

Conception et commande de robots quadrupèdes

Philippe Souères
LAAS-CNRS, Toulouse, France



Solo : Robot quadrupède ouvert commandable en couple



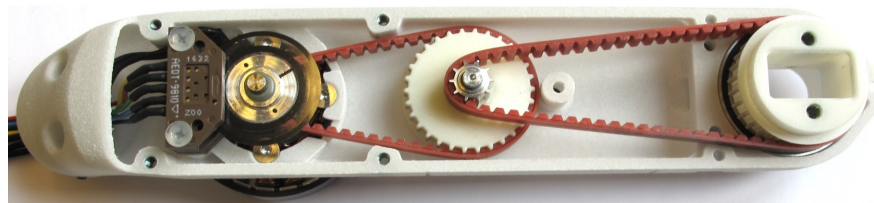
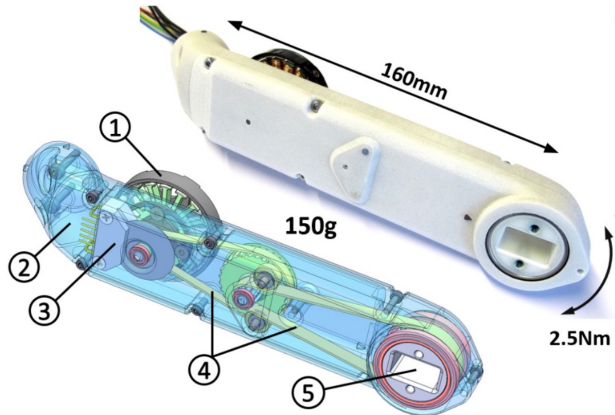
MAX-PLANCK-GESELLSCHAFT



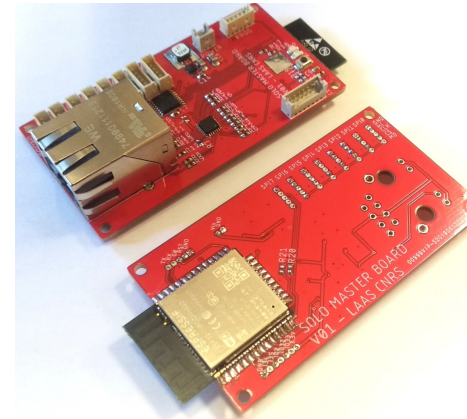
<https://open-dynamic-robot-initiative.github.io/>



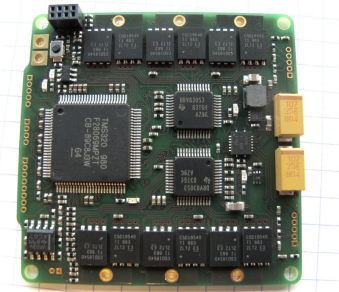
- Matériel, schémas, électronique, microprogrammes et logiciels libres
- Module d'actionnement simple permettant la conception de différents prototypes
- Pièces imprimées 3D et composants sur étagère
- Moteurs PMSM avec commande vectorielle
- Codeurs optiques haute résolution
- Transparence et réversibilité
- IMU (filtrage EKF de l'attitude)



Module d'actionnement ouvert



Carte Mère et
Micro Pilotes



F. Grimminger¹, A. Meduri, M Khadiv, J. Viereck., M Wüthrich, M. Naveau, V. Berenz, S. Heim, F. Widmaier, T. Flayols, J. Fiene, A. Badri-Spröwitz, L. Righetti, *An Open Torque-Controlled Modular RobotArchitecture for Legged Locomotion Research*, IEEE RA-L 5 (2), 3650-3657, 2020

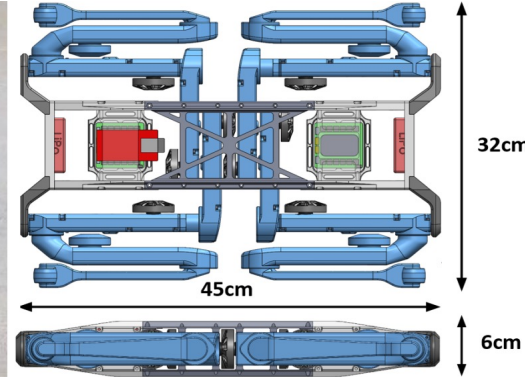




MAX-PLANCK-GESELLSCHAFT

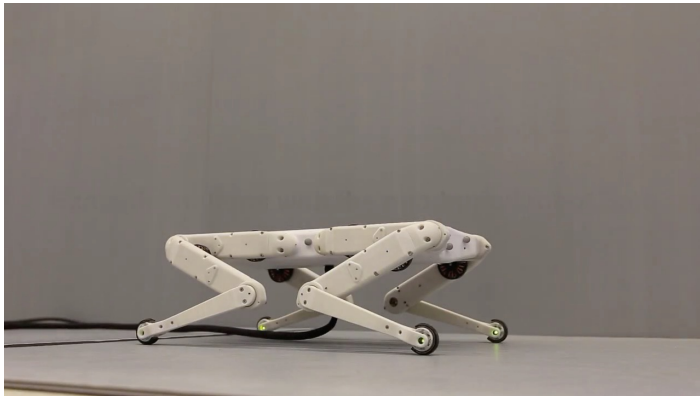


De Solo-8 à Solo-12

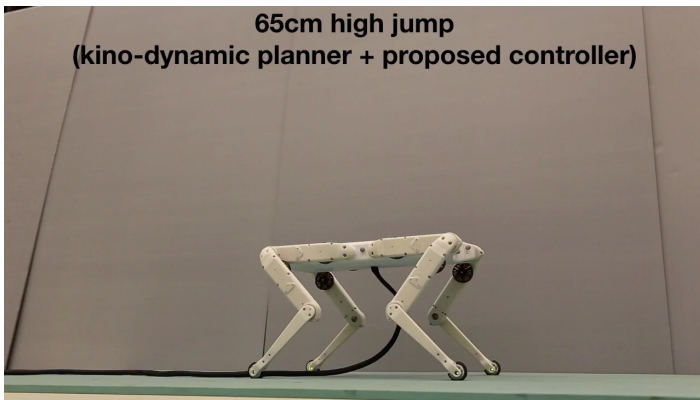


Moins de 2.5 kg, ~10k€,
3 DoFs par patte,
12 modules d'actionnement

- Connexion personnalisée permettant une boucle de commande à 1kHz avec un ordinateur distant (WiFi ou Ethernet).
- Contrôle local en impédance des liaisons (comme sur Mini-Cheetah)
- Réduction moins élevées, couples réduits, et
 - meilleure proprioception (estimation des couples par mesure en courant),
 - amortissement des impacts (restitution de l'énergie aux batteries).



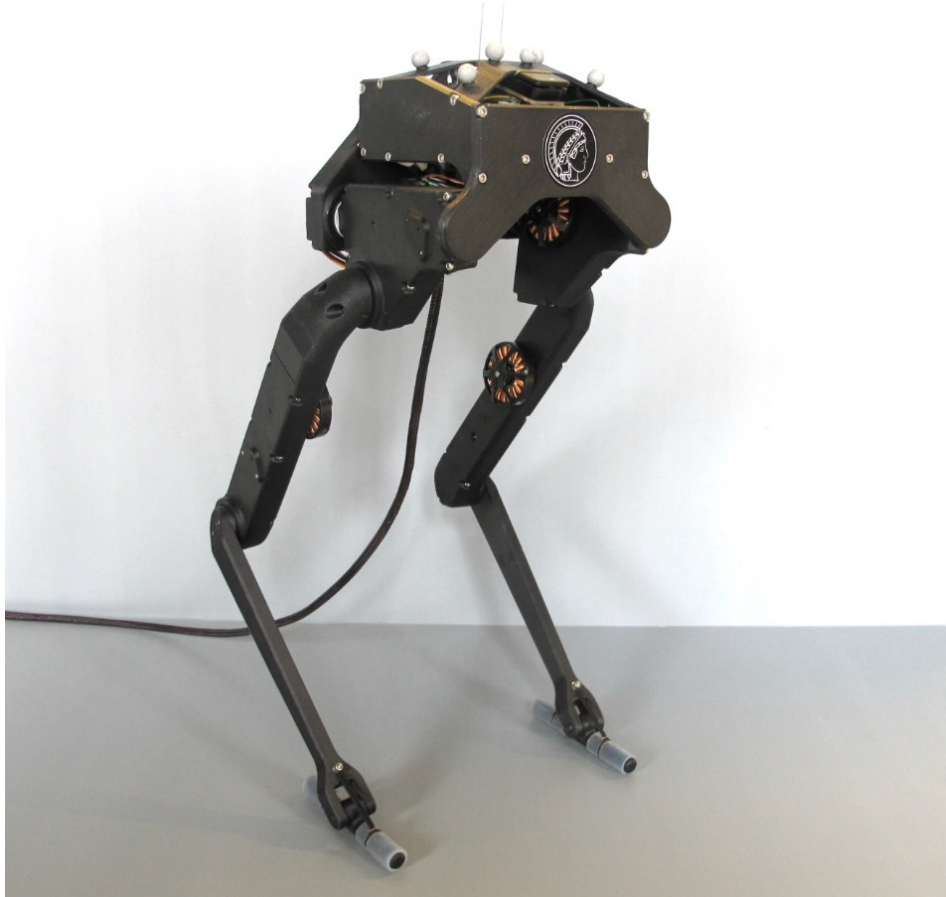
65cm high jump
(kino-dynamic planner + proposed controller)



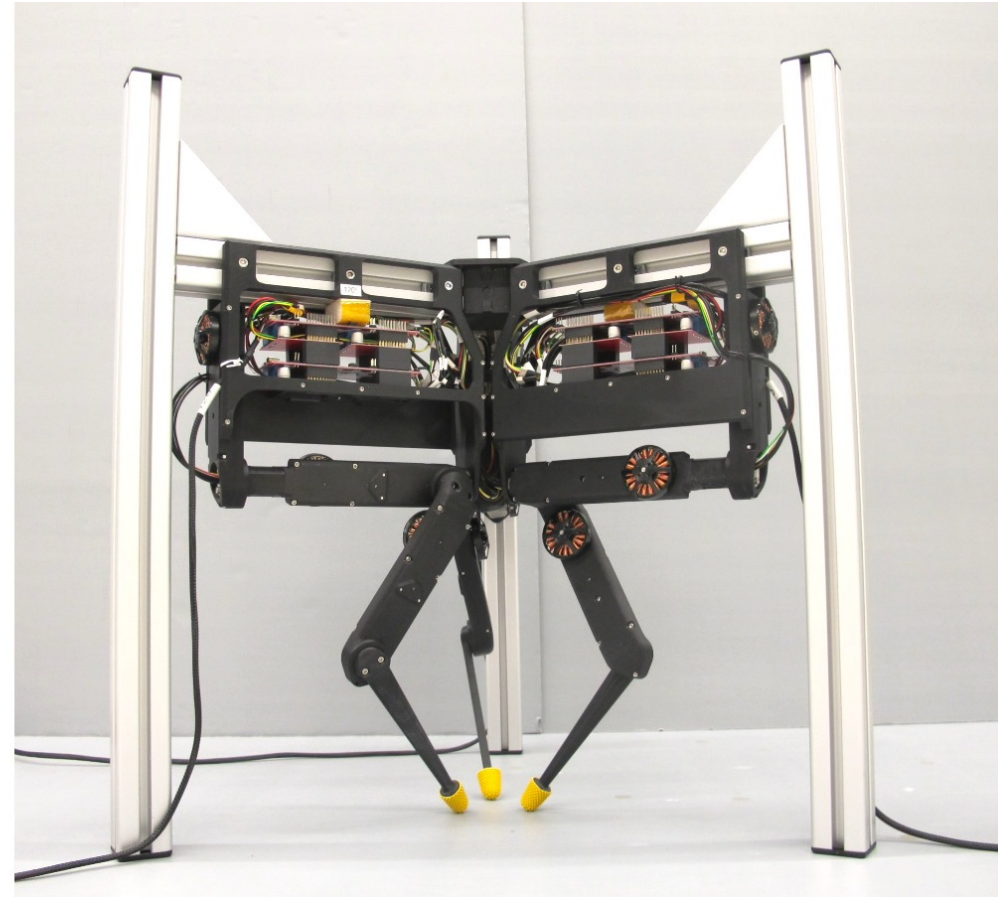
F. Grimmering¹, A. Meduri, M Khadiv, J. Viereck., M Wüthrich, M. Naveau, V. Berenz, S. Heim, F. Widmaier, T. Flayols, J. Fiene, A. Badri-Spröwitz, L. Righetti, *An Open Torque-Controlled Modular RobotArchitecture for Legged Locomotion Research*, IEEE RA-L 5 (2), 3650-3657, 2020



Autres prototypes utilisant le même module d'actionnement



Bolt



TriFinger

Contribution à la co-conception

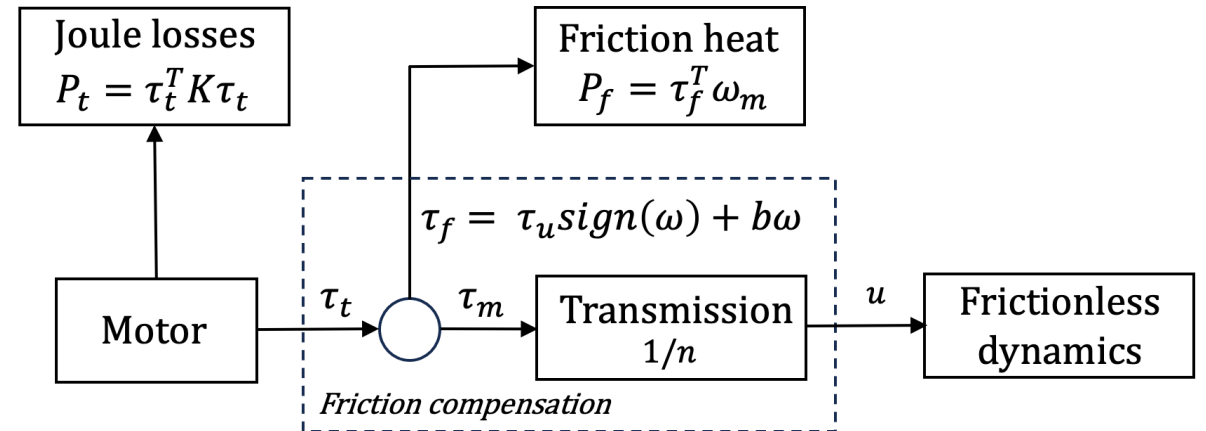
Développement d'un cadre générique

Objectif : Optimiser la conception des robots au regard de tâches spécifiques.

Coût : Puissance dissipée durant la tâche

Paramètres physiques considérés : \mathcal{D}

- Dimension de la structure, masses et inerties
- Capacités de l'actionneur (couples et vitesses)
- Rapport de transmission η
- Masse m_m et inertie I_m des moteurs



Paramètres de commande à optimiser \mathcal{C} : (par exemple : durée des phases de la tâche, gains)

Problème d'optimisation augmenté par les variables de décision + coût terminal garantissant la bonne réalisation de la tâche

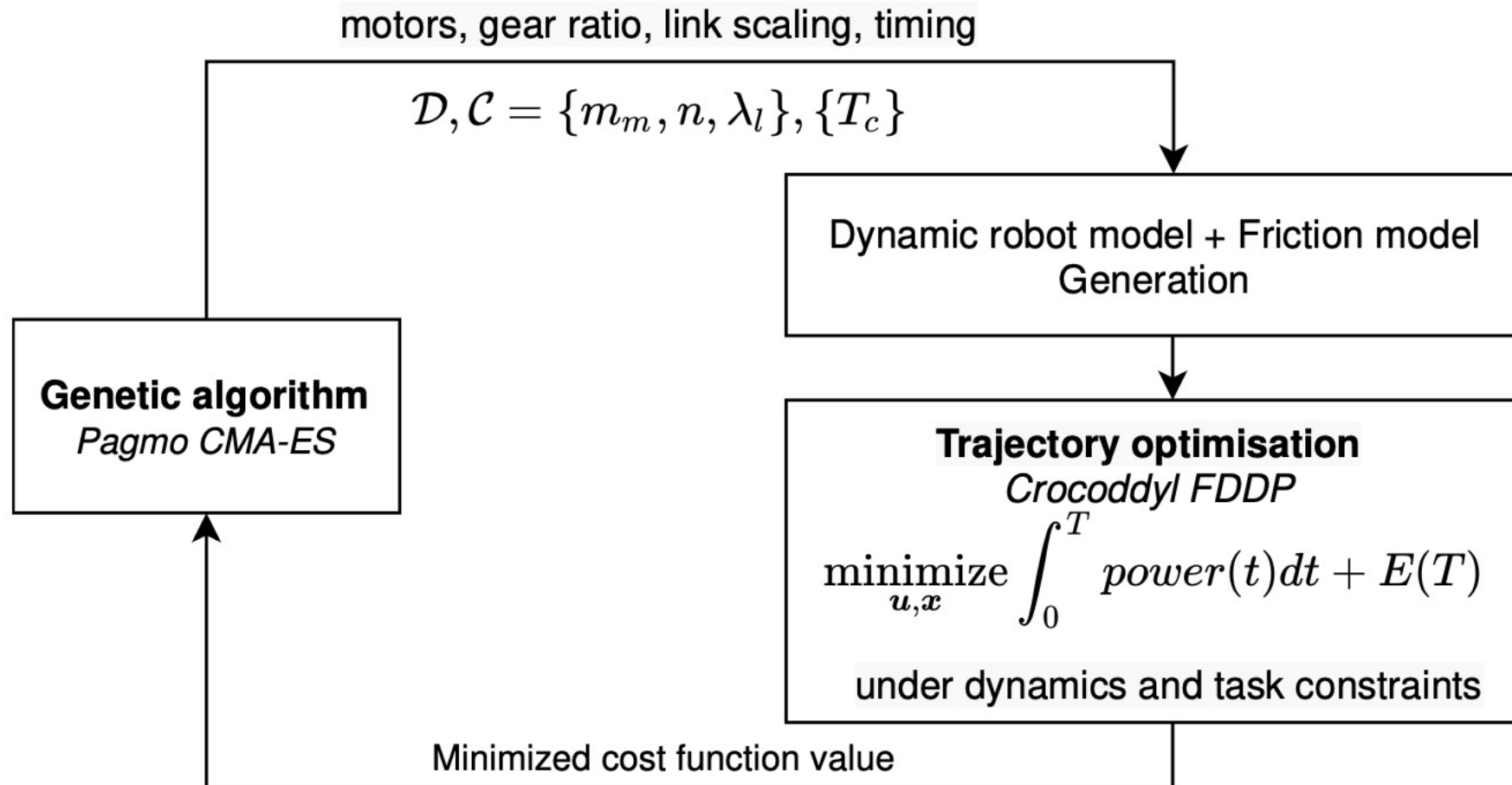
$$\underset{\mathbf{x}, \mathbf{u}, \mathcal{D}, \mathcal{C}}{\text{minimize}} \quad \Phi(\mathbf{x}_T, \mathbf{u}_T, \mathcal{D}, \mathcal{C}) + \int_{t_0}^T \mathcal{L}(\mathbf{x}(t), \mathbf{u}(t), \mathcal{D}, \mathcal{C}) dt$$

$$\dot{\mathbf{x}}(t) = f(t, \mathbf{x}(t), \mathbf{u}(t), \mathcal{D}, \mathcal{C})$$

$$\text{subject to} \quad h(t, \mathbf{x}(t), \mathbf{u}(t), \mathcal{D}, \mathcal{C}) \leq 0$$

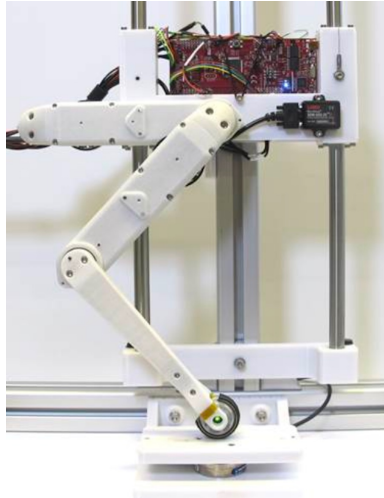
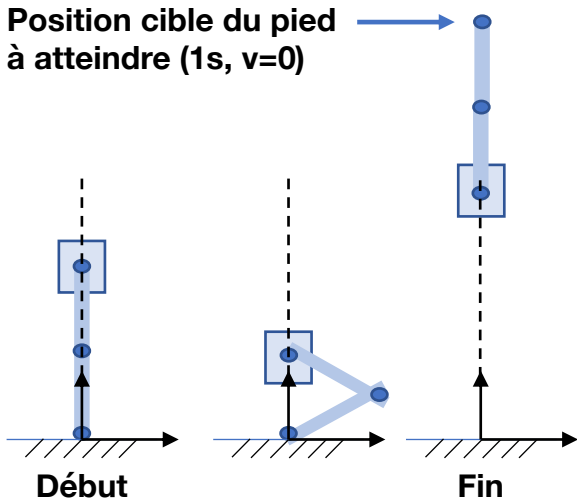
$$g(t_0, T, \mathbf{x}(t_0), \mathbf{x}(T), \mathcal{D}) \leq 0$$

Schéma d'optimisation à deux niveaux



Exemple : Saut vertical avec une patte

Position cible du pied à atteindre (1s, v=0)



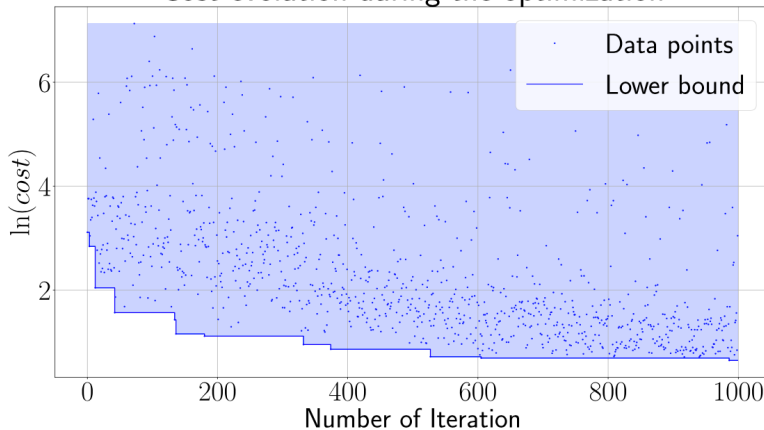
Intervallages des parameters

- $m_m \in [0.05, 1] [kg]$ (nominal $0.37kg$)
- $n \in [3, 20]$ (nominal 9)
- $\lambda_l \in [0.8, 1.2]$ (segment nominal $16cm$)
- $T_c \in [0.2, 0.8] [s]$ (fin de la phase de contact)

Coût (Solveur DDP)

- Courant : Puissance dissipée + pénalités pondérées (contraintes)
- Terminal : Réalisation de la tâche

Cost evolution during the optimization



Overall Energy Minimization

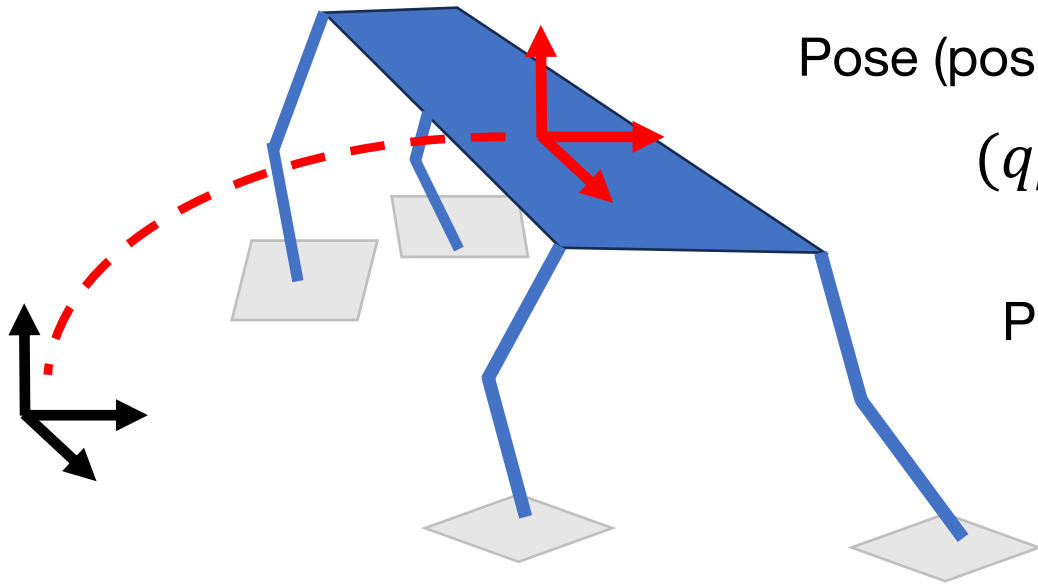
Parameters	all fixed, $T_c = 0.7s$	λ_l, T_c fixed	all free
Mech. energy	0.96 J	1.32 J	1.47 J
Joint friction	0.87 J	1.60 J	0.38 J
Joule dissipation	4.37 J	1.87 J	0.53 J
Total	6.19 J	4.79 J	2.38 J

Optimal solution

$$\left\{ \begin{array}{l} m_m = 9.85 \cdot 10^{-2} [kg] \\ n = 4.5 \\ \lambda_l = (0.83, 0.99) \\ T_c = 0.729 [s] \end{array} \right.$$

Commande de la locomotion

Variables d'état



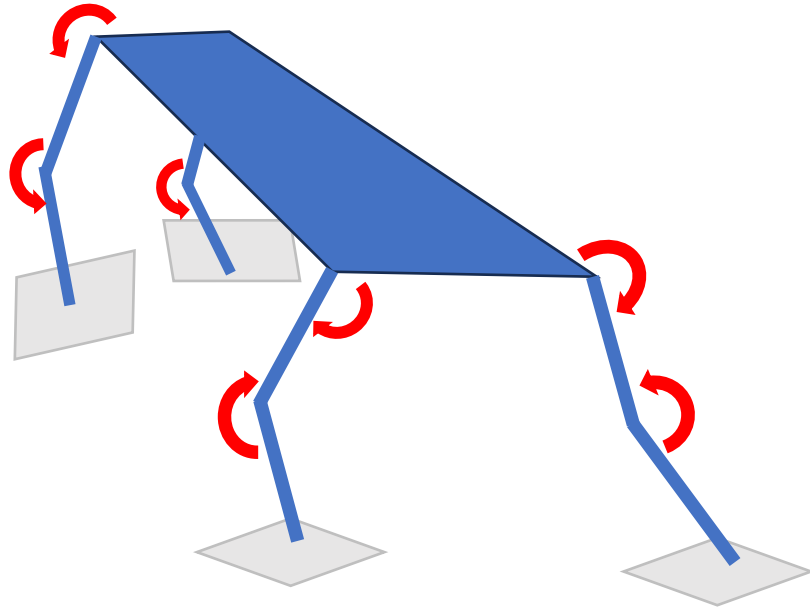
Pose (position, orientation) et vitesse de la base

$$(q_b, \dot{q}_b) \in \mathbb{R}^{12}$$

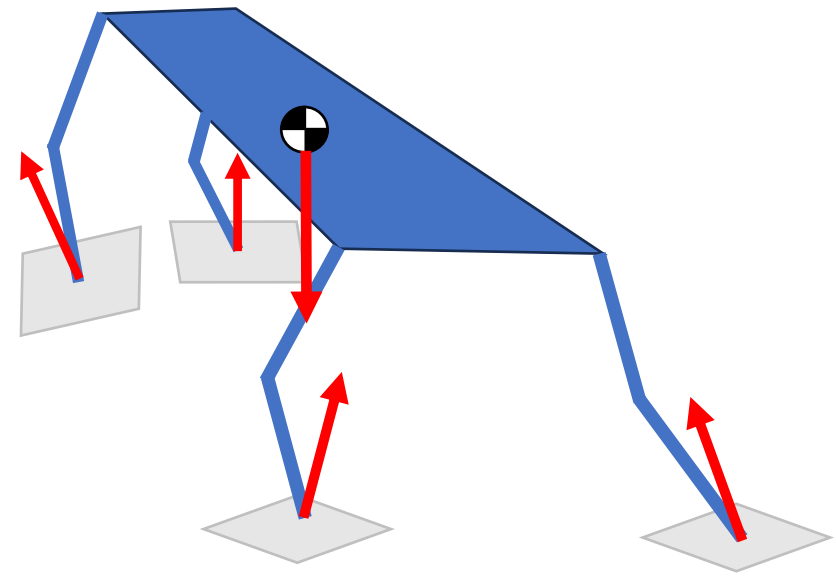
Positions et vitesses articulaires

$$(q_a, \dot{q}_a) \in \mathbb{R}^{24}$$

Modèle dynamique



Partie actionnée



Partie sous actionnée

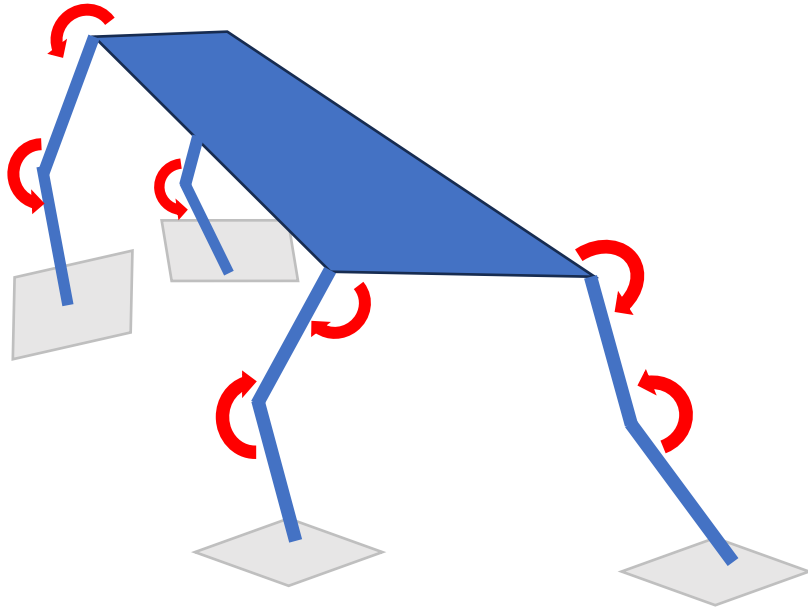
Problème de Commande

- Déterminer les couples articulaires permettant de réaliser une tâche donnée
- Le contrôle de la locomotion nécessite l'anticipation
- Idéalement : Résoudre un problème d'optimisation prédictive (MPC) pour le modèle complet du robot

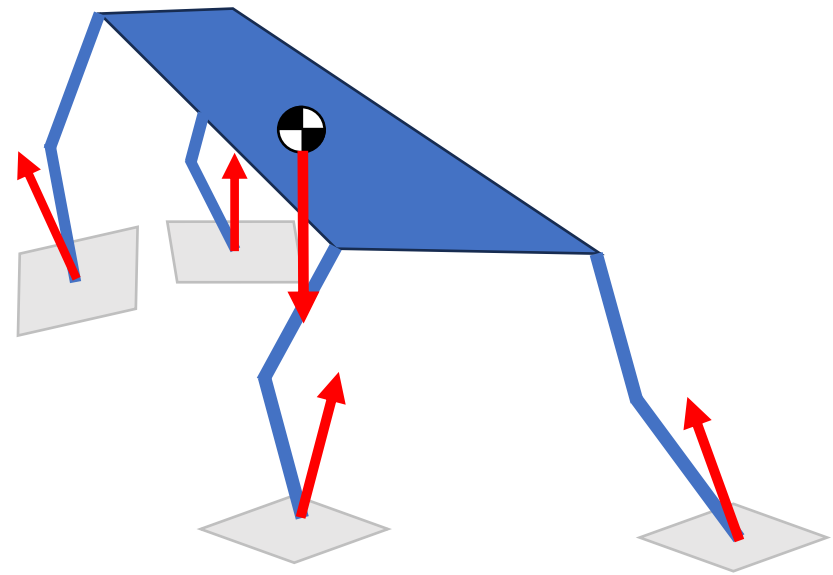
→ La difficulté est de traiter la complexité

$$\begin{cases} M_1(q)\ddot{q} + N_1(q, \dot{q})\dot{q} + G_1(q) = T_1(q)u + S_1^T(q)\lambda & \text{dynamique actionnée} \\ M_2(q)\ddot{q} + N_2(q, \dot{q})\dot{q} + G_2(q) = S_2^T(q)\lambda & \text{dynamique sous-actionnée} \\ g(q, \dot{q}, u) \leq 0 & \\ h(q, \dot{q}, u) = 0 & \text{contraintes} \end{cases}$$

Modèle dynamique

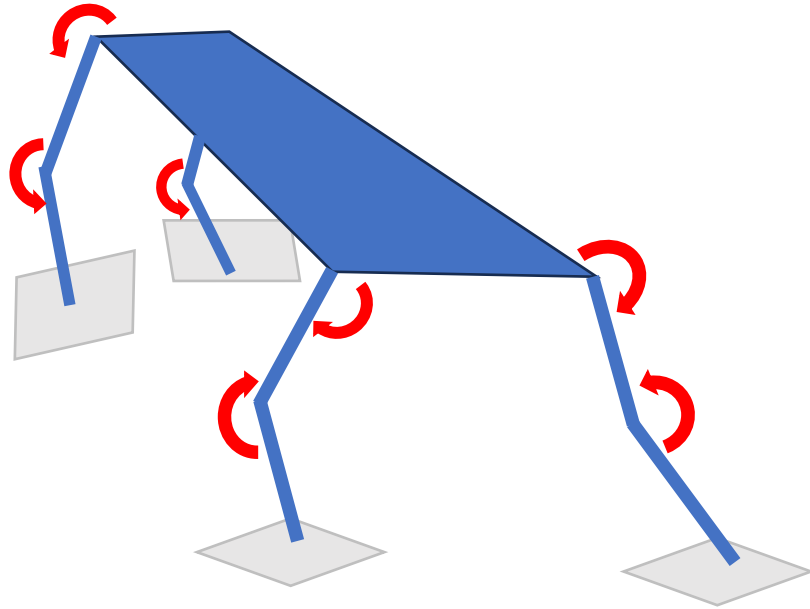


Partie actionnée

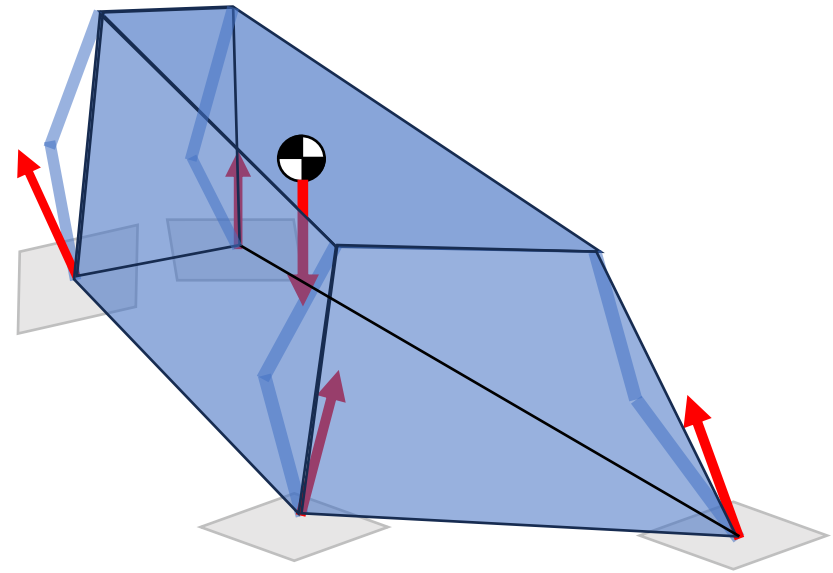


Partie sous actionnée

Modèle Centroidal

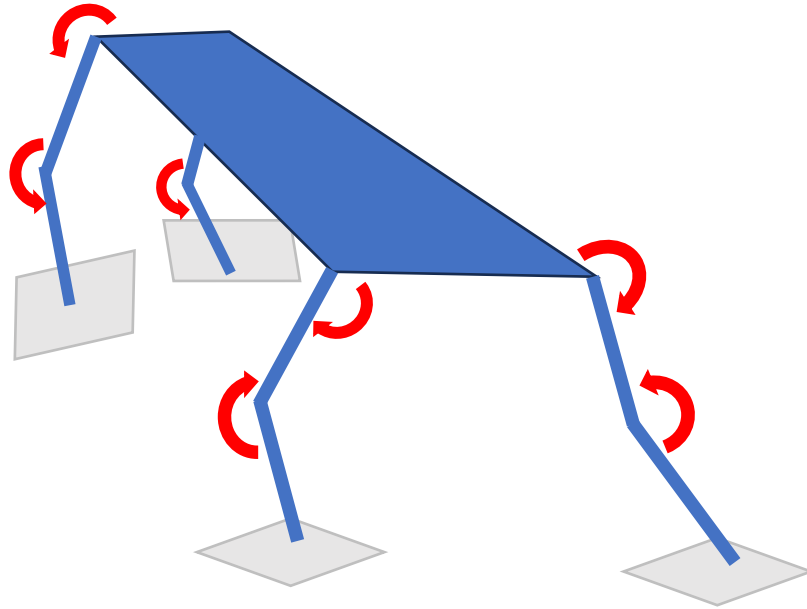


Partie actionnée



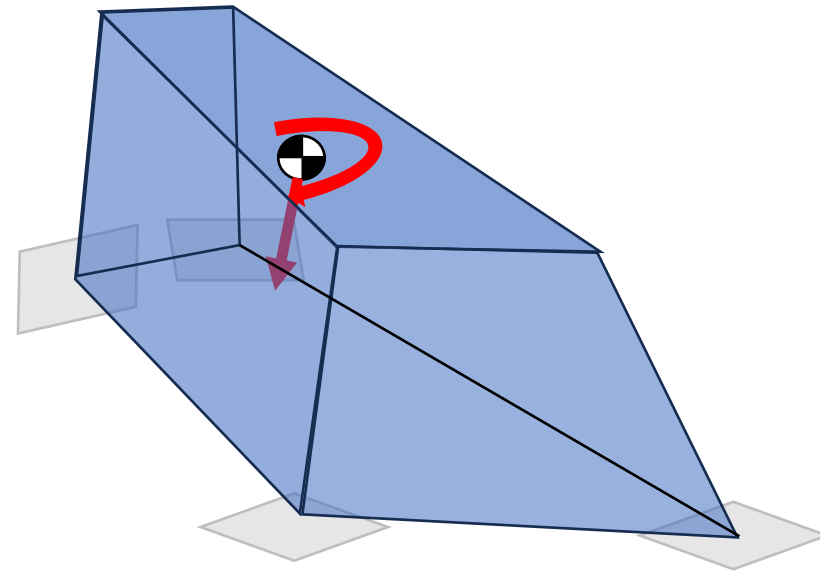
Partie sous actionnée

Modèle Centroïdal



Partie actionnée

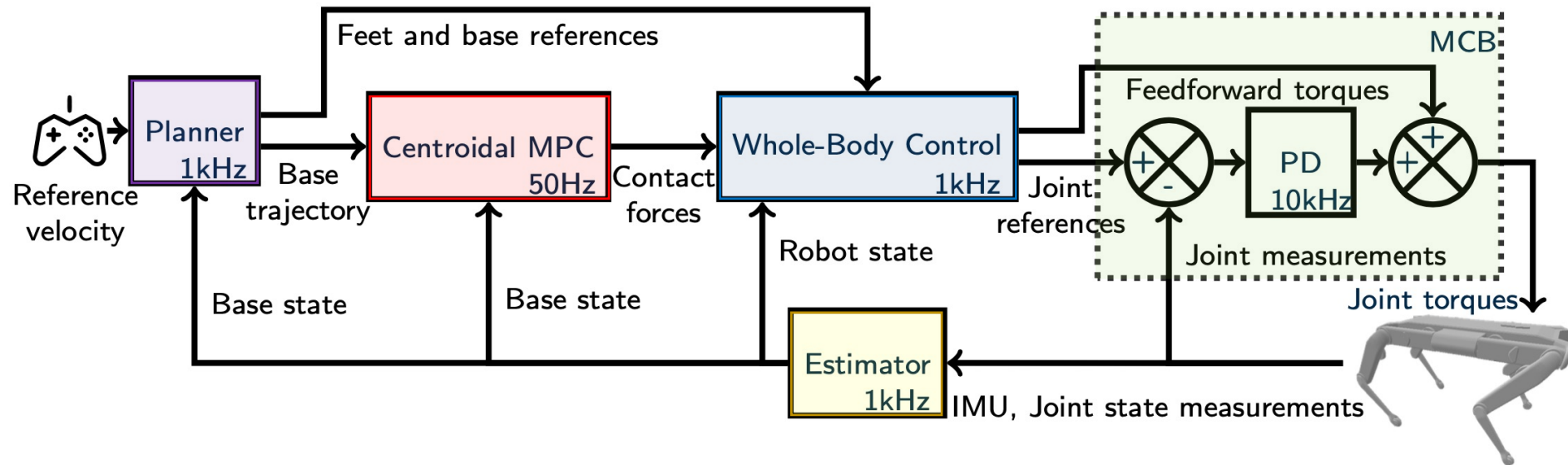
Centroidal model [Orin et al. 2013],
“A key template model for legged-robot control”



Partie sous actionnée

Locomotion basée sur l'optimisation de trajectoire sur un horizon prédictif (MPC)

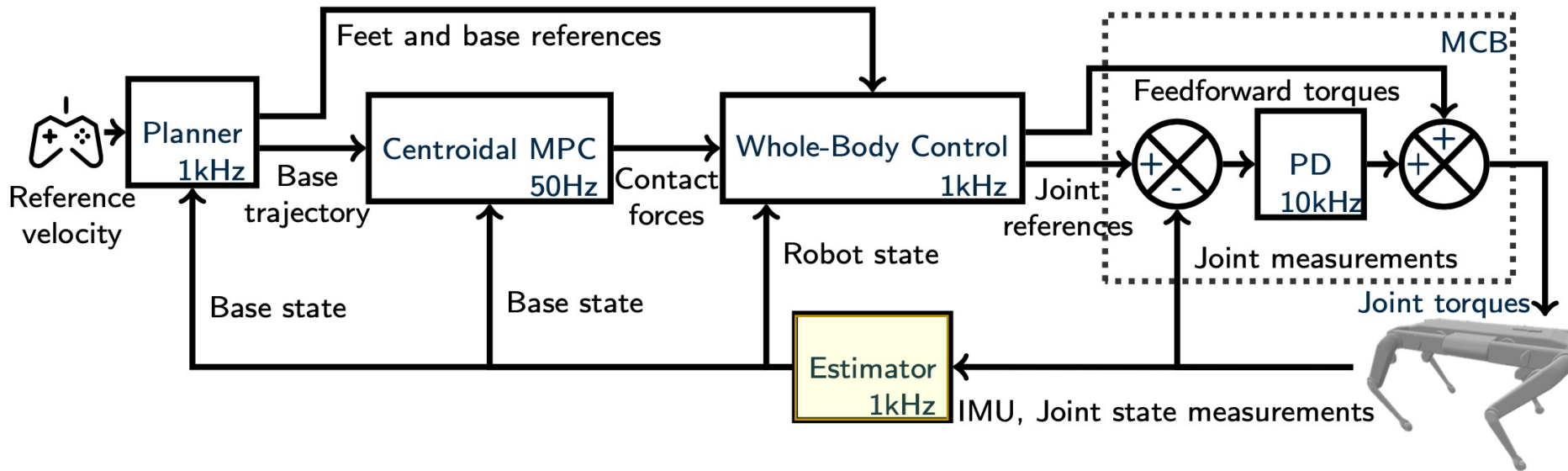
Première Approche : Contrôle prédictif basé-modèle (MPC)



D. Kim, J. Di Carlo, B. Katz, G. Bleedt, and S. Kim, "Highly dynamic quadruped locomotion via whole-body impulse control and model predictive control," arXiv.org, Sep. 2019

P.A. Léziart, T. Flayols, F. Grimmering, N. Mansard, P. Souères, *Implementation of a Reactive Walking Controller for the New Open-Hardware Quadruped Solo-12*, IEEE ICRA, Xi'an, China, 2021

Estimation



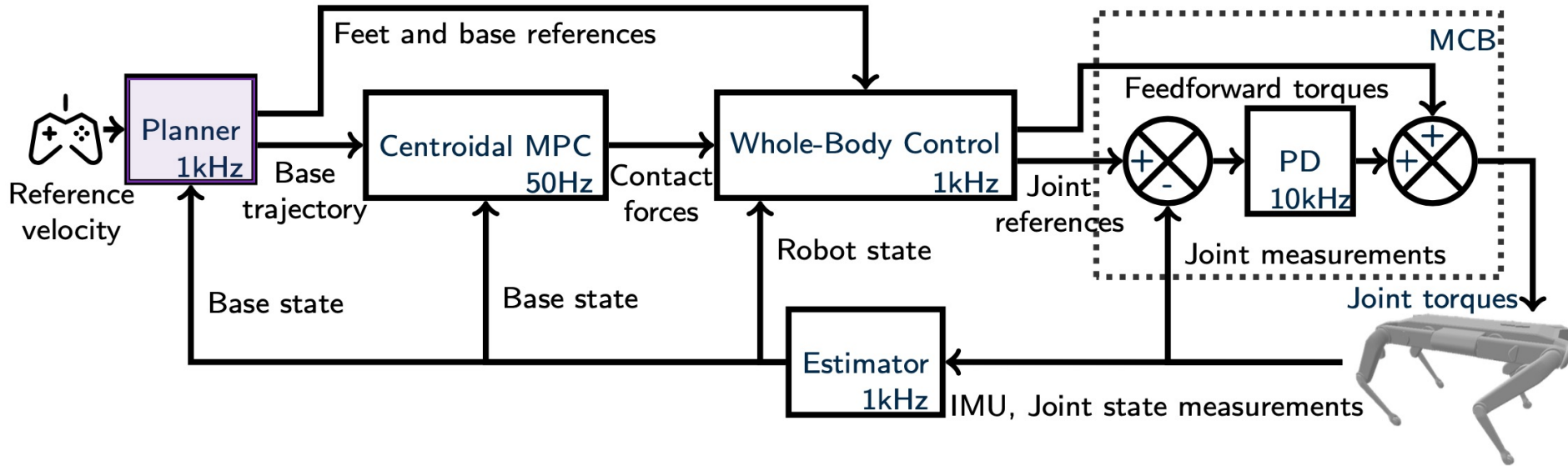
Objectif : Estimer l'état du robot à partir des mesures des capteurs (proprioception + IMU)

- Vitesse et position/pieds de la base : Intégration de l'accélération linéaire (IMU), cinématique directe, géométrie directe.

→ **Fusion multi-capteur pour obtenir des estimations de bonne qualité, non biaisées**

Fitrage complémentaire ou de Kalman + filtrage des oscillations connues liées à l'allure

Planificateur de pas



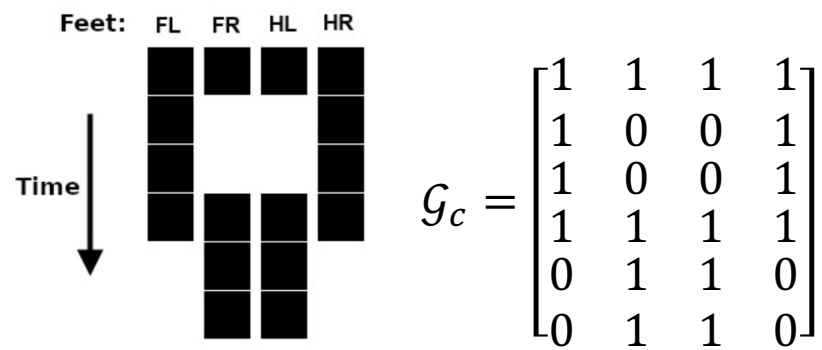
Objectif : Planifier la sequence de pas pour suivre la vitesse de consigne

→ Placement des pieds pour produire les forces de reaction nécessaires

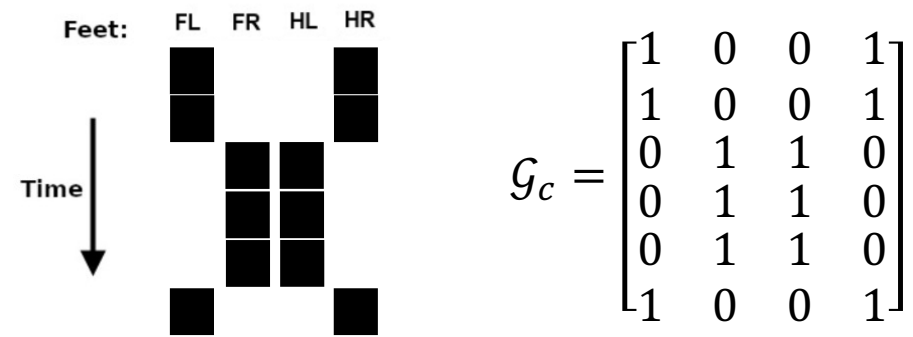
→ Planification la trajectoire du pied en vol

Modèle d'allure prédéfini

Séquence de contacts discrets des pieds avec le sol avec cadence fixée



Trot marché

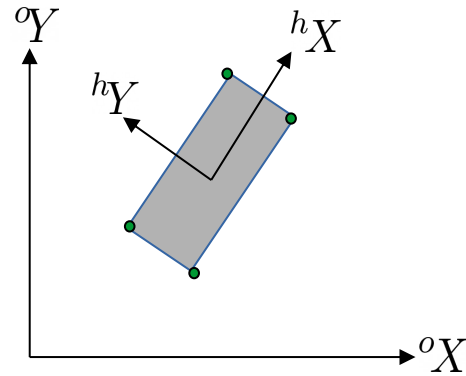


Trot

Placement des pas basé sur l'heuristique de Raibert

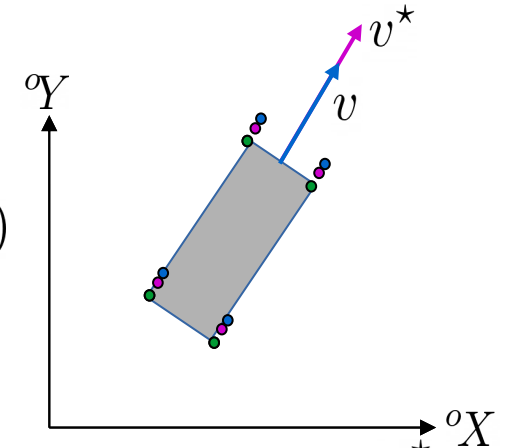
$$r^* = r_{sh} + r_{sym} + r_{fb} + r_c$$

- Terme d'épaule



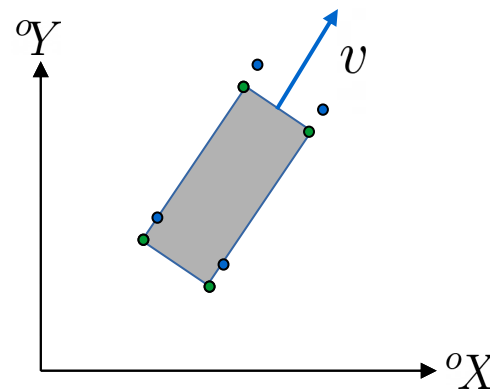
- Terme de feedback

$$r_{fb} = k(v - v^*)$$



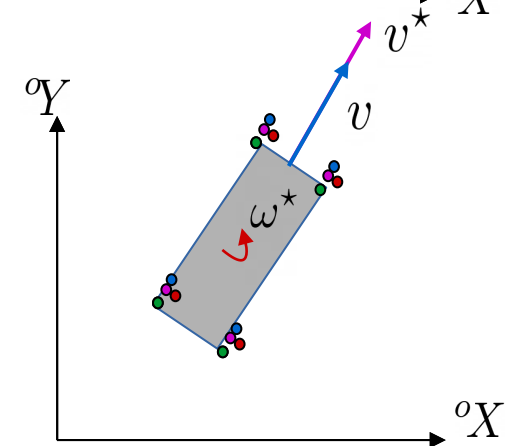
- Terme de Symétrie

$$r_{sym} = \frac{t_{stance}}{2} v$$

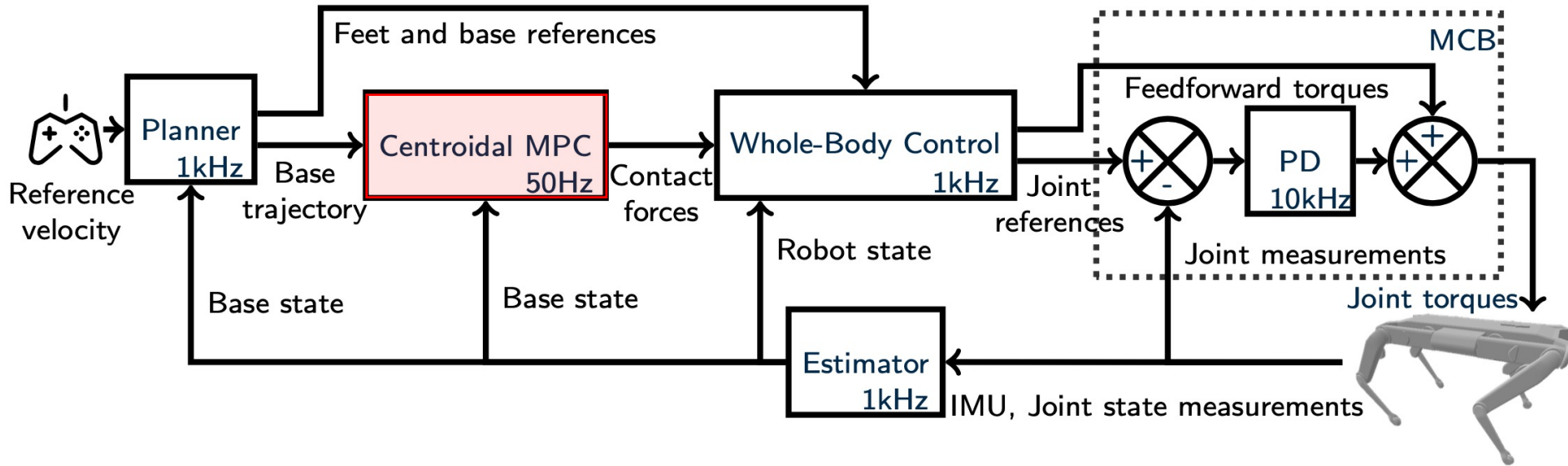


- Terme centrifuge

$$r_c = \frac{1}{2} \sqrt{\frac{h}{g}} v \times \omega^*$$

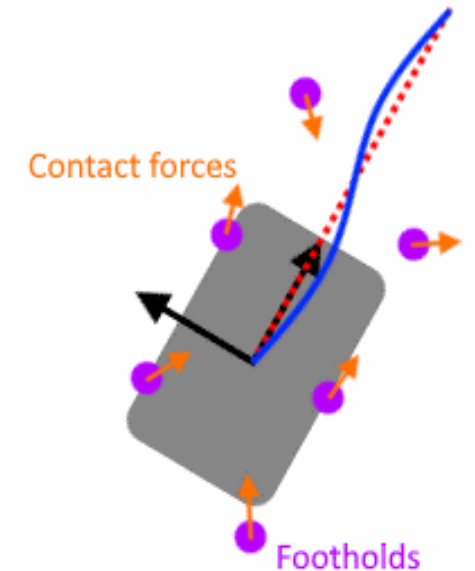


Contrôle Predictif du modèle centroïdal

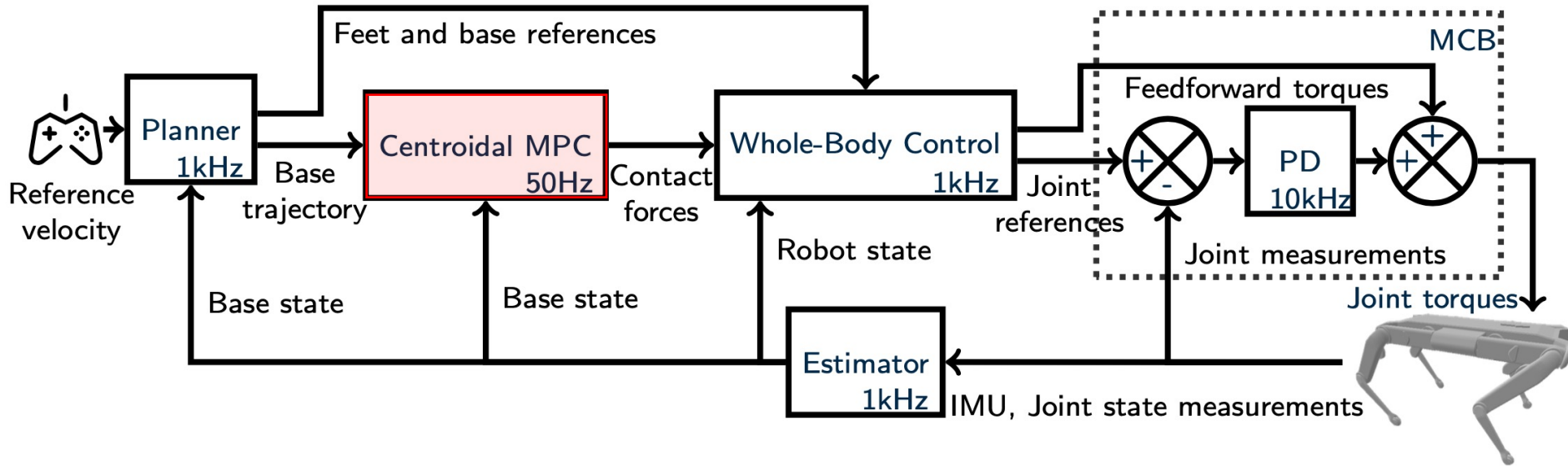


Objectif : Produire la trajectoire centroïdale sur un horizon prédictif à partir des contacts planifiés pour suivre la vitesse de consigne.

- Trajectoire de référence par intégration de la vitesse de consigne
- Optimisation de la séquence de forces de contact pour réaliser le suivi



Contrôle Predictif du modèle centroïdal



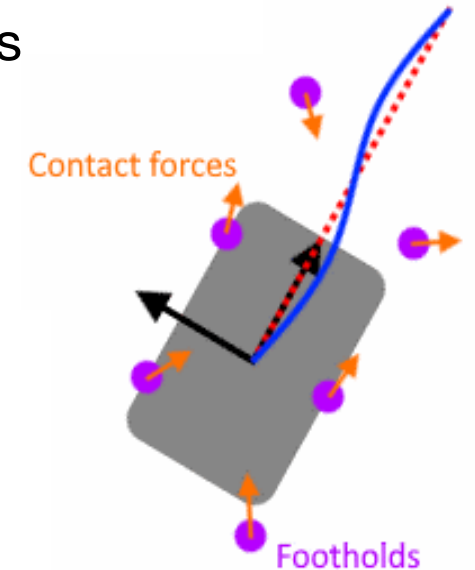
Coûts quadratiques sur l'erreur de suivi de la trajectoire et la norme des forces

$$\min_{\{\mathbf{X}\}, \{\mathbf{F}\}} \sum_{t=0}^T \ell_t(\mathbf{X}_t, \mathbf{F}_t | r_t^*) + \ell_T(x_T)$$

$$\text{s.t. } \mathbf{X}_{t+1} = \mathcal{H}(\mathbf{X}_t, \mathbf{F}_t | r_t^*)$$

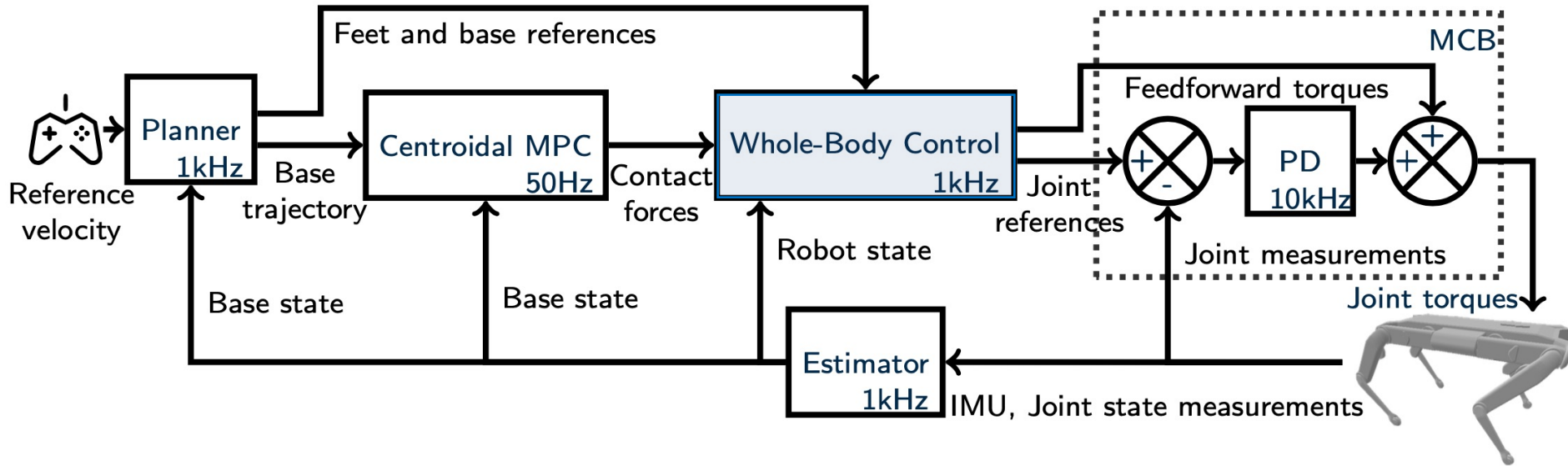
$$\mathbf{X}_t \in \mathcal{F}$$

$$\mathbf{F}_t \in \mathcal{K}$$



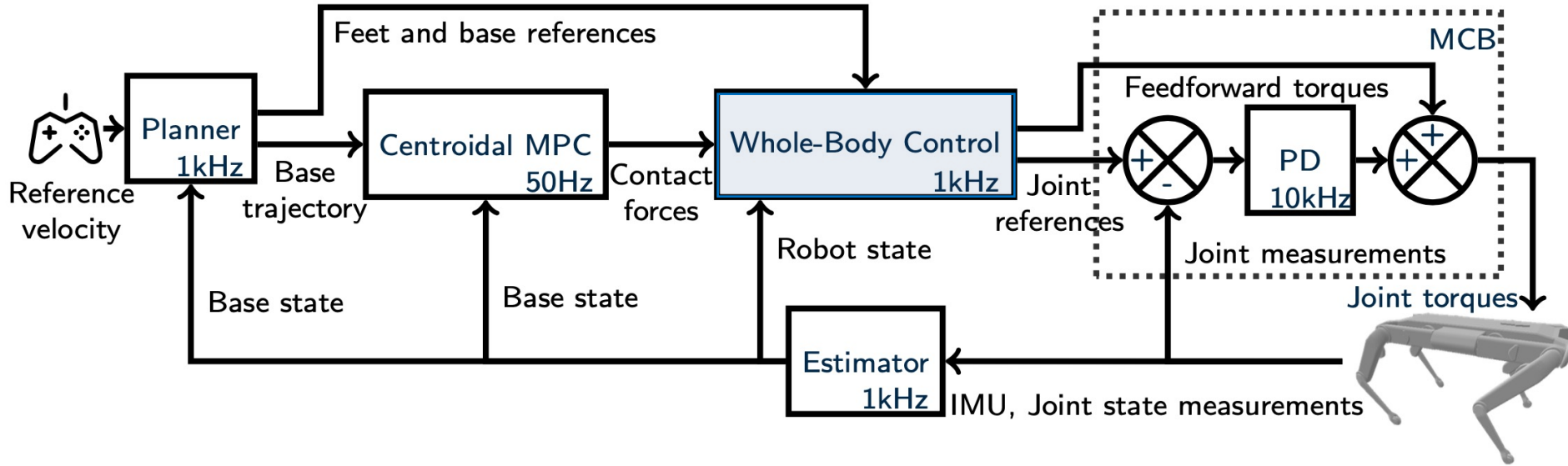
Contraintes exprimées par des log-barrières dans le coût

Contrôle du corps-complet



- **Objectif** : Produire les couples et les trajectoires articulaires de consigne pour la commande en impédance de bas niveau, afin de suivre les profils de forces données par le MPC et les trajectoires des pieds.

Contrôle du corps-complet



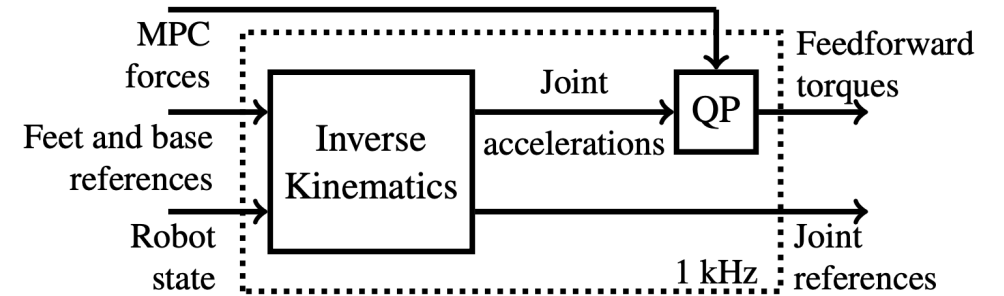
1^{ère} approche : Résolution de la dynamique inverse (5 tâches pondérées) → Instabilité

2^{ème} approche : Compromis entre cinématique inverse et suivi des forces

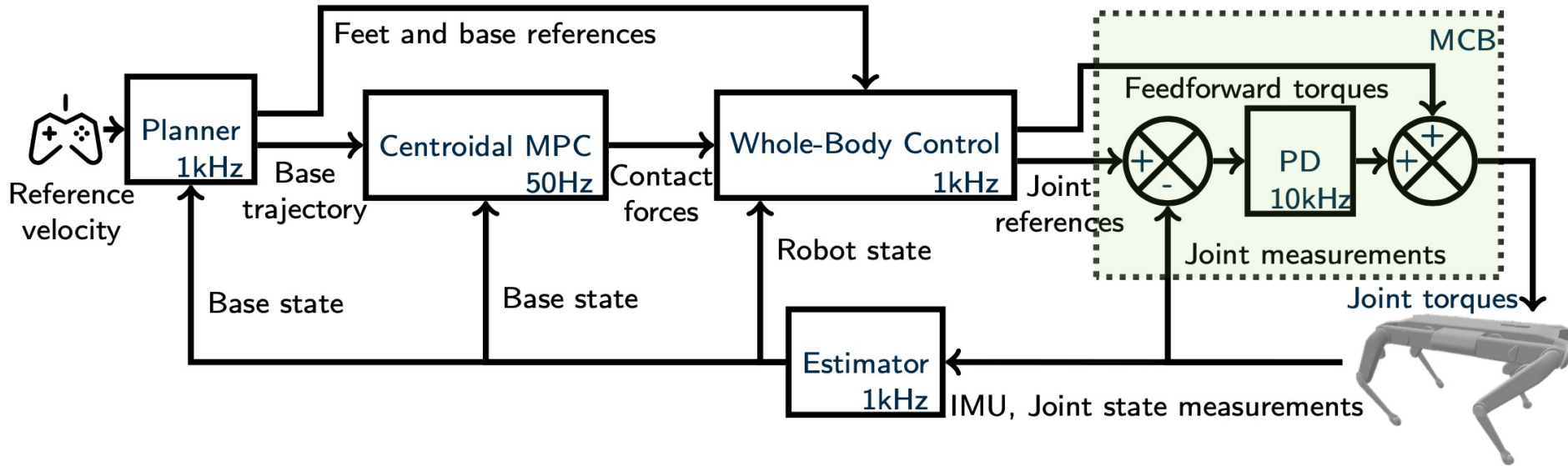
$$\min_{\delta_{\ddot{\mathbf{q}}}, \delta_{\mathbf{F}}} \delta_{\ddot{\mathbf{q}}}^T Q_1 \delta_{\ddot{\mathbf{q}}} + \delta_{\mathbf{F}}^T Q_2 \delta_{\mathbf{F}}$$

$$\text{s.t. } \mathbf{F}_{MPC} + \delta_{\mathbf{F}} \in \mathcal{K}$$

$$S(\mathbf{M} \left(\begin{bmatrix} \ddot{\mathbf{q}}_{IK,u} \\ \ddot{\mathbf{q}}_{IK,a} \end{bmatrix} + \begin{bmatrix} \delta_{\ddot{\mathbf{q}}} \\ 0 \end{bmatrix} \right) + \mathbf{g} + \mathbf{C}) = S\mathbf{J}_c^T (\mathbf{F}_{MPC} + \delta_{\mathbf{F}})$$



Commande en impédance de base niveau



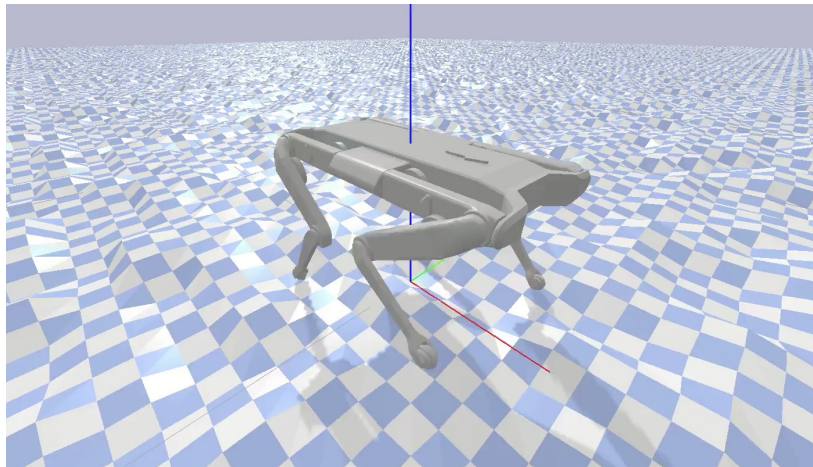
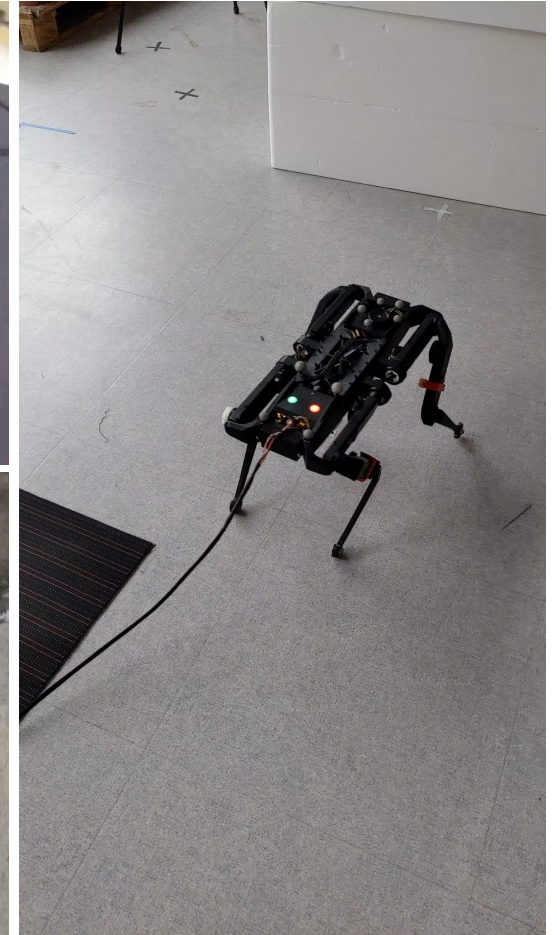
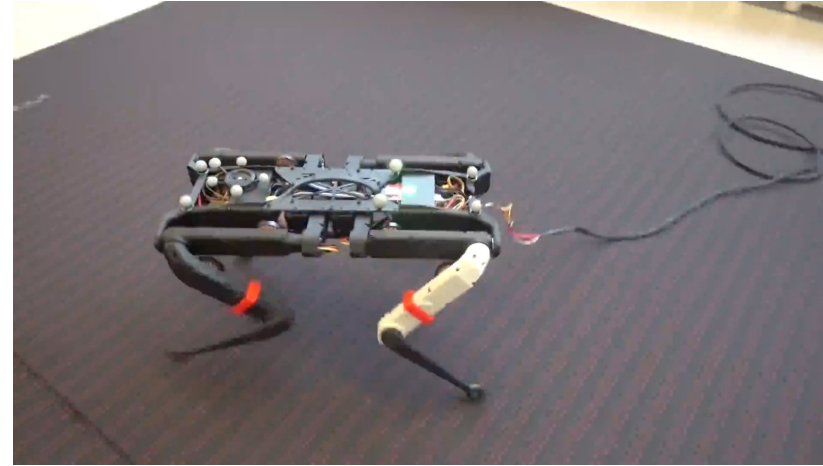
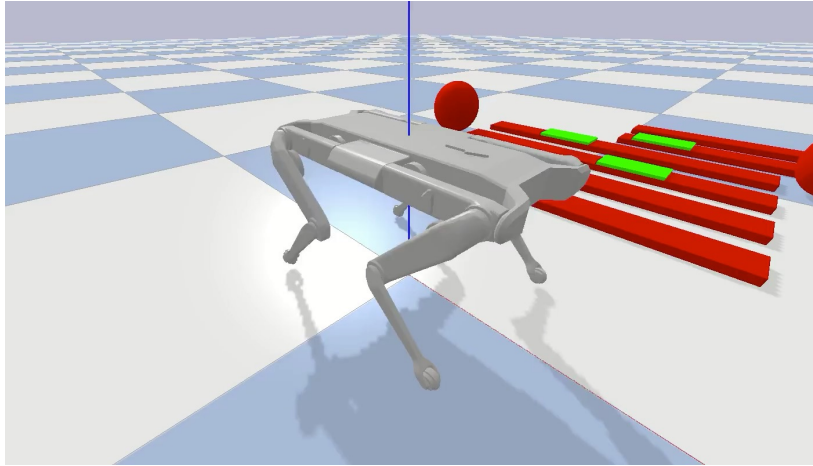
Asservissement haute fréquence des moteurs par des boucles en courant

$$\tau_{mot}^* = \tau^* + K_p(\mathbf{q}_a^* - \mathbf{q}_a) + K_d(\dot{\mathbf{q}}_a^* - \dot{\mathbf{q}}_a) \quad I_{mot}^* = \frac{\tau_{mot}^*}{K_\tau R}$$

$$K_\tau = 0.025 \text{ Nm/A} \text{ et rapport de reduction } R = 9$$

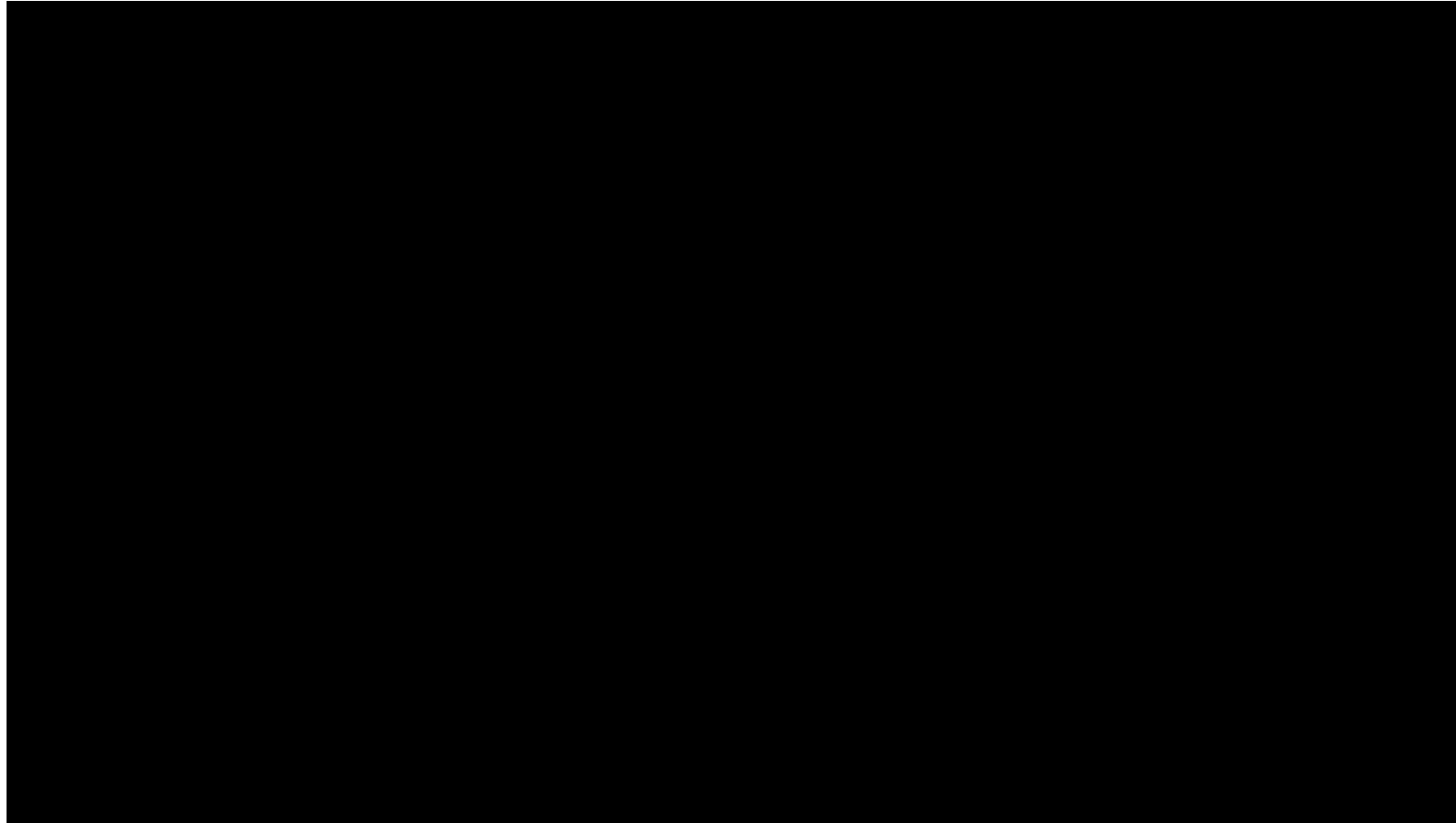
P.A. Léziart, T. Flayols, F. Grimmering, N. Mansard, P. Souères, Implementation of a Reactive Walking Controller for the New Open-Hardware Quadruped Solo-12, IEEE ICRA, Xi'an, China, 2021

Résultats expérimentaux



P.A. Léziart, T. Flayols, F. Grimmering, N. Mansard, P. Souères. Implementation of a Reactive Walking Controller for the New Open-Hardware Quadruped Solo-12, IEEE ICRA, Xi'an, China, May 2021

Comparaison de différents schémas de MPC



P.A. Léziart, P. A. Léziart, T. Corbères, T. Flayols, S. Tonneau, N. Mansard and P. Souères, "Improved Control Scheme for the Solo Quadruped and Experimental Comparison of Model Predictive Controllers," IEEE RA-L 2022,

Approche mixte MPC/Apprentissage par Renforcement (RL)

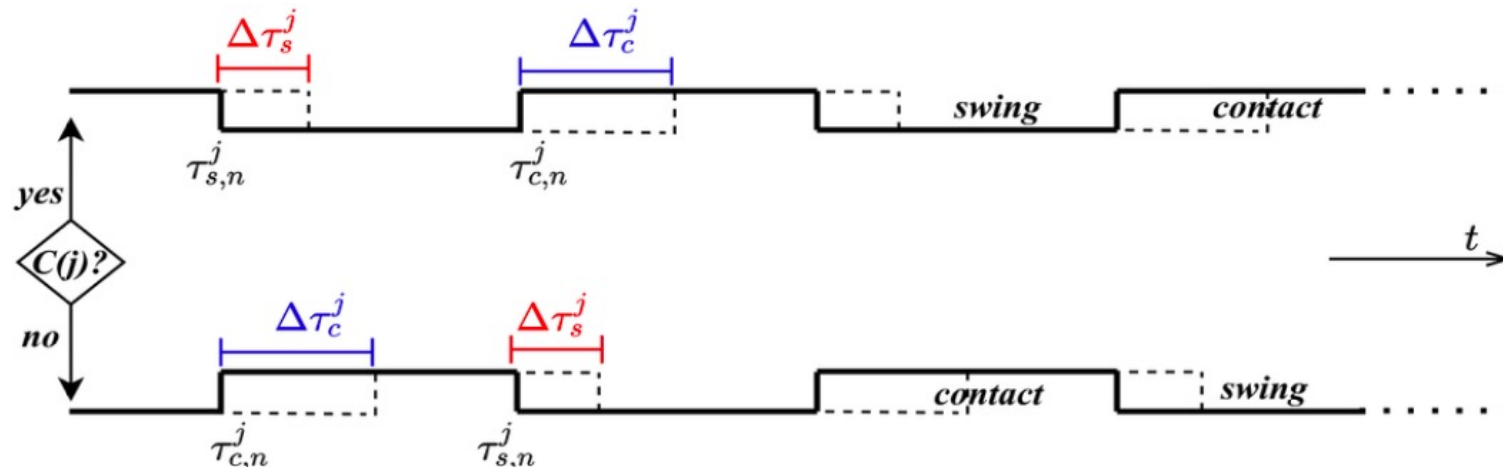
Idée : Insérer une part d'apprentissage dans l'architecture MPC

Objectif : Utiliser l'apprentissage par renforcement (RL) pour adapter temporellement l'allure (temps du cycle, phases de contact et de vol) afin d'améliorer les performances

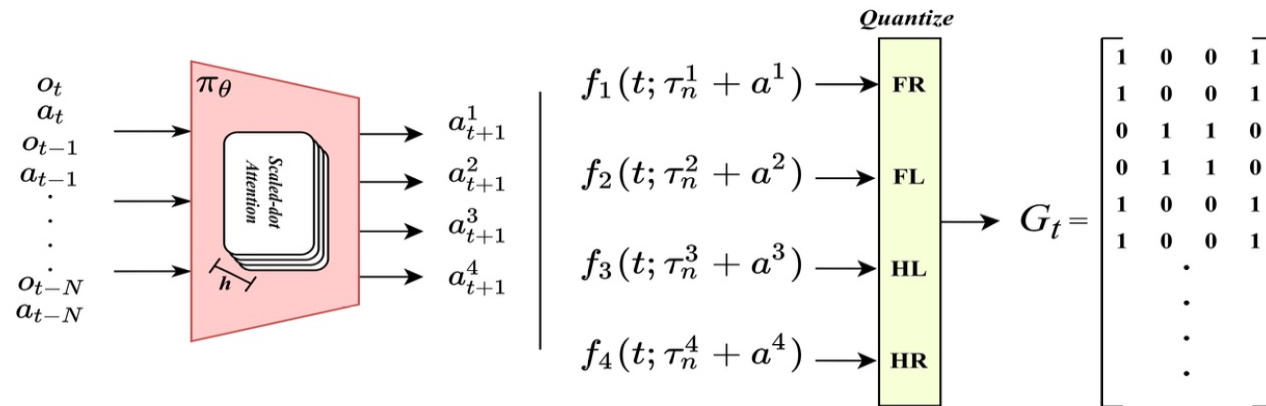
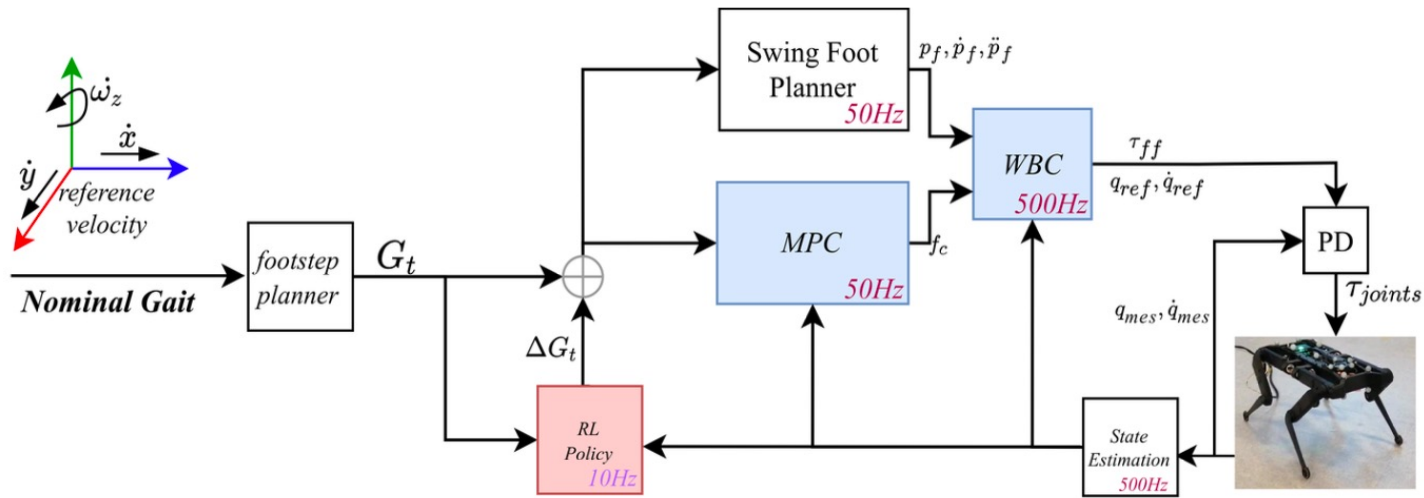
→ stabilité vs économie d'énergie

- A partir d'un trot nominal, une action $a = (a^1, \dots, a^4) \in \mathcal{A}$ est un décalage des commutations contact/vol sur chaque pied

$$a^i = (\Delta\tau_s^i, \Delta\tau_c^i)$$

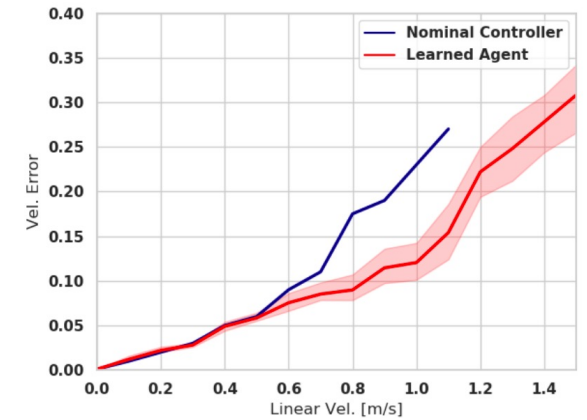
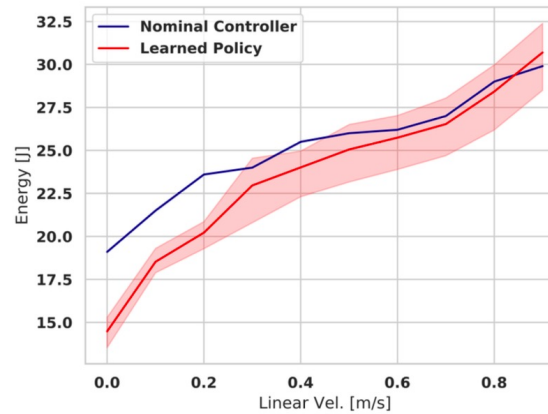
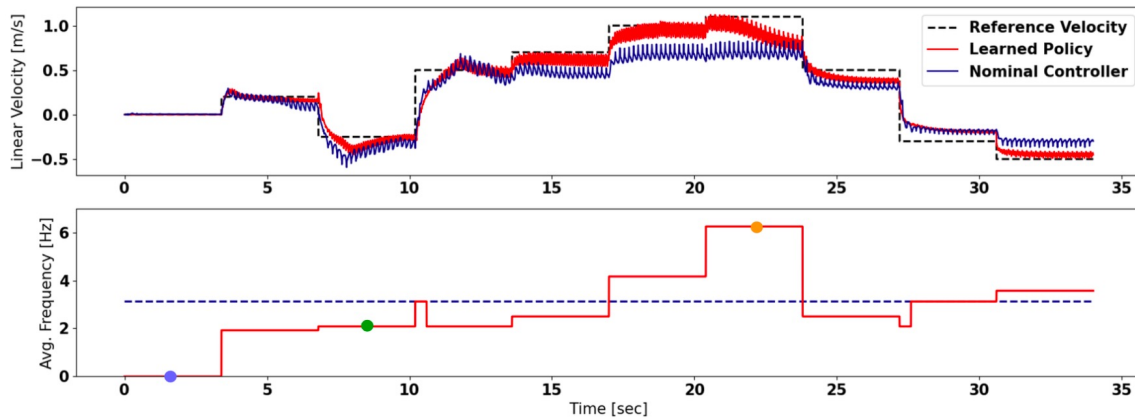


Adaptation basée-RL du schéma temporel de l'allure



Récompense :
$$R_t = c - \int_{t-T_{RL}}^t \|\dot{q}_{ref,v} - \dot{q}_{base,v}\|^2 dv - \lambda E_t$$

- La politique permet d'adapter le trot nominal à la vitesse de consigne, le changement de fréquence améliore la performance (meilleur suivi, économie d'énergie, stabilité)



Energie consommée et erreur de suivi en vitesse en fonction de v_{ref}

- Mais l'approche ne permet pas de produire des allures différentes. Un résultat similaire pourrait être obtenu en augmentant linéairement la fréquence en fonction de la vitesse de consigne
- Non suffisamment convaincant pour un déploiement sur le robot réel.

→ **Nouvel objectif** : Apprendre des politiques de locomotion basées RL de bout en bout

Apprentissage par renforcement de la locomotion

Problème : Développer une approche basée RL pour apprendre au quadrupède à suivre la vitesse de consigne

Etat : Orientation et vitesse de la base, positions vitesses articulaires + historique (N)

$$\{q_{t-j}^e = (q_{t-j}^{target} - q_{t-j}) \in \mathbb{R}^{12}\}^{j=1..N}, \{\dot{q}_{t-j} \in \mathbb{R}^{12}\}^{j=1..N}, \{a_{t-j} \in \mathbb{R}^{12}\}^{j=1..N-1}$$

Actions et Politique :

→ Apprentissage difficile de la commande en couple (profils complexes, haute fréquence requise)

→ On cherche à apprendre une politique π donnant la variation des positions articulaires par rapport à une configuration nominale q_{init}

$$\pi_{\beta}(s_t) = \Delta_{q_t}^{\beta} \longrightarrow \text{cible articulaire : } q_t^{target} = q_{init} + \lambda_q \Delta_{q_t}^{\beta}$$

Le contrôleur calcule ensuite les couples $\tau_t = K_p(q_t^{target} - q_t) - K_d \dot{q}$

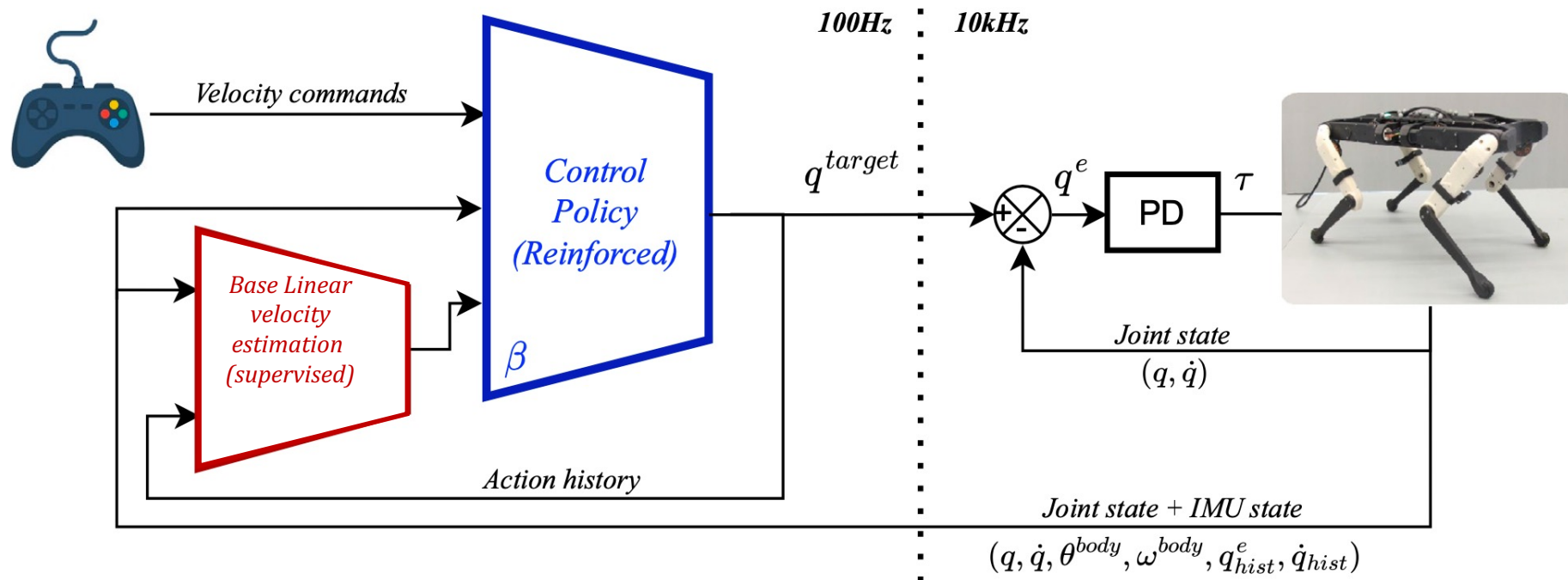
Reward : Some pondérée d'un terme de récompense et de plusieurs pénalités [Li et al. 2022]

- + Suivi de la vitesse de consigne: $r_{vel} = c_{vel} e^{-\|V^{cmd} - V_{x,y,wz}\|^2}$
- Dégagement du pied : $r_{clear} = c_{clear} \sum_{i=1}^4 (p_{z,i} - p_z^{max})^2 \|\dot{p}_{xy,i}\|^{0.5},$
- Glissement du pied : $r_{slip} = c_{slip} \sum_{i=1}^4 C_i \|\dot{p}_{xy,i}\|^2,$
- Attitude et vitesse verticale de la base : $r_{base} = c_{orn}(roll^2 + pitch^2) + c_{vz} V_z^2.$
- Puissance dissipée : $r_E = C_E \sum_{j=1}^{12} P_{Fj} + P_{Tj}$
- Positions articulaires : $r_{joint} = c_q \|q_t - q_{init}\|^2$
- Lissage du mouvement : $r_{smooth} = c_{a1} \|q_t^{target} - q_{t-1}^{target}\|^2 + c_{a2} \|q_t^{target} - 2q_{t-1}^{target} + q_{t-2}^{target}\|^2$

Apprentissage par renforcement de la locomotion

→ Politique codée par un réseau de neurone recevant l'état en entrée et donnant les actions en sortie.
Optimisation des paramètres du réseau par l'algorithme PPO (Proximal Policy gradient Objective)

→ Un second réseau est entraîné par apprentissage supervisé pour prédire la vitesse de la base



Randomisation de l'état et de la dynamique

Objectif : faciliter le transfert de la simulation vers le robot réel

Technique: Ajoût de bruit

→ Randomiser l'état est essentiel pour compenser le bruit des capteurs

→ Pour Solo12, randomiser les gains du contrôleur suffit pour rendre compte de la stochasticité du bas-niveau de commande

Observation noise	
θ_{body}	$U^3(-0.05, 0.05)$
ω_{body}	$U^3(-0.10, 0.10)$
v_{body}	$U^3(-0.10, 0.10)$
q	$U^{12}(-0.05, 0.05)$
\dot{q}	$U^{12}(-0.50, 0.50)$
Dynamics noise	
K_p	$U(-1.0, 3.0)$
K_d	$U(-0.1, 0.1)$

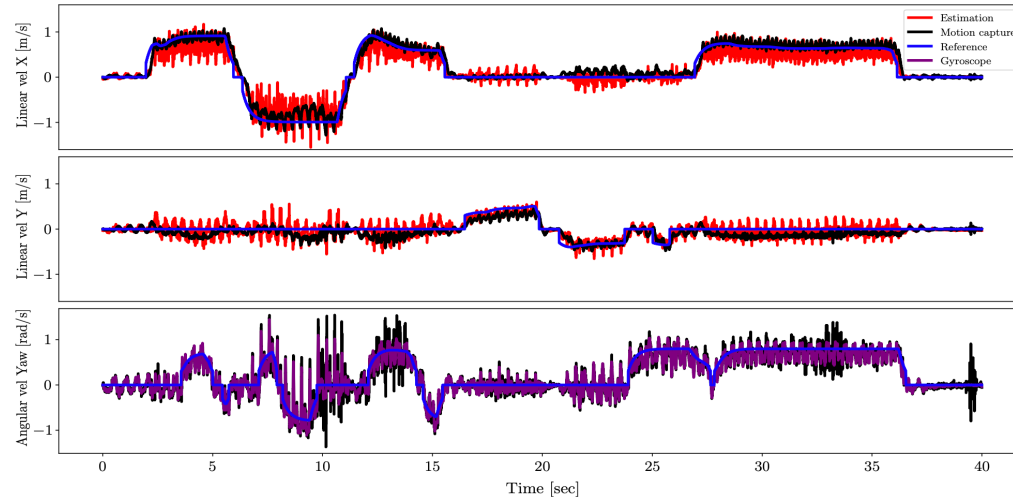
Curriculum

Objectif: faciliter l'apprentissage en introduisant les tâches les plus faciles au début puis en augmentant progressivement le niveau de difficulté.

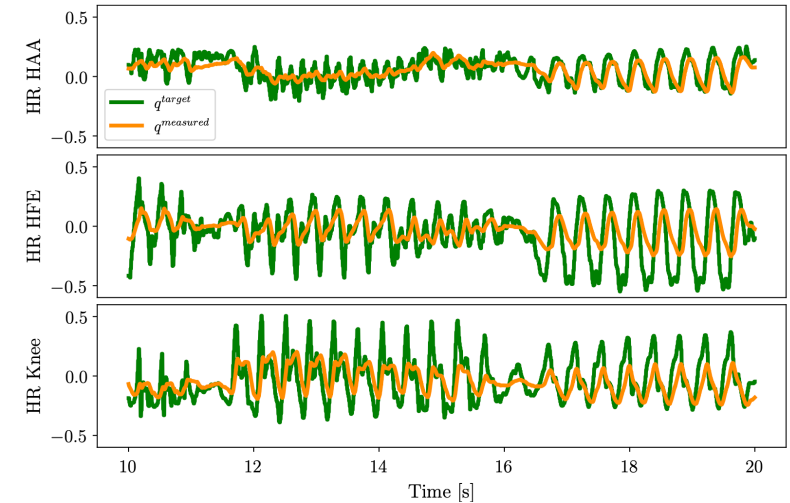
- Reward Curriculum : multiplication des pénalités par $k_c \in [0,1]$ progressivement croissant $r_{total} = r_{vel} - k_c r_{pen}$
- Noise curriculum : multiplication du bruit injecté par $k_{c,noise} \in [0,1]$ croissance plus lente que k_c
- Terrain curriculum : introduction de terrains irréguliers à la fin de l'entraînement. Certains paramètres de PPO sont réduits en parallèle pour effectuer des mises à jours plus modérées de la politique.

Résultats expérimentaux

Suivi de la vitesse



Profils de vitesse (trot aléatoire)



Valeurs articulaires cibles vs. mesurées (patte HR)

Consommation énergétique

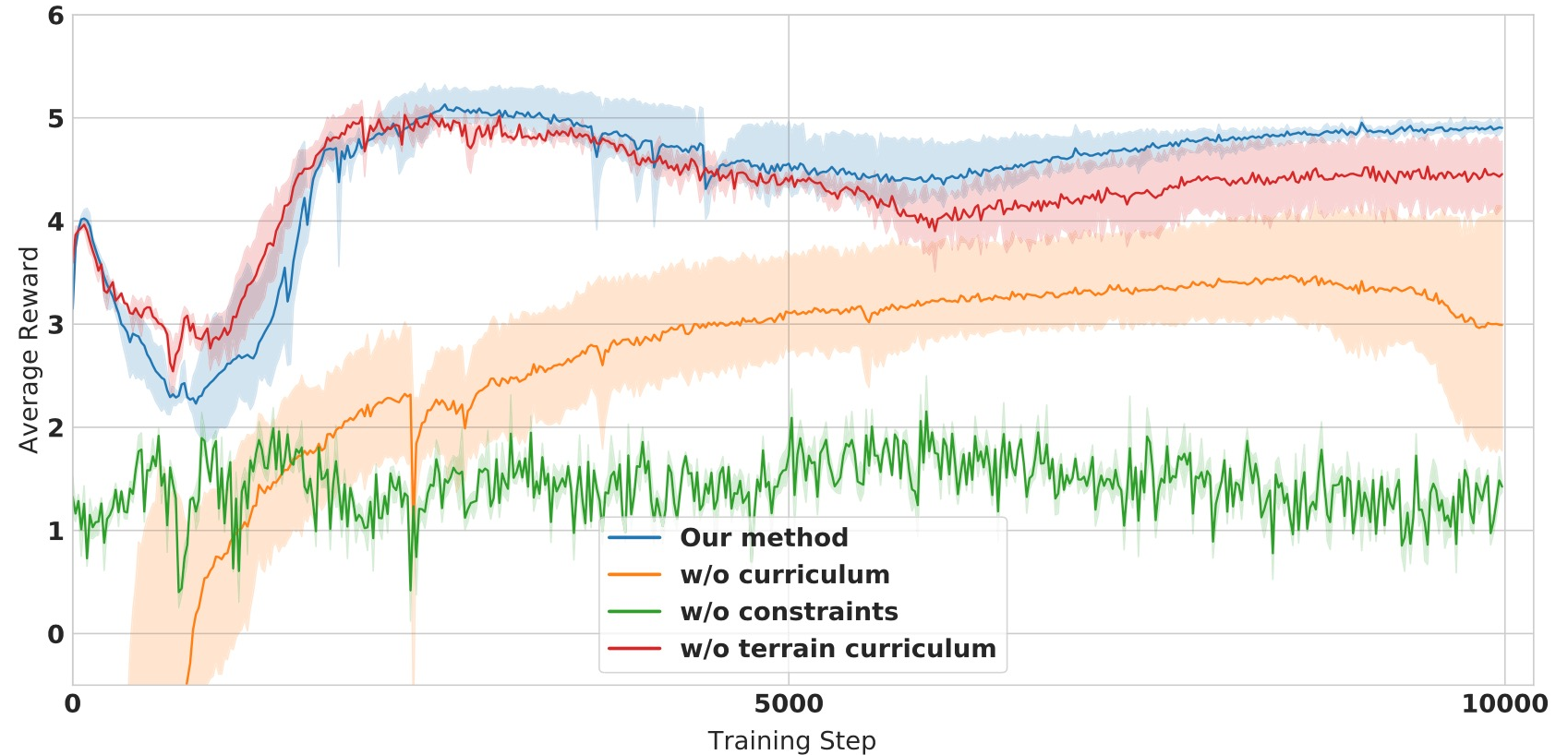
Performances au regard de la variation de la pondération c_E de la pénalité sur la puissance dissipée (vitesse 1.5m/s)

c_E	Power [W]	Velocity error [m/s] (%)	Base height [m]
0.0	17.7	0.079 ± 0.054 (7.9%)	0.23 ± 0.004
0.1	16.2	0.083 ± 0.067 (8.3%)	0.23 ± 0.006
1.0	13.7	0.092 ± 0.065 (9.2%)	0.24 ± 0.009
2.0	12.0	0.121 ± 0.064 (12.1%)	0.24 ± 0.007
3.0	11.0	0.141 ± 0.086 (14.1%)	0.24 ± 0.014
4.0	10.2	0.145 ± 0.091 (14.5%)	0.25 ± 0.004
10.0	7.7	0.198 ± 0.164 (19.8%)	0.25 ± 0.014
20.0	5.5	0.275 ± 0.113 (27.5%)	0.23 ± 0.005

Résultats expérimentaux

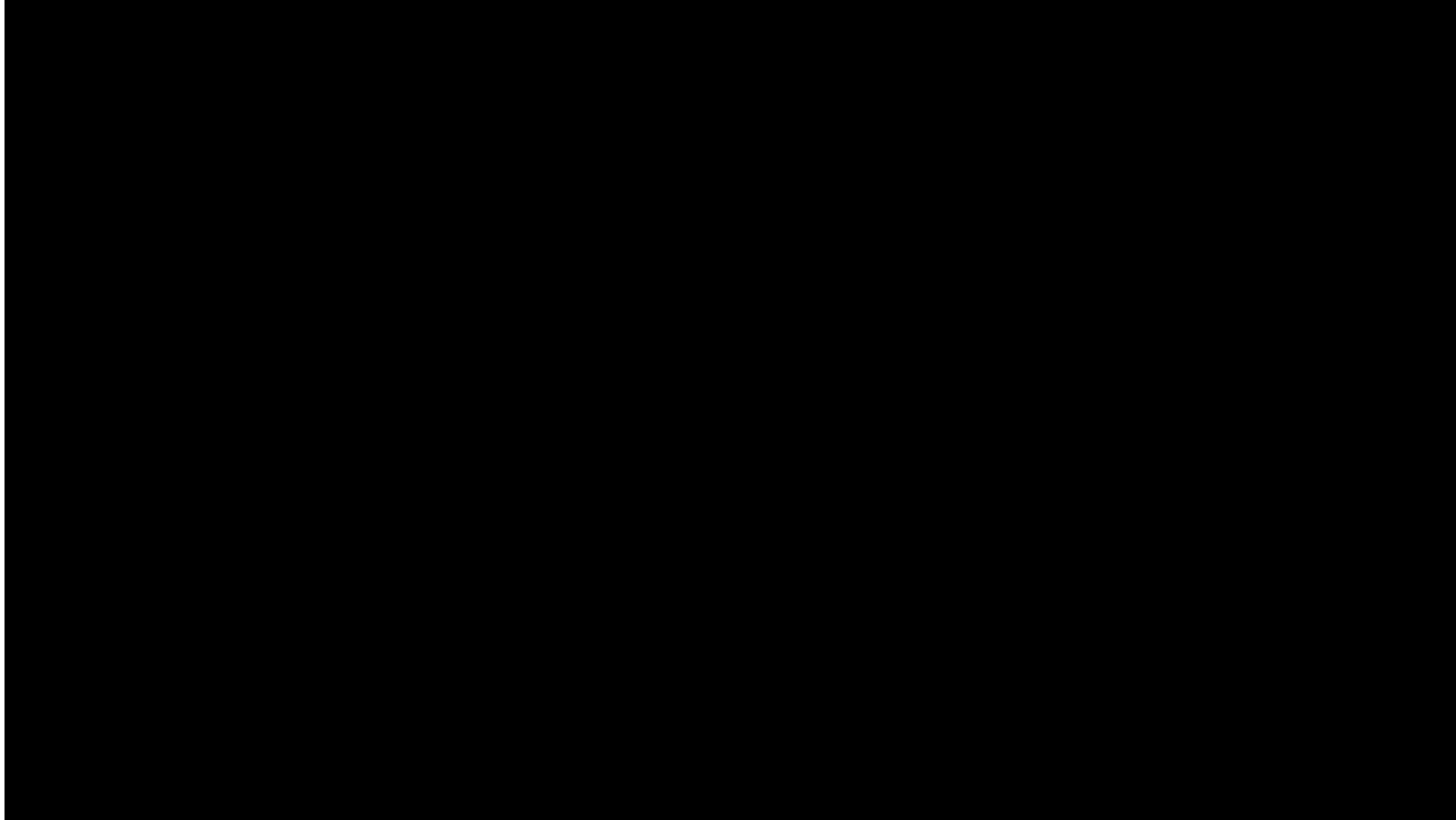
Effet du Curriculum

Apprentissage moyen au fil de l'entraînement pour différentes approches



→ L'apprentissage fait naturellement émerger un lien proportionnel entre la fréquence de l'allure et la vitesse de consigne.

Résultats expérimentaux



M. Aractingi, P.A. Léziart, T. Flayols, J. Perez, T. Silander, P. Souères. Controlling the Solo12 Quadruped Robot with Deep Reinforcement Learning, Scientific Reports, 11945, July 2023.

La nouvelle génération...



La nouvelle génération “Sassa”



*Oréotrague/
Klipspringer*

Entièrement ouvert mais en cours de développement

- Poids \approx 13kg
- Moteurs : 12 RMD X8-H Myactuator avec réducteurs planétaires
- Contrôleurs sur Raspberry Pi 4
- Batteries 40V Li-Ion
- Electronique développée au LAAS

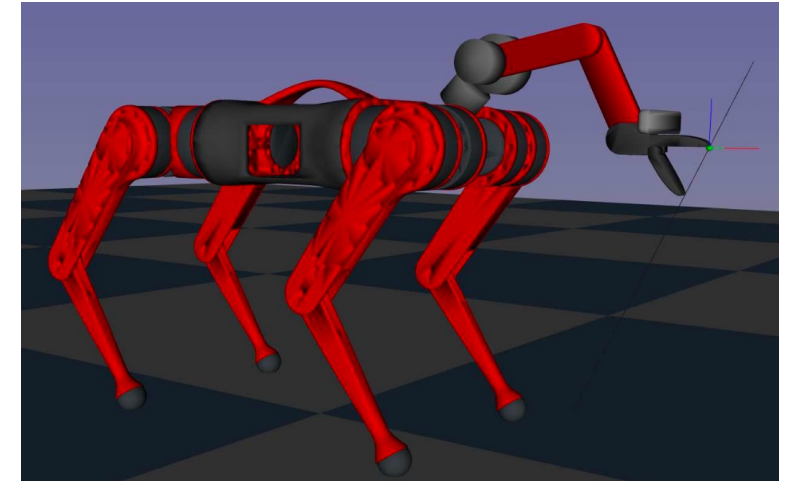


Quadruped challenge ICRA 2023

La nouvelle génération “Sassa”

Project QUADRUPES

- Co-design
- Locomotion basée MPC et RL (approches hiérarchiques et hybrides)
 - Vers un apprentissage sur le robot réel
- Tête mobile équipée de capteurs extéroceptifs (vision, IMU, audio) + préhenseur de type “bouche”
- Capacités de navigation et de préhension
- Inspiration biologique
- Education ouverte



Merci !



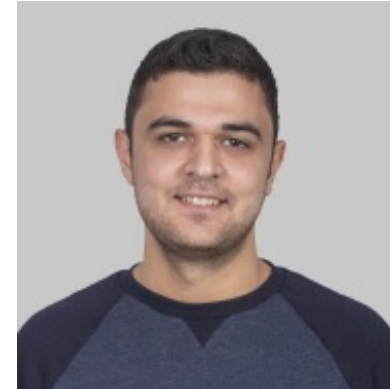
Thomas Flayols



Pierre-Alexandre Léziart



Gabriele Fadini



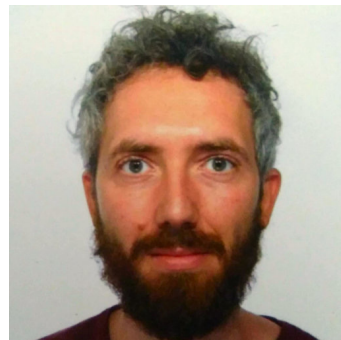
Michel Aractingi



Thomas Corbères



Tomi Silander



Andrea Del Prete



Steve Tonneau



Nicolas Mansard