

# Apprentissage par Renforcement Appliqué à l'Opération des Systèmes RF du PS au CERN

*Journées Accélérateurs de la SFP  
05/10/2023*

A. Lasheen, J. Wulff

Remerciements:

H. Damerau, V. Kain, N. Madysa, opérateurs CPS



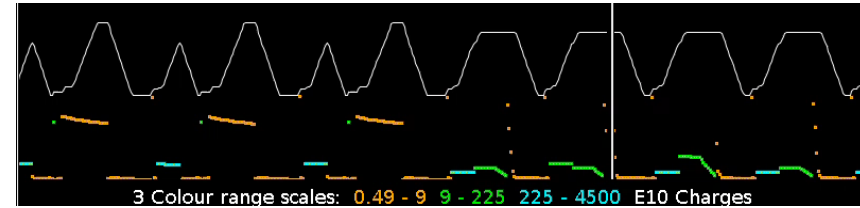
# Plan

---

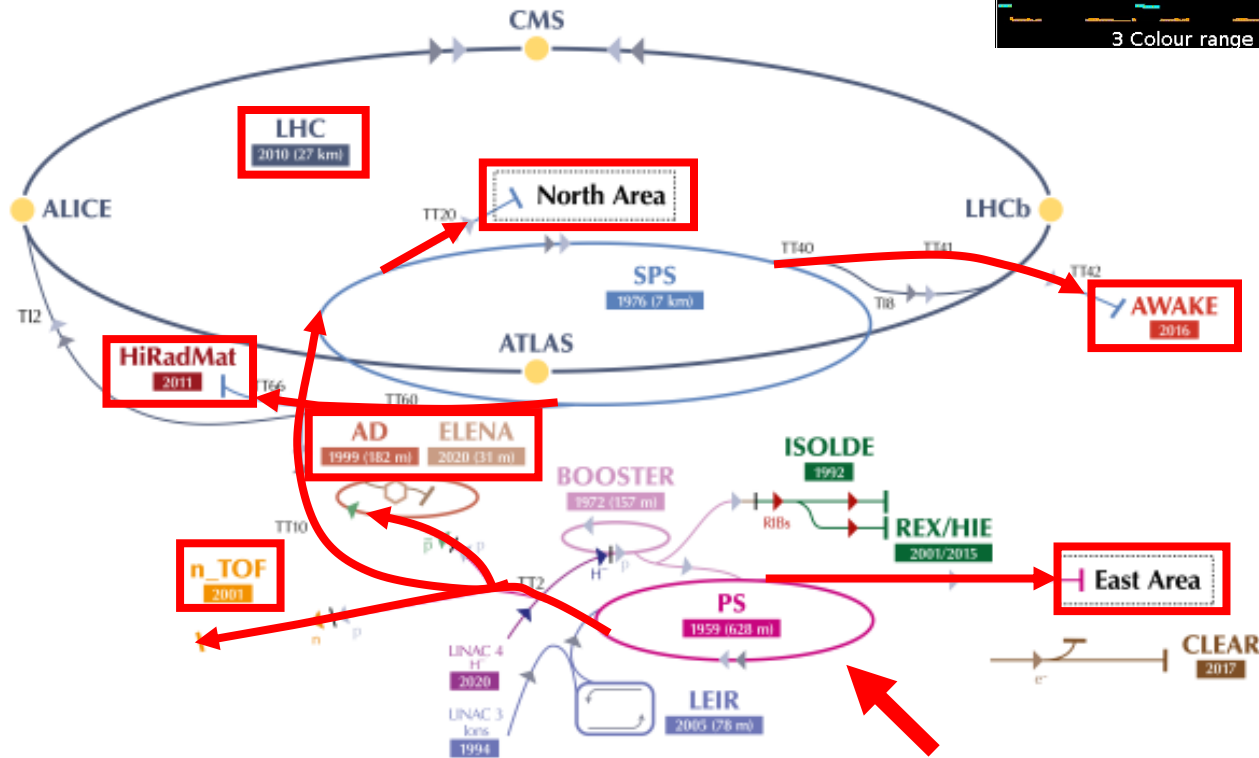
- Introduction
  - *Le PS au CERN et les manipulations RF*
- L'apprentissage par renforcement appliqué au « triple splitting »
  - *Implémentation des agents*
  - *Validation expérimentale*
  - *Limitations et défis*
- Déploiement à large échelle
  - *Développer une approche adaptée aux accélérateurs*
- Conclusions

# Le PS au cœur du complexe du CERN

The CERN accelerator complex  
 Complexe des accélérateurs du CERN



[Vistars \(cern.ch\)](http://Vistars.cern.ch)



- PS (1959)  
 $p = 26 Z \text{ GeV}/c$
- Espèces  
 $p+$  et ions (Pb54+)
- Destinations (10+ modes de production!)
  - EAST
  - nTOF
  - AD/ELENA
  - SPS North Area
  - HiRadMat
  - AWAKE
  - LHC (\*multiple beam variants)

▶  $H^-$  (hydrogen anions)  
 ▶  $p$  (protons)  
 ▶ ions  
 ▶ RIBs (Radioactive Ion Beams)  
 ▶  $n$  (neutrons)  
 ▶  $\bar{p}$  (antiprotons)  
 ▶  $e^-$  (electrons)

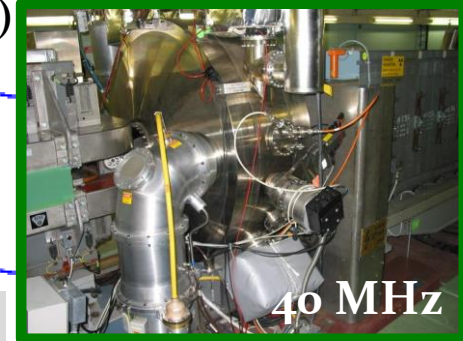
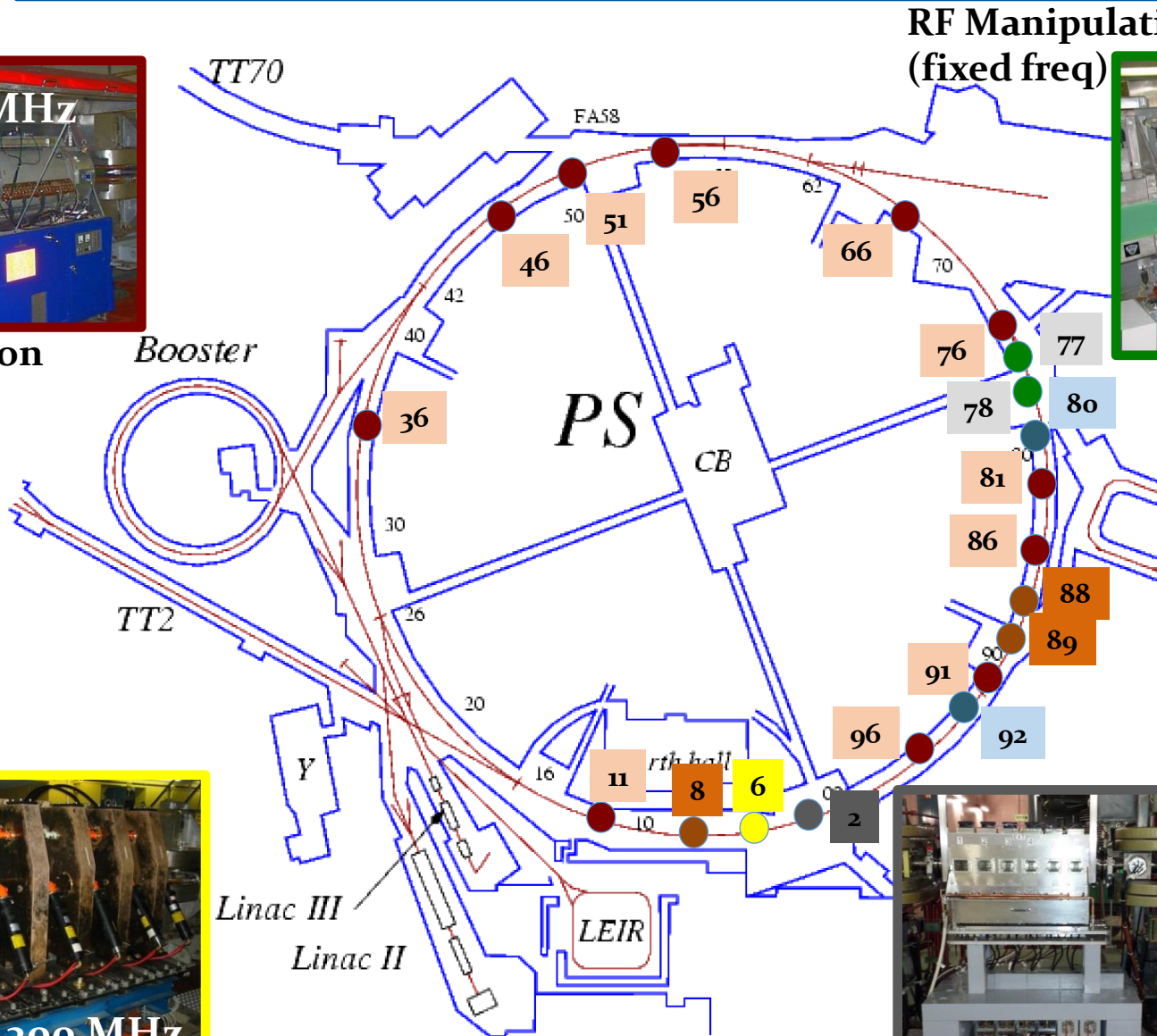
LHC - Large Hadron Collider // SPS - Super Proton Synchrotron // PS - Proton Synchrotron // AD - Antiproton Decelerator // CLEAR - CERN Linear Electron Accelerator for Research // AWAKE - Advanced WAKEfield Experiment // ISOLDE - Isotope Separator OnLine // REX/HIE - Radioactive Experiment/High Intensity and Energy ISOLDE // LEIR - Low Energy Ion Ring // LINAC - LINEar ACcelerator // n\_TOF - Neutrons Time Of Flight // HiRadMat - High-Radiation to Materials

# Les systèmes RF du PS



2.8 - 10 MHz

Acceleration  
(tuning)  
to SPS



40 MHz



80 MHz



200 MHz

Longitudinal blow-up



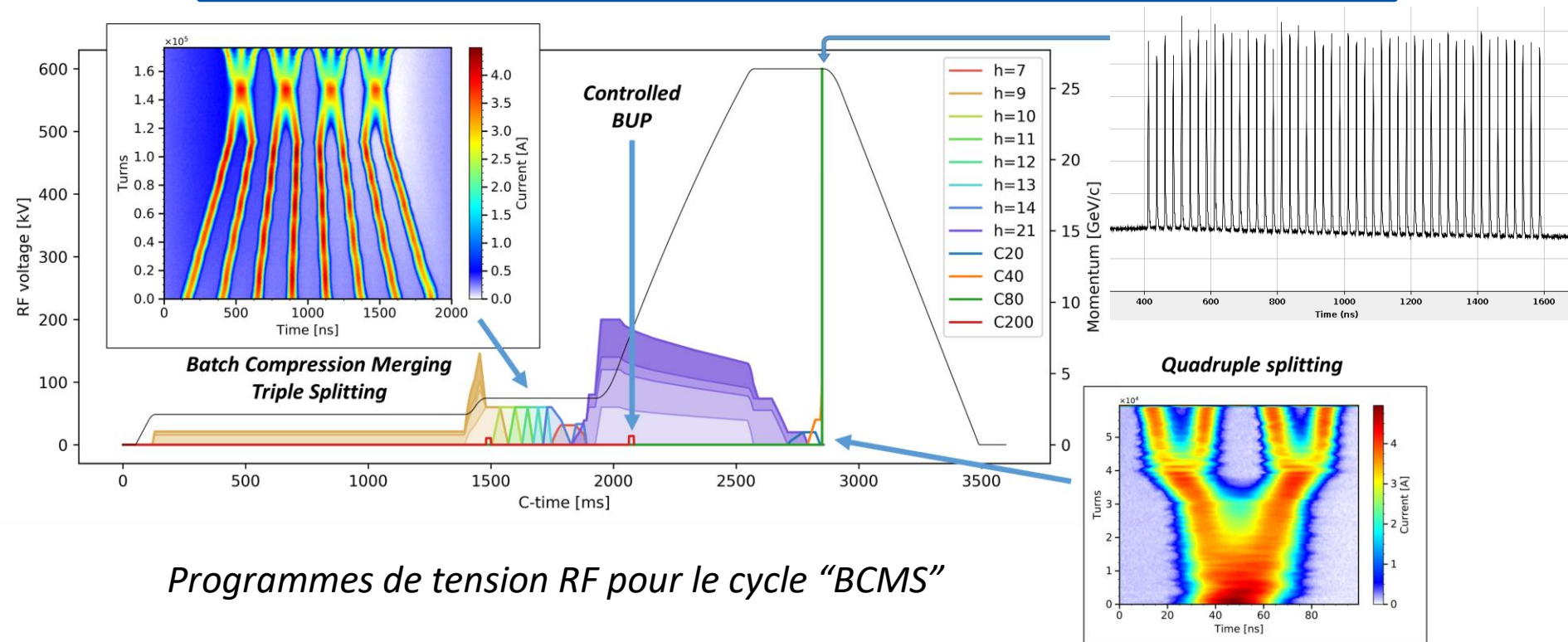
0.4 - 5 MHz

Damper



20 MHz

# Manipulations RF pour les faisceaux LHC



*Programmes de tension RF pour le cycle "BCMS"*

- Les cavités RF du PS permettent de couvrir un grand nombre d'harmoniques permettant de faire des manipulations RF.
- Le PS définit l'espace entre les paquets de 25 ns pour les faisceaux allant au LHC.
- L'objectif primaire en qualité est de produire des paquets égaux en intensité et en émittance longitudinale, pour éviter les pertes en aval et les variations en luminosité dans le LHC.



# Les ajustements RF au PS, une tâche quotidienne

- Certaines manipulations RF sont linéaires, avec une observable pour un paramètre à optimiser. Pourquoi utiliser de l'apprentissage machine? Un coup d'œil au journal du PS pour les «splittings»...

- Les critères de performance peuvent être insuffisants

Checked out the 72b for the SPS. Looks fine, may need slight splitting adjustment

We correct injection oscillations, the splitting looks good enough.

We have to equalize the PSB batch intensity, then we get good beam, no losses and bunch splitting not too bad.

The SPS BQM complained from time to time about the splitting. We investigate but it seems as good as usual.

SPS complained about the splitting, but finally, they were happy.

- Les paramètres varient avec le temps

Quite a few changes needed to get it acceptable for SPS, but still not really great.

Unsure of the reason for the drift, since we're using the same cavities as previous fills.

Following shifts may wish to consider rolling back to before, if splitting still needs adjusting; drifts may only be temporary.

- Toutes les manipulations ne sont pas triviales à optimiser et dépendent de l'intensité du faisceau

The triple splitting is very bad we try to correct but it's not easy.

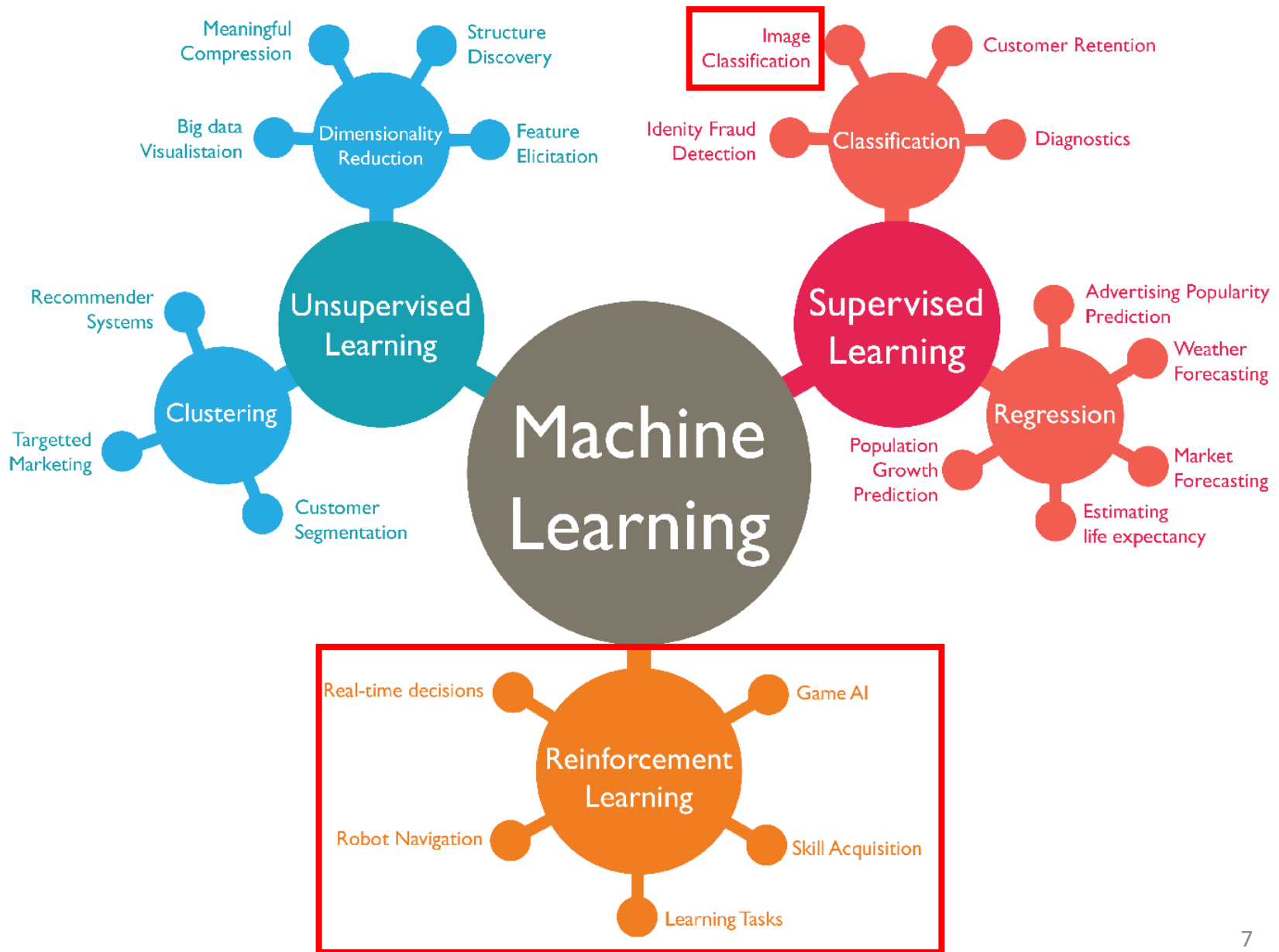
Alexandre Lasheen and Heiko will have a look

- Finalement...

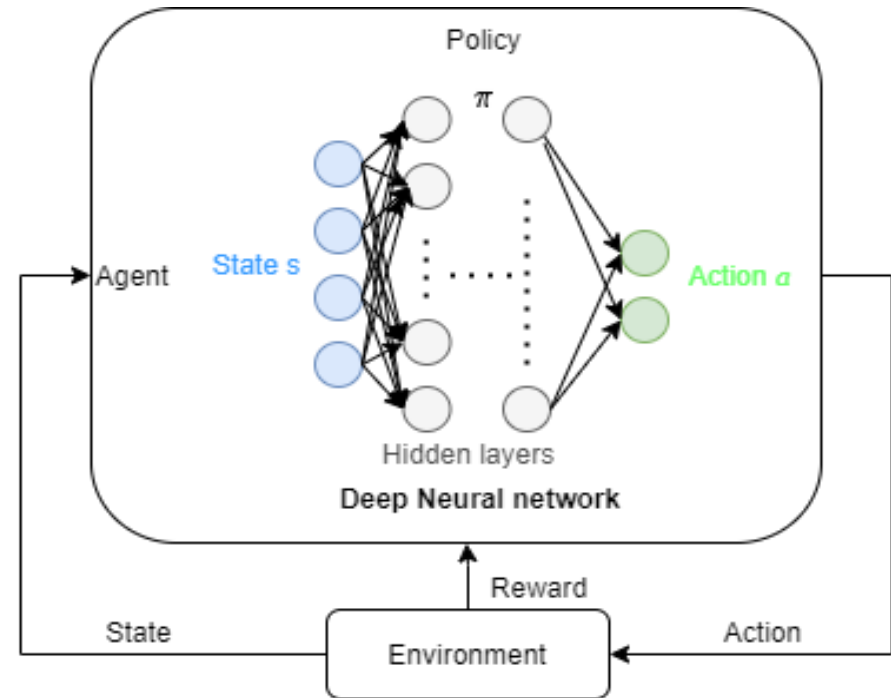
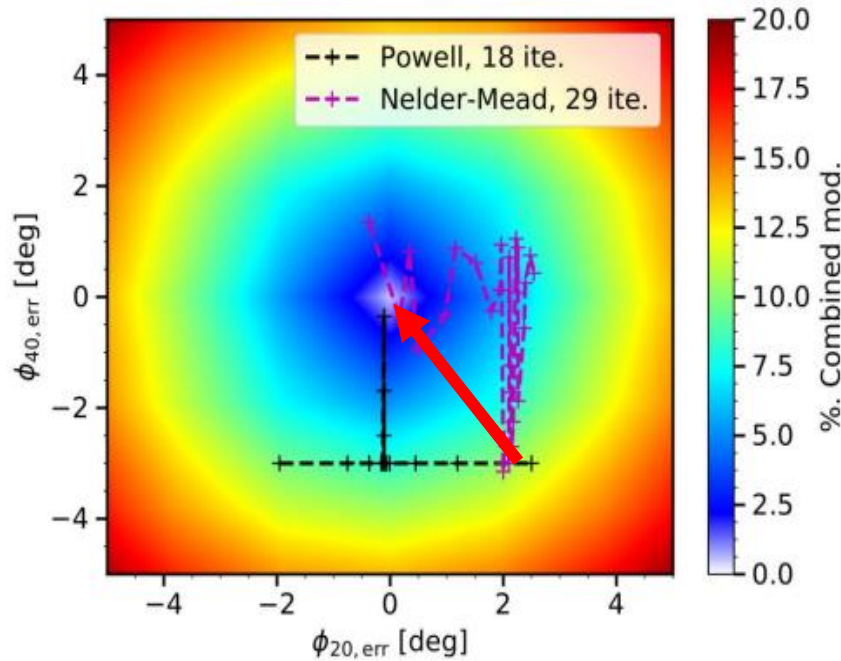
Everybody is happy now.

All is good.

# Apprentissage par renforcement profond (RL)



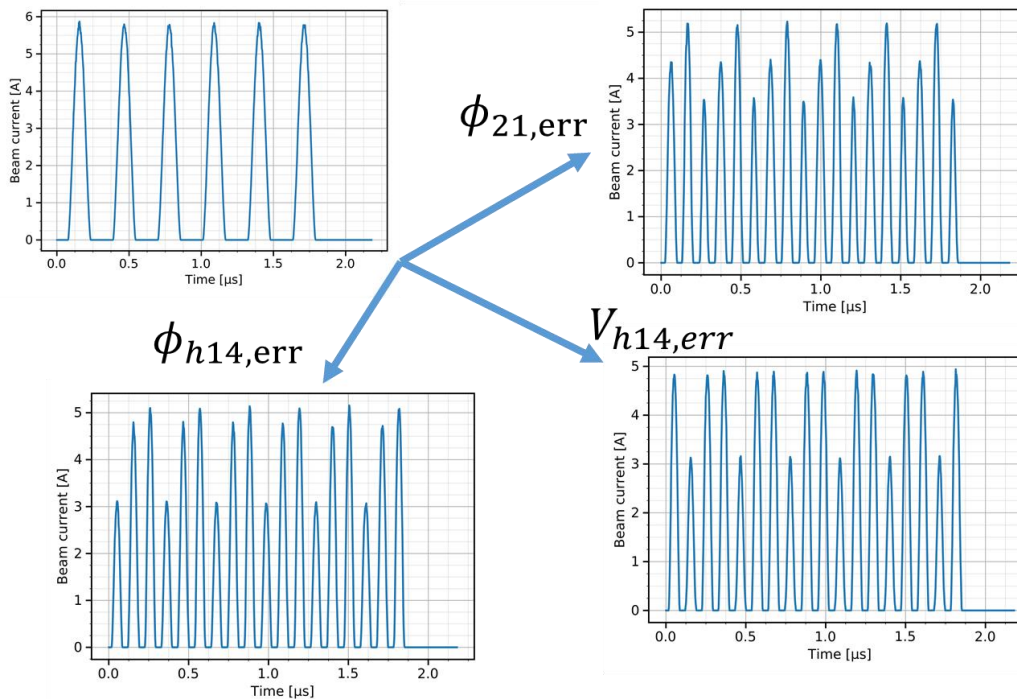
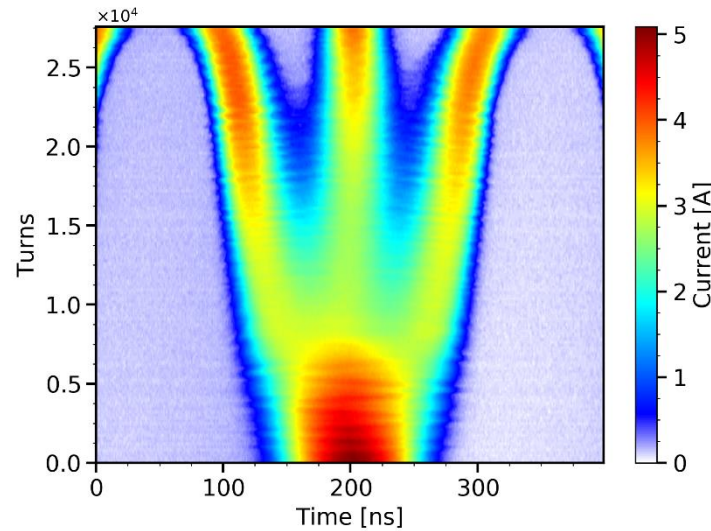
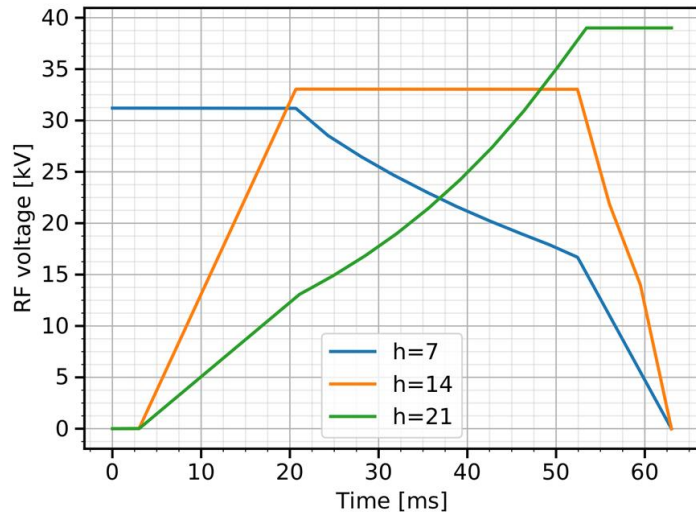
# Apprentissage par renforcement profond (RL)



- L'apprentissage par renforcement (RL) vise à obtenir une optimisation efficace en capitalisant sur l'expérience acquise, mesurée ou simulée.
- L'optimisation standard prend de l'ordre de 20-30 itérations, l'objectif de l'apprentissage par renforcement est d'améliorer la performance par un ordre de grandeur pour des tâches complexes.

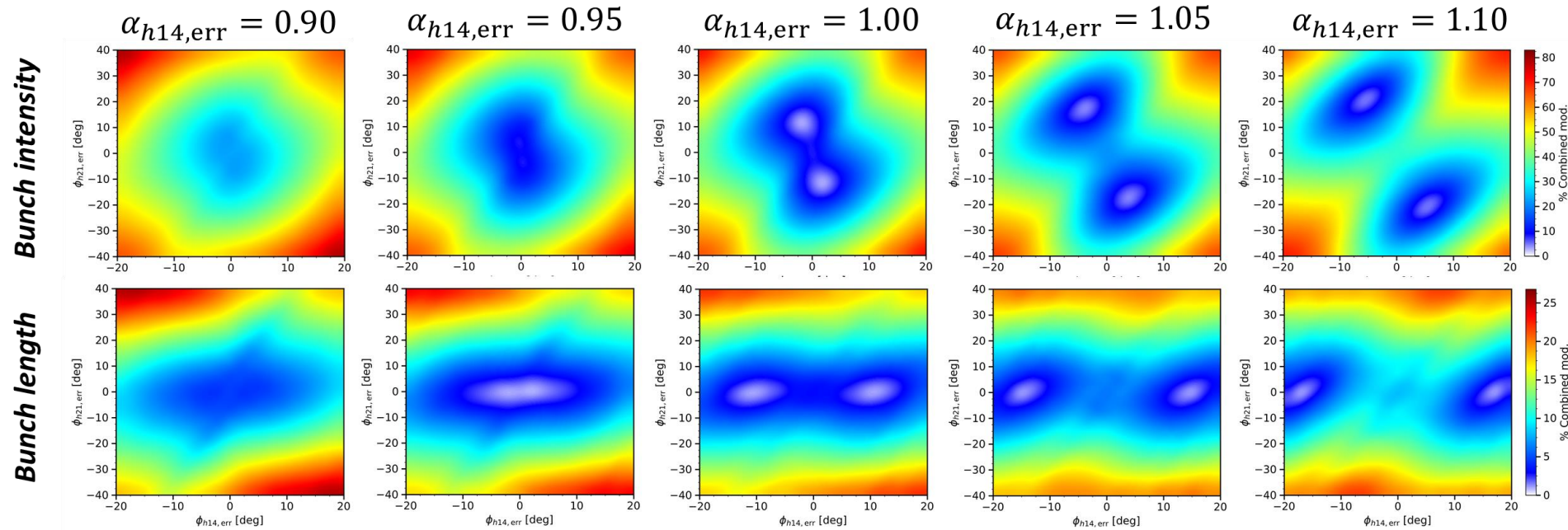


# Le triple splitting: un problème équilibré (1)



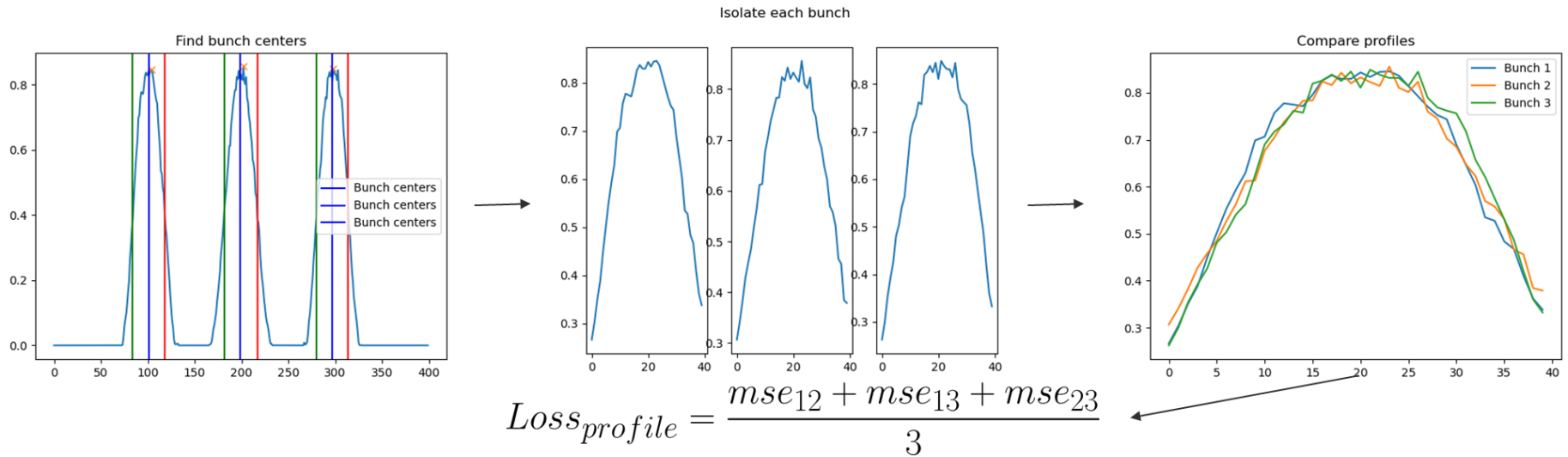
- Le triple splitting est un challenge équilibré pour l'apprentissage par renforcement.
- Trois paramètres interdépendants doivent être ajustés pour obtenir 3 paquets identiques.

# Le triple splitting: un problème équilibré (2)



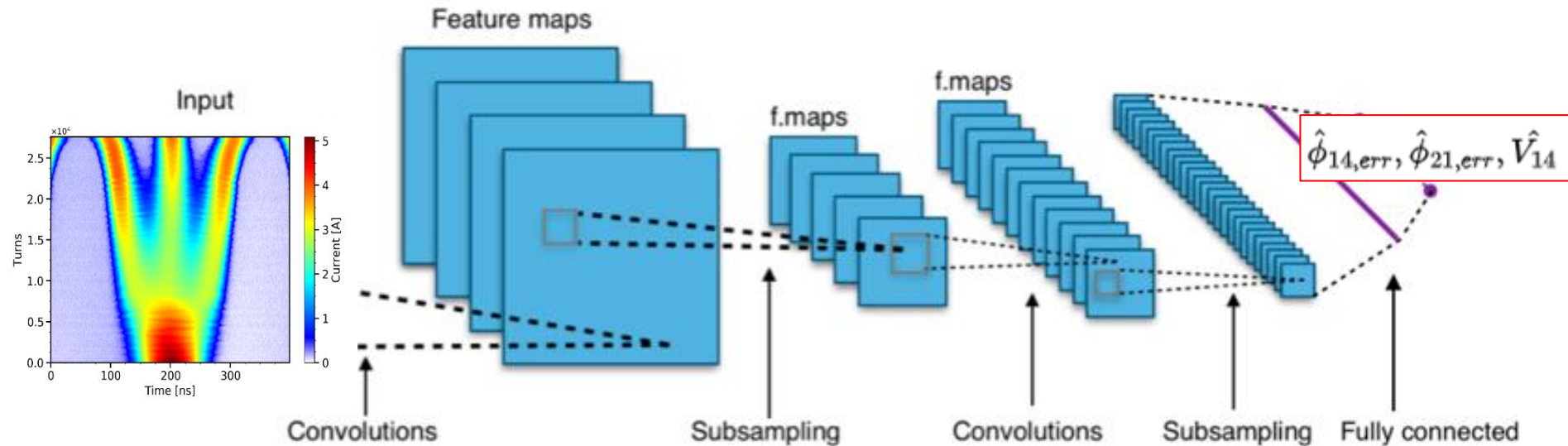
- L'optimisation du triple splitting est non linéaire et présente de nombreux minima locaux.
- Un seul optimum global existe pour obtenir des paquets avec la même intensité et émittance longitudinale.
- Pour certaines études, le triple splitting doit être optimisé à chaque intensité et peut ralentir considérablement l'opération.

# Expérimentations préliminaires (1)



- Entraînement d'un agent sur les profils finaux, visant à obtenir des paquets identiques.
- Les données obtenues en simulations. Divers algorithmes et implémentations testés (DDPG, SAC, ...).
- Echec de convergence, même en simulations! L'agent ne trouve pas de cheminement avec ces seules données.

# Expérimentations préliminaires (2)

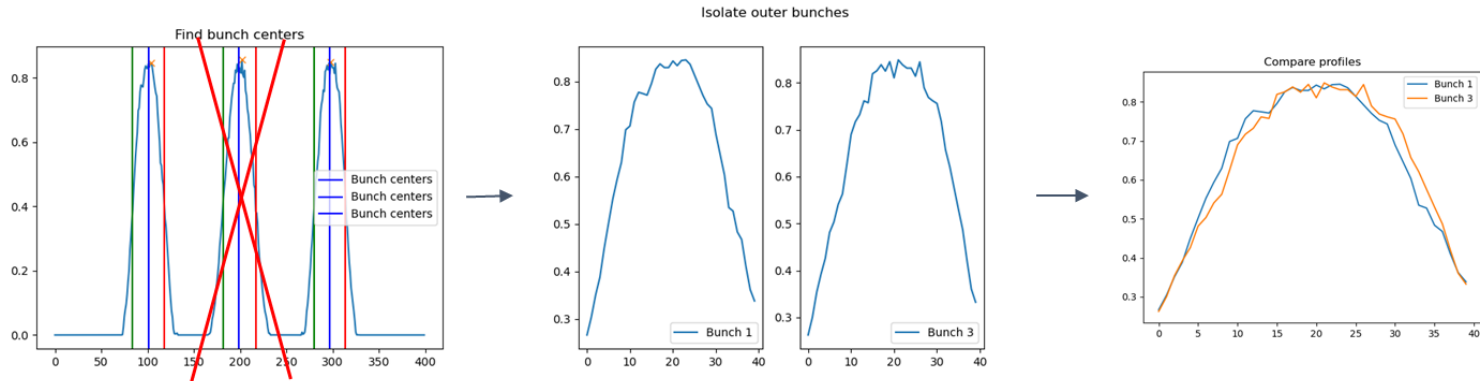


- L'évolution des paquets peut être analysée à l'aide de réseaux neuronaux convolués (CNN), traitant l'évolution des paquets comme une image.
- L'agent obtient les paramètres en sortie du CNN et applique la correction telle quelle.
- Fonctionnel en simulations, échec de l'implémentation avec le faisceau! L'agent n'arrive pas à généraliser par rapport aux biais de mesure absent des simulations.

# Segmentation du problème (1)

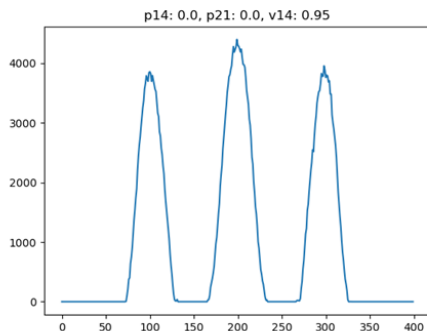
Deux agents indépendants lancés de façon séquentielle.

1. Exclusion du paquet central et ajustement uniquement des phases pour régler l'asymétrie gauche/droite

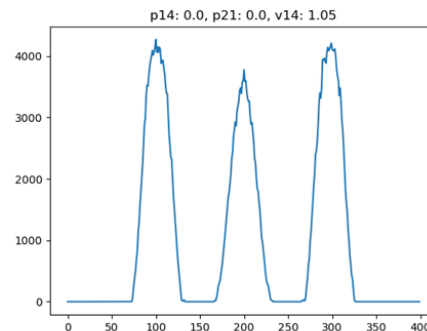


2. Réglage de la tension du palier intermédiaire pour équilibrer le paquet central

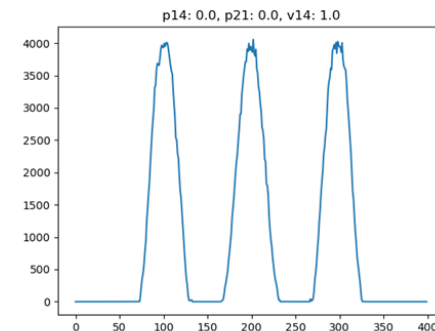
A) Middle profile larger than the others  
→ Voltage too low!



B) Middle profile smaller than the others  
→ Voltage too high!



C) Middle profile equal to the others  
→ Voltage optimised!



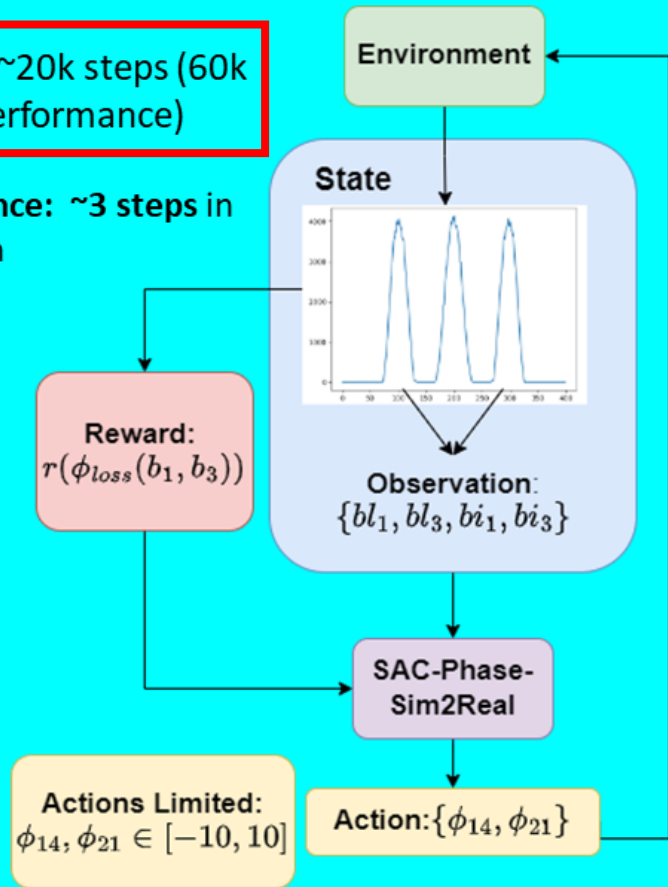


# Segmentation du problème (2)

## Phase Agent

**Training:** ~20k steps (60k to best performance)

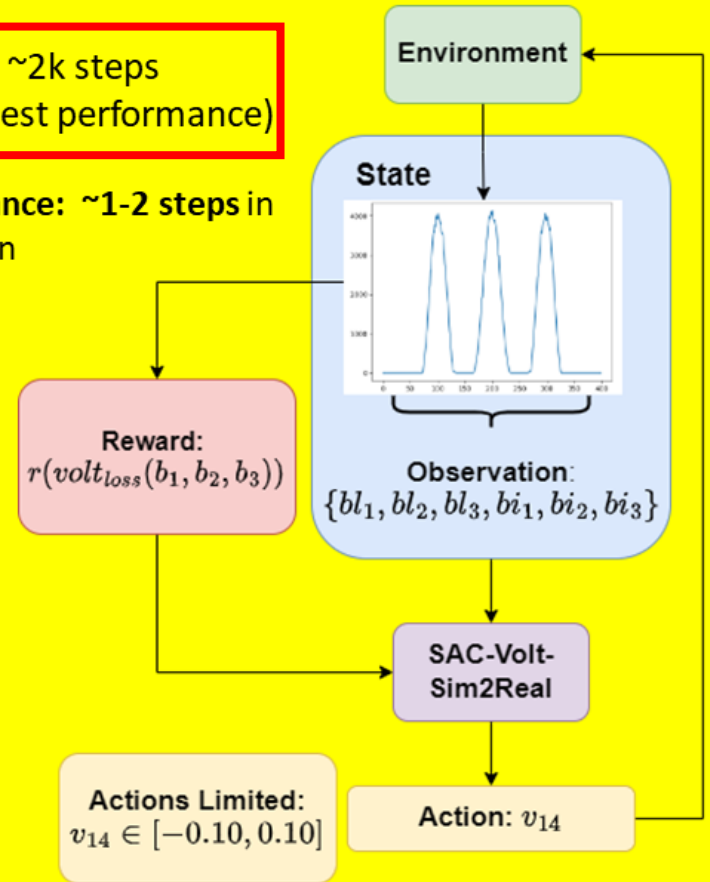
**Performance:** ~3 steps in simulation



## Voltage Agent

**Training:** ~2k steps (20k to best performance)

**Performance:** ~1-2 steps in simulation

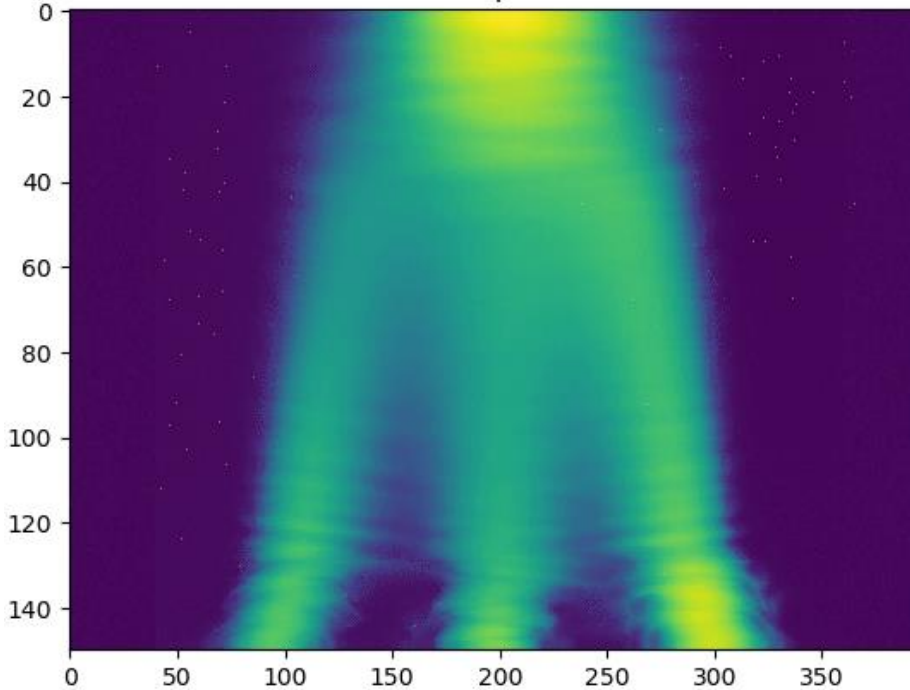


- Les deux agents requièrent plusieurs milliers d'itérations d'entraînement pour obtenir des performances très satisfaisantes en simulations

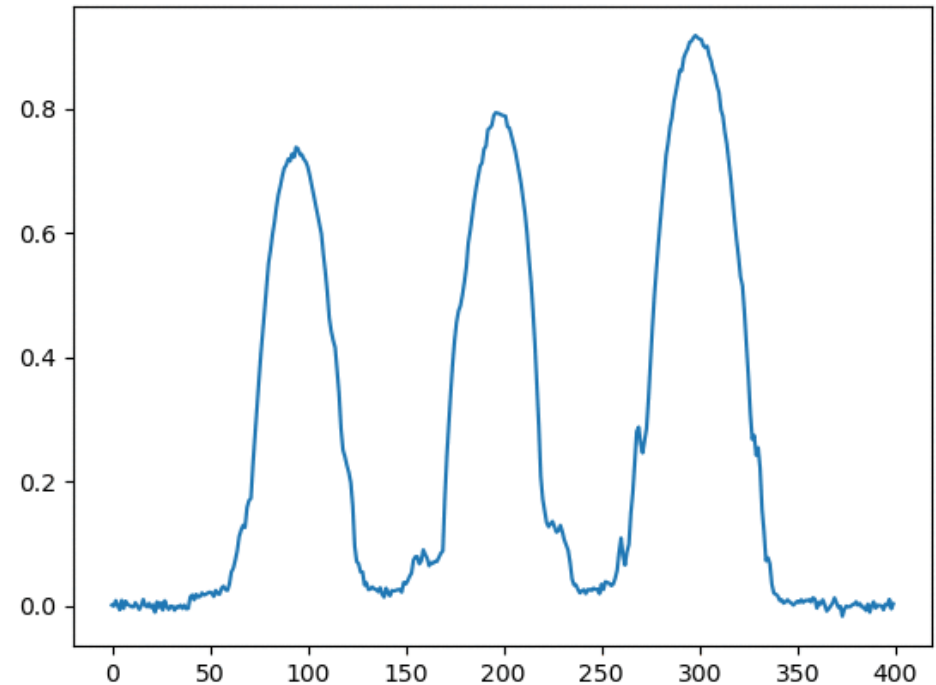


# Premiers résultats avec deux agents

Step: 0



Step: 0, Loss: 0.023790403209074773

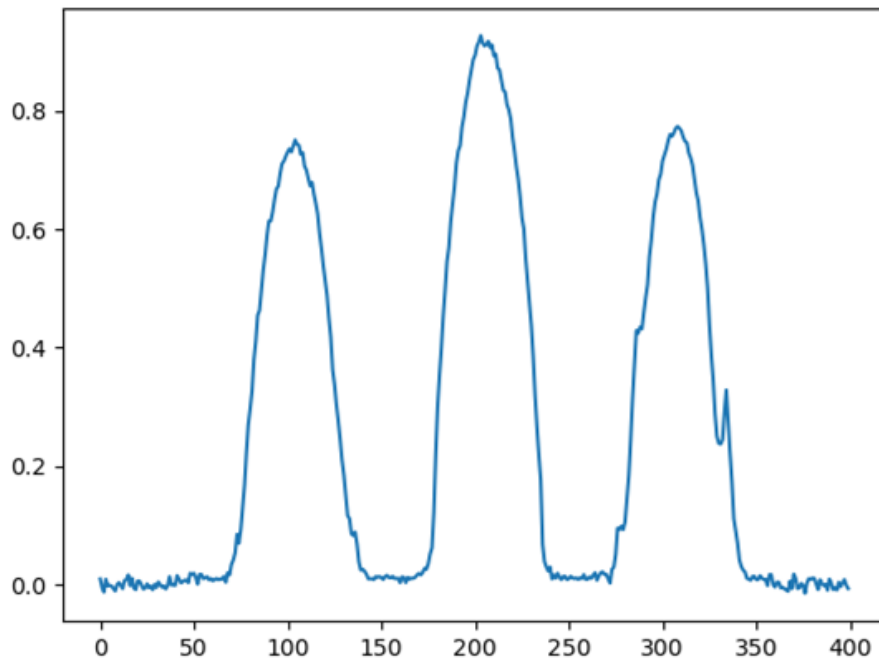


■ Les premiers résultats étaient très encourageants!

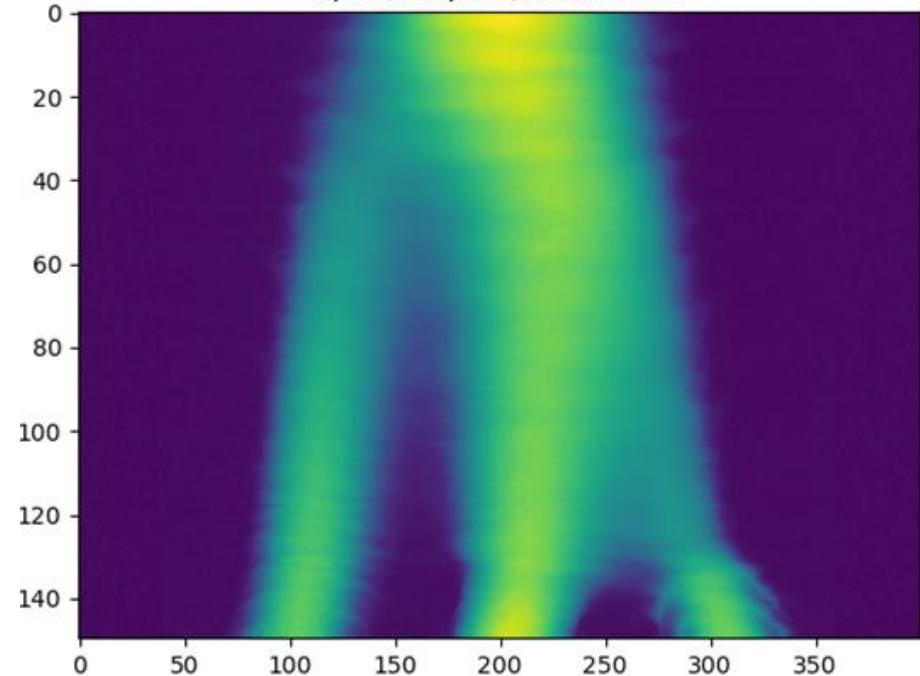
➤ Mais assez rapidement certaines limites...

# Combinaison de modèles CNN + RL

Ep. 2, step 11, profile

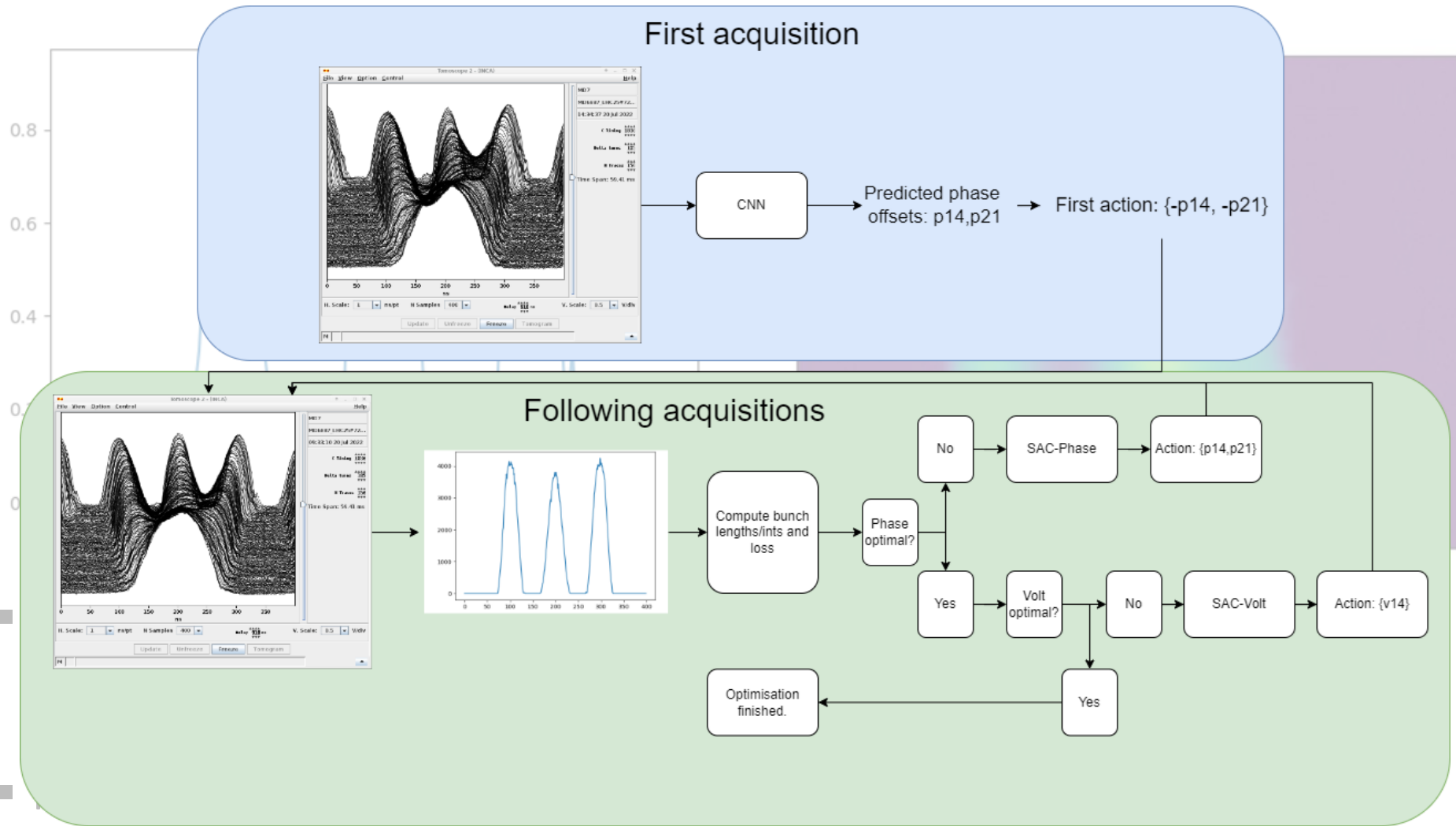


Ep. 2, step 11, datamatrix



- Après optimisation de la symétrie gauche/droite par le premier agent ajustant la phase, certains épisodes d'optimisation échoue à la deuxième étape.
- Pour des cas très loin de l'optimum, l'agent se heurte à un minimum local!

# Combinaison de modèles CNN + RL



➤ Ce cas de figure est solutionné en utilisant le CNN présenté précédemment pour faire un premier pas pour se rapprocher de l'optimum global!

# Résultats en opération (1)

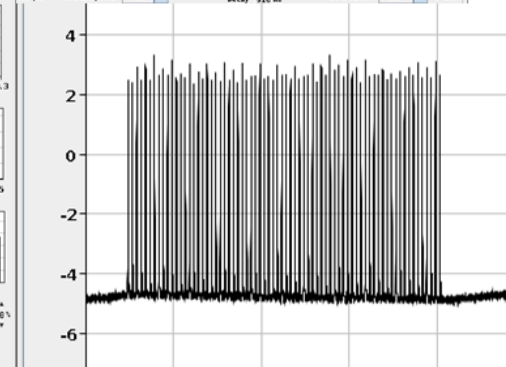
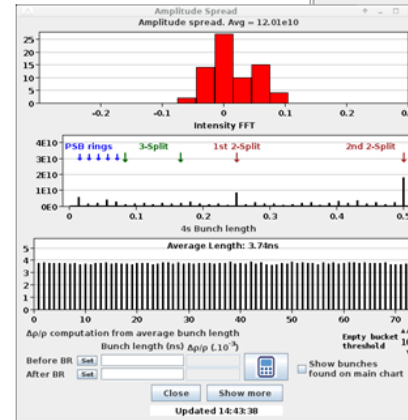
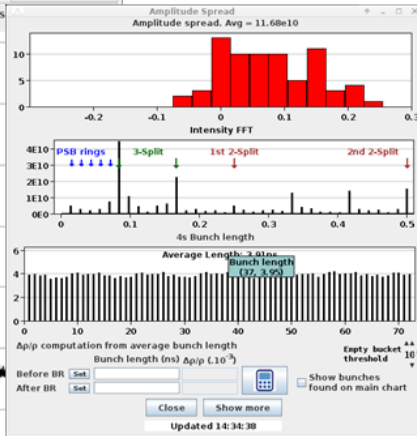
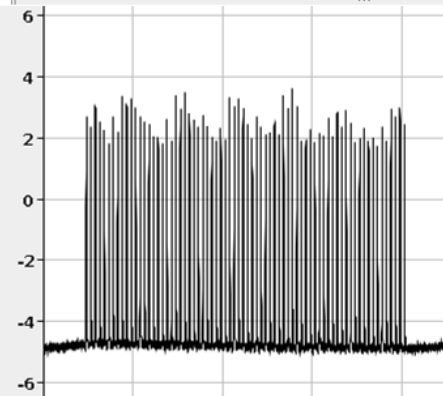
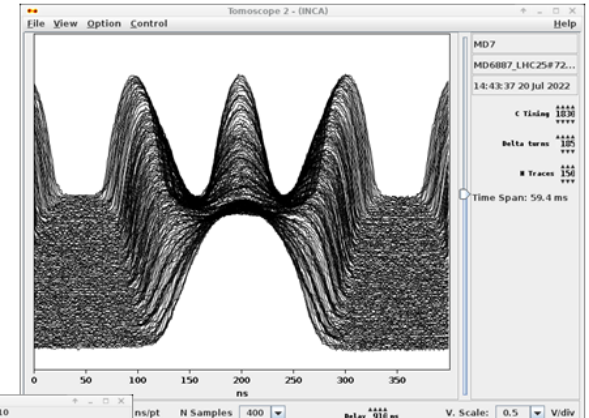
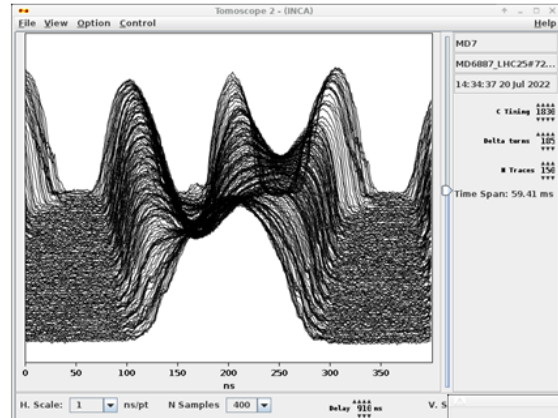
Example episode:

Approx. initial offset:  $\phi_{14} = 10$ ,  $\phi_{21} = -20$ ,  $V_f = 1.08$

Init

Final

Phase opt. steps: 3  
 Volt opt. steps: 4  
 Total iterations required: 7

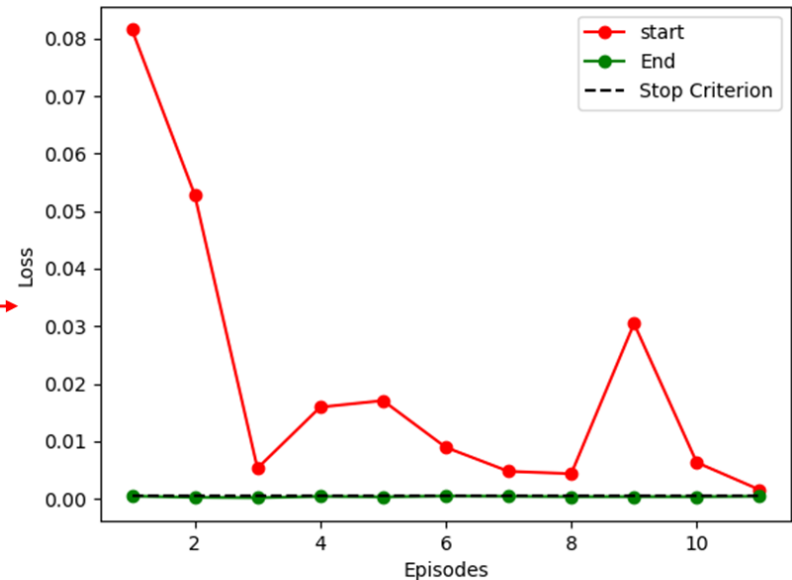


# Résultats en opération (2)

Episode	Init settings [ $p14, p21,$ $v14\_offset$ ]	Phase opt.	Voltage opt.	Success
1	20, 20, -0.08	5	7	Yes!
2	-20, -20, 0.10	8	3	Yes!
3	-20, 20, 0.10	8	3	Yes!
4	20, -20, -0.01	9	5	Yes!
5	15, -5, -0.05	3	6	Yes!
6	15, -5, 0.05	1	1	Yes!
7 (2.6e11)	10, -10, -0.10	8	10	Yes!
8 (2.6e11)	-10, 10, 0.10	2	2	Yes!
9 (2.6e11)	10, 10, -0.05	5	10	Yes!
10 (BCMS)	10, -10, 0.05	3	2	Yes!
11 (BCMS)	-10, 10, -0.05	2	3	Yes!

Figure: Start and end criterion for episodes 1-11. Computed using full profile loss, i.e. comparing all three bunches.

MD 6887: Start and end Agent Criterion (Comparing all bunches)



➤ Des performances très satisfaisantes sont aussi obtenues en mesures!

➤ L'implémentation est désormais en test depuis fin 2022 avec un **taux de succès > 95%** et **des performances supérieures à l'ajustement manuel d'un expert** (2-3 itérations pour des ajustements standards).

# Conclusions, perspectives, limitations (1)

- Un algorithme d'automatisation des « splittings » au CERN PS basé sur l'apprentissage par renforcement et entraîné sur des simulations a été développé.
- L'approche adoptée consiste à segmenter le problème en plusieurs tâches contrôlées par des agents RL différents.
- L'algorithme atteint des performances supérieures ou comparables à l'ajustement par un expert ou opérateur expérimenté, et est testé en opération.
- Une limite importante est le nombre d'itérations pour entraîner le modèle  $\sigma(10^4)$ , qui rend l'entraînement sur des données mesurées inadapté.



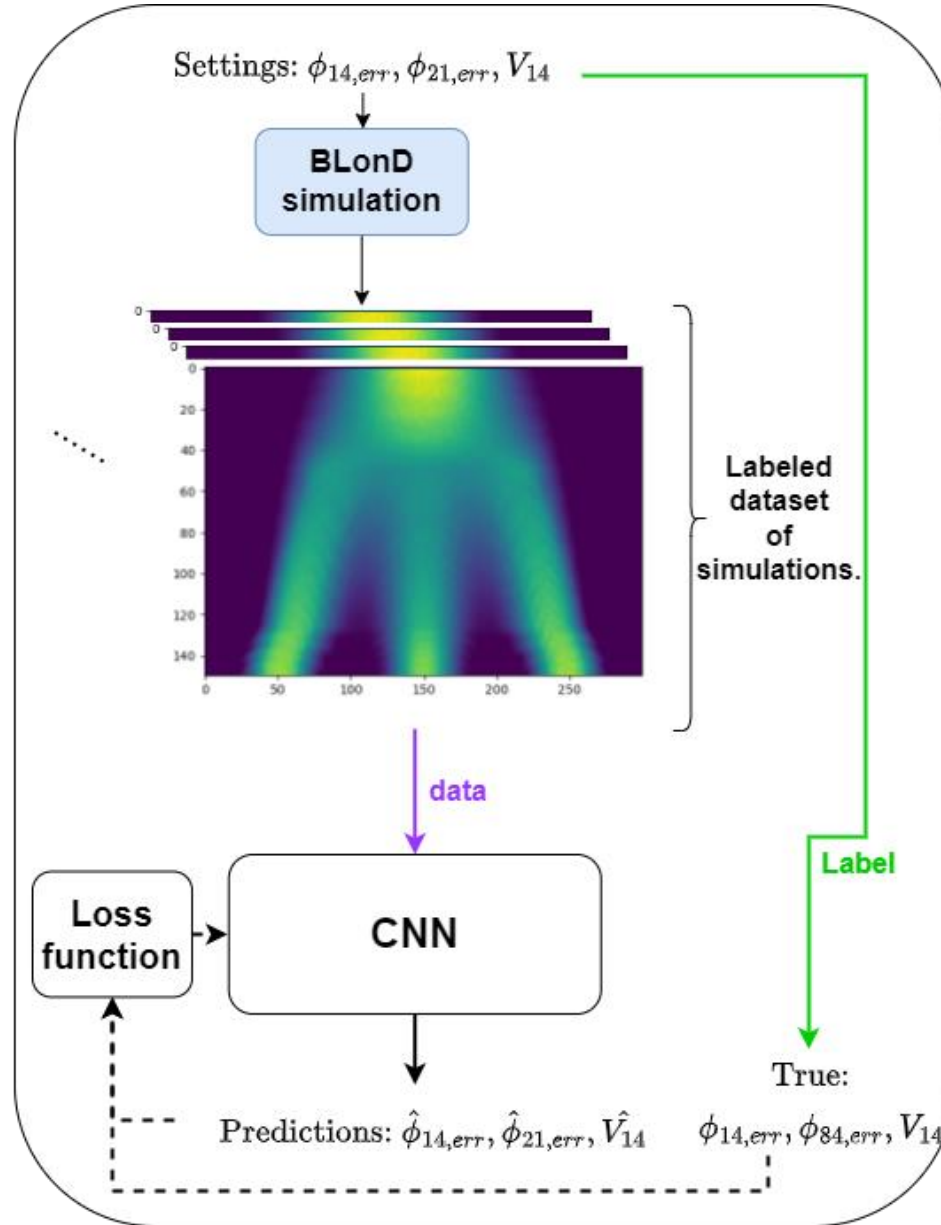
# Conclusions, perspectives, limitations (2)

---

- Plusieurs axes d'investigations pour une implémentation plus globale doivent être exploré
  - *Développement d'agents hiérarchisés pour orchestrer les séquences d'optimisation et éviter les conflits (-> sécurité de l'accélérateur)*
  - *Transfert de connaissances de modèles pour déployer des nouveaux modèles en réduisant la nécessité de reconstruire des sets de données (beaucoup de faisceaux et process à ajuster).*
  - *Ajustement fin des modèles en combinant données simulées et mesurées*
  - **Développer une stratégie et des standards pour les accélérateurs!**

**Merci pour votre attention!**

# Exemple d'entraînement du CNN



# Exemple d'architecture CNN

