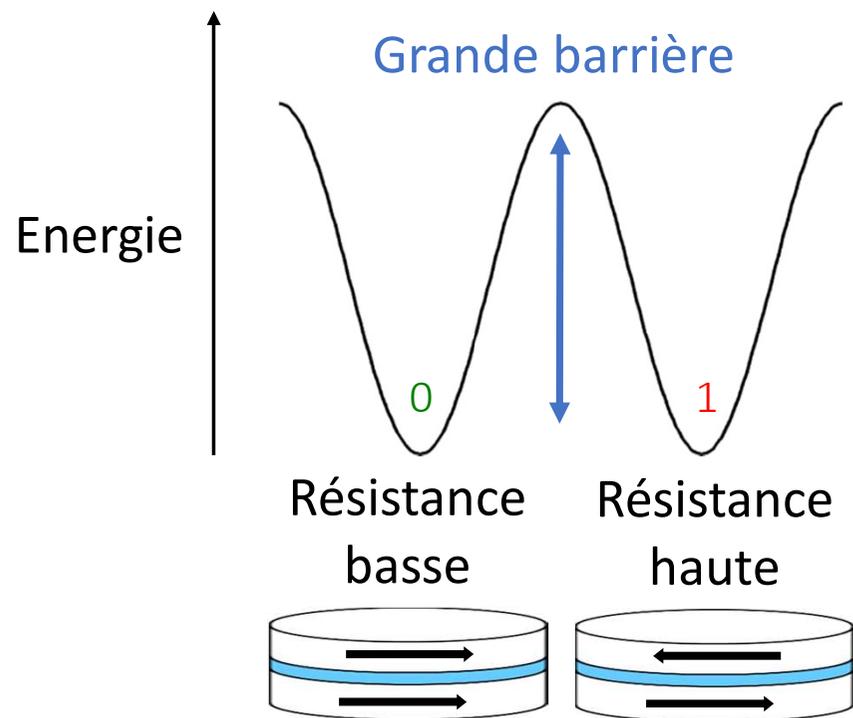
Certains neurones calculent avec des signaux aléatoires !



Réponse non-linéaire

Kumano et al., J. Neurophysiology 2010

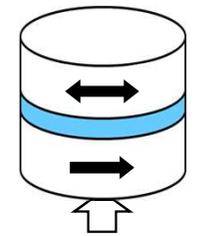
Un nano-oscillateur aléatoire comme neurone artificiel



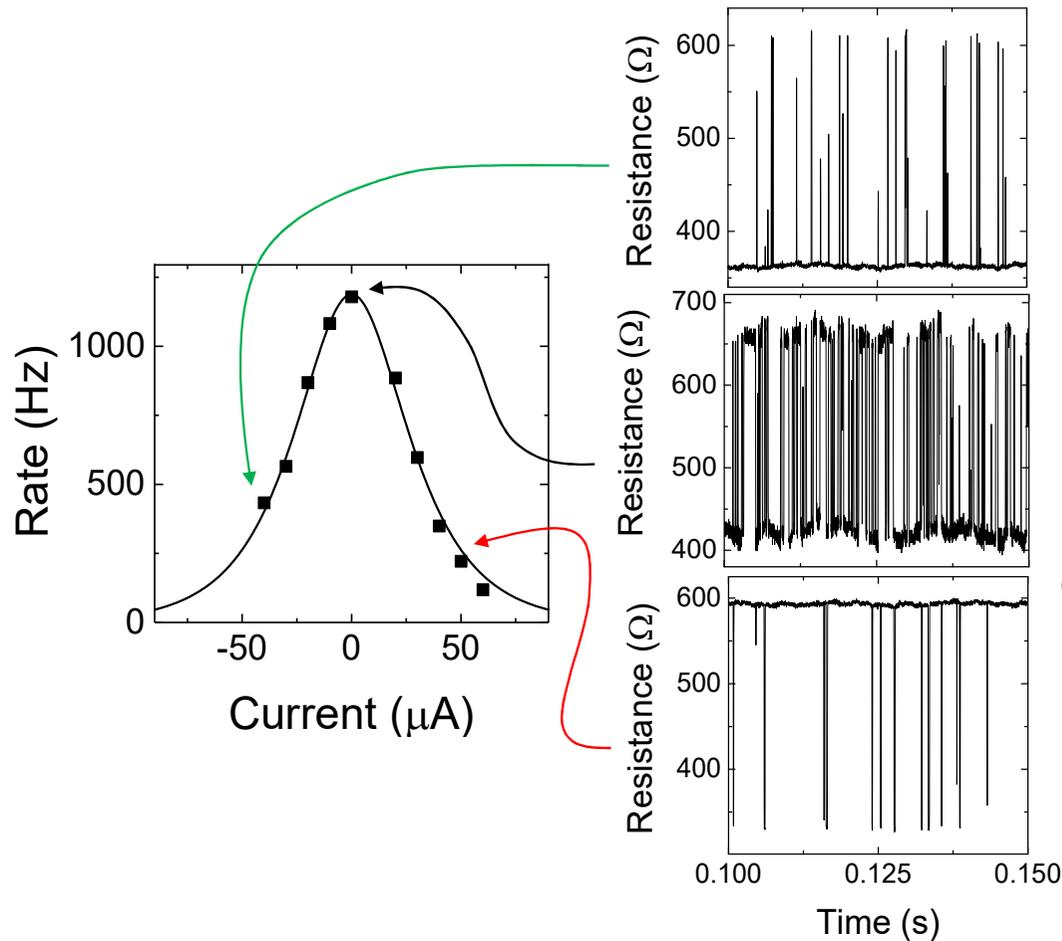
« 0 » et « 1 » sont des états stables

Mémoire non-volatile

Un nano-oscillateur aléatoire comme neurone artificiel



Courant électrique



Même réponse non-linéaire que le neurone

Peut faire du calcul avec

Consomme moins que circuit électrique équivalent sans nano-oscillateur