Iranian Society of Acoustics and Vibration
The 15th ISAV2025 International Conference on Acoustics and Vibration
24-25 Dec 2025 Tehran- Iran

An Integral Reinforcement Learning Approach for Adaptive Chatter Control in Milling Processes

Pourya Shadkami Ahvazi[a*], Hossein Mohammadi[a]

[a] *School of mechanical engineering, Shiraz university, Shiraz, Iran.*

** Corresponding author e-mail: p40032529@hafez.shirazu.ac.ir*

Abstract

Chatter vibration in milling is a self-excited phenomenon that deteriorates surface finish, shortens tool life, and limits productivity. Conventional model-based controllers rely heavily on accurate identification of system dynamics and cutting force coefficients, which requires extensive modal testing and chip formation analysis. To overcome these limitations, this study proposes a partially model-free adaptive control approach based on Integral Reinforcement Learning (IRL) for chatter suppression in milling processes. Unlike traditional controllers, the proposed IRL-based design depends solely on the equivalent mass matrix, eliminating the need for stiffness, damping, or cutting force parameters. The control framework adopts an actor–critic learning structure in which the critic network estimates the value function, while the actor iteratively updates the control policy to minimize the Bellman error. To guarantee closed-loop stability in the presence of state delay, a Lyapunov–Razumikhin function is employed, ensuring asymptotic convergence of the estimation and control errors. The resulting control law drives both the control force and the state variables to operate within their optimal ranges, minimizing the quadratic performance index without prior model knowledge. Simulation results show that the proposed adaptive controller effectively suppresses regenerative chatter and substantially enlarges the chatter-free region in the Stability Lobe Diagram, allowing stable cutting at higher spindle speeds and greater axial depths.

Keywords: Active Chatter Suppression, Partially Model-free Control, Actor-Critic, Milling.

1. Introduction

Chatter vibration in milling is a self-excited phenomenon caused by the regenerative effect between the tool and the workpiece. It deteriorates surface quality, accelerates tool wear, and may damage the spindle system. As spindle speeds increase and structures become lighter, chatter suppression becomes essential for maintaining machining accuracy and productivity [1].

Various control approaches have been developed to mitigate chatter. Passive control methods aim to modify the structural properties of the system by increasing stiffness or damping, or by introducing variable spindle speed modulation [2]. Although these methods are simple and cost-effective,

they provide limited adaptability under changing cutting conditions. In contrast, active control strategies utilize actuators such as piezoelectric stack actuators [3], magnetic bearings [4], or active fixtures to apply counteracting forces [5]. Classical feedback controllers, including PID [6], LQR, and model predictive control [4], have been implemented to stabilize tool vibrations; however, these approaches typically depend on accurate identification of the cutting force coefficients, equivalent stiffness, and damping parameters. Obtaining these parameters requires extensive modal testing and chip formation analysis, making real-time implementation challenging [7].

To improve controller performance in systems with inherent time delays, several studies have investigated memory-based feedback control frameworks. For instance, an optimal memory-based LQR controller using a Lyapunov functional and Linear Matrix Inequalities (LMIs) was recently proposed for chatter suppression in milling, demonstrating significant improvement in the stability region compared to memoryless controllers [8]. Despite these advances, such methods still rely on a full knowledge of system matrices and cutting parameters.

Recent developments in machine learning have motivated the application of data-driven adaptive controllers to machining stability enhancement. Reinforcement Learning (RL), in particular, provides a powerful framework for learning optimal control policies through direct interaction with the system. Building on this concept, IRL offers a continuous-time formulation suitable for systems with partially unknown dynamics [9,10]. However, time-delay effects—intrinsic to regenerative chatter—require additional consideration to ensure closed-loop stability.

To address these challenges, this study proposes a partially model-free adaptive control approach based on Integral Reinforcement Learning (IRL) combined with the Lyapunov–Razumikhin stability theorem [11]. The proposed method depends solely on the equivalent mass matrix while eliminating the need for stiffness, damping, and cutting coefficients. The controller adaptively updates its policy through actor–critic learning to minimize the Bellman error, guaranteeing asymptotic stability under time delay. Simulation results confirm that the proposed IRL–Razumikhin controller effectively suppresses chatter and expands the chatter-free region in the Stability Lobe Diagram.

2. Time Delay Dynamic Modeling of Milling Chatter

The dynamic behavior of the milling process can be effectively represented by a two-degree-of-freedom model in the feed x and normal y directions, where regenerative chatter arises from the periodic interaction between the tool and the workpiece, as illustrated in Figure 1.

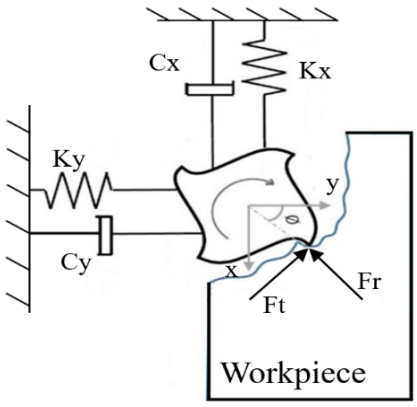

Figure 1. Two-dimensional schematic of the milling process and regenerative effect.

The cutting tool, shown schematically in Figure 2, is assumed to have four teeth N = 4 equally spaced around the spindle with a pitch angle $\phi_p = \frac{2\pi}{N}$. The time delay $\tau = \frac{60}{N\Omega}$ defines the interval between successive tooth engagements and depends on the spindle speed Ω. The instantaneous imersion angle determines when each cutting edge enters and exits the material, specifying the engagement region in which the dynamic cutting forces act and the regenerative effect develops.

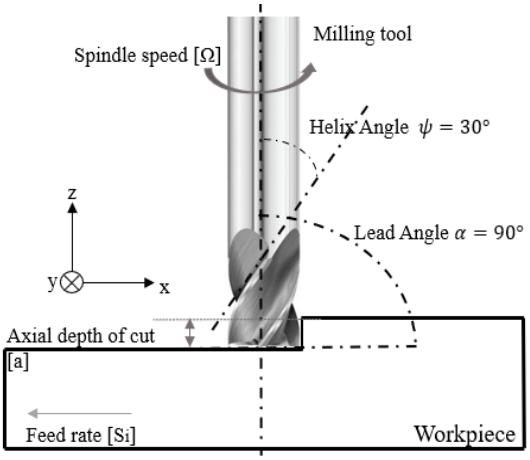

Figure 2. Cutter geometry showing helix angle, lead angle, and axial depth of cut [a]

The instantaneous chip thickness consists of a static component (h_s), arising from the nominal feed per tooth, and a dynamic component (h_d) caused by the regenerative effect, as expressed in Eq. (1). Since the static chip thickness does not influence the chatter vibration mechanism, it is neglected in the simulation [12]. The dynamic chip thickness, defined in Eq. (2) by the relative displacement between current and delayed tool positions, serves as the primary source of chatter excitation, while the engagement of each tooth is governed by the step function $w(\phi_i)$ given in Eq. (3) [13].

$$h(\phi_i) = h_d(\phi_i) + h_s(\phi_i) \quad (1)$$

$$h_d(\phi_i) = \big([x(t) - x(t-\tau)]sin(\phi_i) + [y(t) - y(t-\tau)]cos(\phi_i)\big)w(\phi_i) \quad (2)$$

$$w(\phi_i) = \begin{cases} 1 & \phi_s \le \phi_i \le \phi_e \\ 0 & \phi_s > \phi_i \ \ or \ \ \phi_e < \phi_i \end{cases} \quad (3)$$

ϕ_i is the immersion angle of cutting. Also ϕ_s and ϕ_e are the start and end angles of cutting. The cutting forces acting on the tool are modeled in the tangential (F_t) and radial (F_r) directions as proportional to the instantaneous dynamic chip thickness, as shown in Eq. (4).

$$F_t(\phi_i) = K_t a h(\phi_i) \quad (4\text{-a})$$

$$F_r(\phi_i) = K_r a h(\phi_i) \quad (4\text{-b})$$

a is the axial depth of cut. K_t , K_r represent the tangential and radial cutting force coefficients, respectively. Through trigonometric transformation, these force components are projected onto the Cartesian coordinates to obtain the corresponding forces in the feed x and normal y directions, as presented in Eq. (5).

$$F_{x_i} = -F_{t_i}\cos(\phi_i) - F_{r_i}\sin(\phi_i) \quad (5\text{-a})$$

$$F_{y_i} = F_{t_i}\sin(\phi_i) - F_{r_i}cos(\phi_i) \quad (5\text{-b})$$

After averaging the first term of the Fourier series expansion, the equivalent cutting coefficient λ is obtained [14].

$$F = \big[F_x, F_y\big]^T = a\begin{bmatrix}\lambda_{xx} & \lambda_{xy} \\ \lambda_{yx} & \lambda_{yy}\end{bmatrix}[x(t) - x(t-\tau), y(t) - y(t-\tau)]^T \quad (6)$$

By defining the state variables as $x_1 = x,\ x_2 = y,\ x_3 = \dot{x},\ x_4 = \dot{y}$, the milling system in Eq. (7) can be represented as a time-delay state-space model that captures both the regenerative effect and the external control input (u).

$$\dot{x}(t) = A_0 x(t) + A_1 x(t-\tau) + B\big[u_x, u_y\big]^T \quad (7)$$

Where,

$$A_0 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -\frac{(k_x + a\lambda_{xx})}{m_x} & -\frac{a\lambda_{xy}}{m_x} & -\frac{c_x}{m_x} & 0 \\ -\frac{a\lambda_{yx}}{m_y} & -\frac{(k_y + a\lambda_{yy})}{m_y} & 0 & -\frac{c_y}{m_y} \end{bmatrix},\ A_1 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \frac{a\lambda_{xy}}{m_x} & \frac{a\lambda_{xy}}{m_x} & 0 & 0 \\ \frac{a\lambda_{yx}}{m_y} & \frac{a\lambda_{yy}}{m_y} & 0 & 0 \end{bmatrix},\ B = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ \frac{1}{m_x} & 0 \\ 0 & \frac{1}{m_y} \end{bmatrix}$$

3. Adaptive Memoryless Controller Based on Integral Reinforcement Learning

At the initial stage, consider a time-delay linear time-invariant (LTI) system with a known initial function, where the matrices A_0 and A_1 are unknown, while only the input matrix B is available. To design the optimal control policy u*(t), the following quadratic cost function is defined:

$$J = \int_0^\infty \left(x^T(t)\tilde{Q}x(t) + u^T(t)\tilde{R}u(t) \right) dt \quad (8)$$

where $\tilde{Q} \in \mathbb{R}^{4\times4}$ and $\tilde{R} \in \mathbb{R}^{2\times2}$ are symmetric positive-definite weighting matrices. In reinforcement learning problems, which operate based on the agent–environment interaction, a reward function r(x,u) is required. In this framework, the quadratic cost function J is treated as the reward to be minimized, ensuring that both the state and control inputs remain within their optimal operating ranges.

The Lyapunov–Razumikhin function is then introduced as the value function candidate:

$$\mathcal{V}\big(x(t)\big) = x^T(t)\mathcal{P}x(t) \quad (9)$$

where $\mathcal{P} \in \mathbb{R}^{4\times4}$ is a symmetric positive-definite matrix. Taking the derivative of $\mathcal{V}\big(x(t)\big)$ yields:

$$\dot{\mathcal{V}}\big(x(t)\big) = 2x^T(t)\,\mathcal{P}\left(A_0x(t) + A_1x(t-\tau) + Bu(t)\right) \quad (10)$$

Substituting this expression into the Hamiltonian equation results in the optimal control policy:

$$u^*(t) = -R^{-1}B^T\mathcal{P}x(t) \quad (11)$$

Since the system dynamics A_0, A_1 and the matrix $\mathcal{P}$ are unknown, the controller parameters must be determined through the Integral Reinforcement Learning process rather than explicit model identification[9]. Considering the cost function J as the reward function and the Lyapunov–Razumikhin function as value function over a long time horizon T, the Eq.(12) relation holds:

$$\mathcal{V}\big(x(t)\big) - \mathcal{V}(x(T)) = \int_t^T \left(x^T(t)\tilde{Q}x(t) + u^T(t)\tilde{R}u(t) \right) dt \quad (12)$$

By vectorizing the quadratic term $\mathcal{P}$ using the Kronecker product $\bar{x}(t) = x(t)\otimes x(t)$, the unknown vector $vec(\mathcal{P})$ can be estimated through:

$$vec(\mathcal{P})(\bar{x}(t) - \bar{x}(T)) = \int_t^{t+T} \left(x^T(t)\tilde{Q}x(t) + u^T(t)\tilde{R}u(t) \right) dt \quad (13)$$

However, due to the presence of time delay, this conventional estimation is no longer valid. To overcome this, a Bellman error formulation is adopted following the approach of Moghadam et al. [9], allowing approximate estimation of $vec(\mathcal{P})$ for delayed systems through iterative learning:

$$E_{BE} = \int_t^T \left(x^T(t)\tilde{Q}x(t) + u^T(t)\tilde{R}u(t) \right) dt + vec(\mathcal{P})(\bar{x}(t) - \bar{x}(T)) \quad (14)$$

Here, $\Delta\bar{x}(t) = \bar{x}(t) - \bar{x}(T)$, and the ideal critic parameter is defined as $vec(\mathcal{P})$, which is unknown due to the lack of explicit system information. Therefore, the parameter must be approximated using the critic function defined as Eq.(15).

$$\hat{\mathcal{V}}\big(x(t)\big) = x^T(t)\hat{\mathcal{P}}x(t) \quad (15)$$

By substituting the estimated critic parameter $\hat{P}_v = vec(\mathcal{P})$ into the Bellman error Eq.(14), it follows that if a feasible control policy ($u_0(t)$) is applied, the Bellman error converges to zero, and the estimated matrix $\hat{\mathcal{P}}$ approaches the optimal value $\mathcal{P}$:

$$E_{BE} = \int_t^T \left(x^T(t)\tilde{Q}x(t) + u^T(t)\tilde{R}u(t) \right) dt + \hat{P}_v\Delta\bar{x}(t) \quad (16)$$

A feasible control policy is an initial gain that stabilizes the closed-loop system, though it is not necessarily optimal; such a stabilizing policy can be obtained by assigning negative poles. To derive the update law for the estimated parameter $\hat{P}_v$, we first define the (squared) Bellman error norm in Eq. (17); partial derivative of E with respect to $\hat{P}_v$ and applying a gradient-descent scheme then yields the update rule given in Eq. (18).

$$E = \frac{1}{2} E_{BE}^2 \tag{17}$$

$$\dot{\hat{P}}_v = -\beta \Delta\bar{x}(t) E_{BE} \tag{18}$$

where β denotes the learning rate. To avoid divergence or slow convergence—resulting respectively from large or small values of $\Delta\bar{x}(t)$—a quadratic normalization term is incorporated. In addition, a small positive constant γ is introduced to prevent unnecessary parameter growth. The final normalized update law is expressed in Eq. (19). From this relation, the estimated parameter vector $\hat{P}_v$ is updated iteratively, and the corresponding estimated matrix $\hat{\mathcal{P}}$ is then employed to compute the optimal control policy given in Eq.(20).

$$\dot{\hat{P}}_v = -\beta \left(\frac{\Delta\bar{x}(t)}{\left(1+\Delta\bar{x}(t)^T \Delta\bar{x}(t)\right)^2} \right) E_{BE} - \gamma\hat{\theta} \tag{19}$$

$$u^*(t) = -R^{-1} B^T \hat{\mathcal{P}} x(t) \tag{20}$$

This iterative process continues until the Bellman error satisfies $E_{BE} < \varepsilon$. The overall structure of the adaptive memoryless controller based on IRL is illustrated in Figure 7.

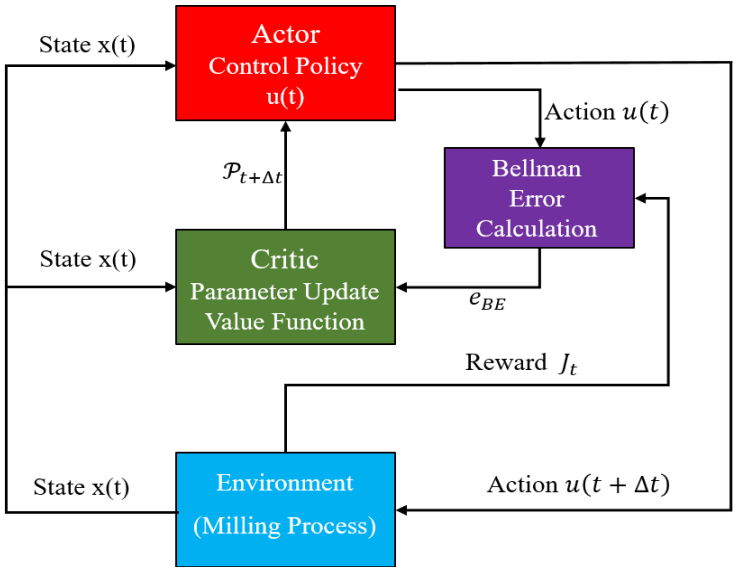

Figure 3. Block diagram of the IRL–based controller using the actor–critic algorithm.

4. Results and Discussion

To validate the effectiveness of the proposed adaptive control approach, a series of numerical simulations were performed under realistic cutting conditions. The system parameters used in the simulations are listed in Table 1, which provides the mechanical and cutting characteristics of the milling setup. These parameters were identified from experimental measurements on a CNC milling machine to ensure consistency between the simulated and actual dynamic responses of the process. The simulation environment thus offers a reliable platform for assessing the performance of the IRL-based controller in suppressing regenerative chatter.

Table 1. Dynamic parameters and cutting coefficients used in the milling process model [14].

Parameter	Value	Parameter	Value
m_x	$13.4\ gr$	K_y	$7.36 \times 10^5 \frac{N}{m}$
C_x	$19.6\ \frac{N.s}{m}$	λ_{xx}	$117.22 \frac{N}{mm^2}$
K_x	$7.14 \times 10^5 \frac{N}{m}$	λ_{xy}	$93.5 \frac{N}{mm^2}$
m_y	$30.2\ gr$	λ_{yx}	$-30 \frac{N}{mm^2}$
C_y	$30\ \frac{N.s}{m}$	λ_{yy}	$-76.78 \frac{N}{mm^2}$

4.1 Closed-Loop Analysis of Chatter Suppression

In this simulation, the axial depth of cut was set to 2 [mm], and the spindle speed was maintained at 3500 [rpm]. The weighting matrices in the quadratic cost function were empirically selected as $Q = I_{2\times2}$ and $R = 5 \times I_{2\times2}$. Based on these parameters and the proposed Integral Reinforcement Learning algorithm, the optimal feedback gain matrix was iteratively learned and applied to the milling process. The final control gain obtained after convergence is given by:

$$K = \begin{bmatrix} 17697.54 & -646223.58 & -16.46 & 17.70 \\ 591.56 & -286734.96 & 15.85 & -28.90 \end{bmatrix}$$

Figure 4 illustrates the time-domain response of the milling system under the proposed IRL-based adaptive controller. As observed, the controller effectively suppresses the chatter-induced oscillations within a short time interval, restoring the system to a stable cutting condition.

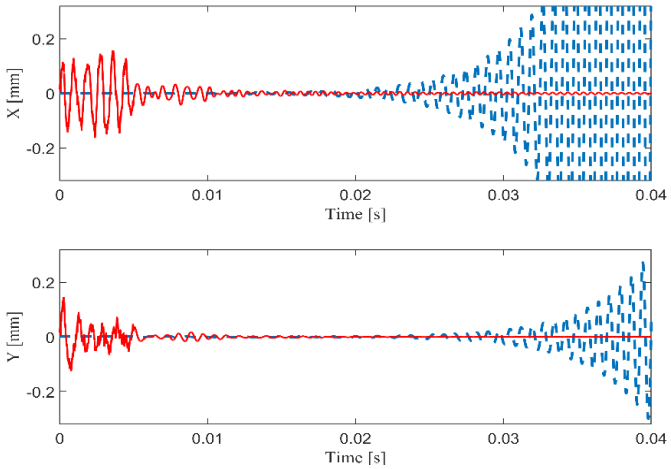

Figure 4. Time-domain response of tool vibration. The blue dashed line represents the uncontrolled case, while the red solid line shows the response under the adaptive IRL controller

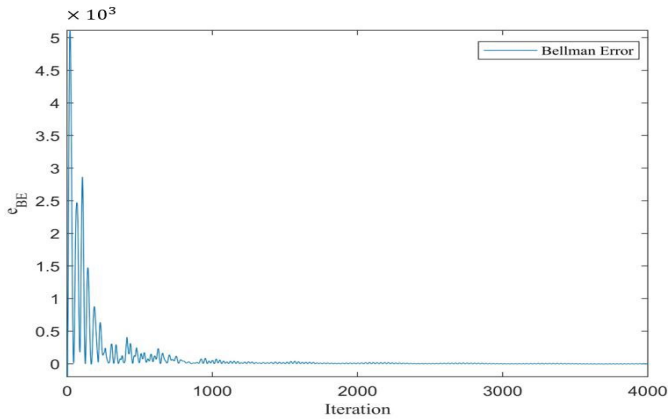

Figure 5. Bellman error versus iteration index

The convergence trend of the Bellman error is shown in Figure 5, where the estimated parameter reaches near-zero after about 1000 iterations, indicating that the learning process has stabilized and the control parameters are accurately identified.

As shown in Figure 6, the proposed controller generates a maximum control force of approximately 40 N, which is considerably lower than that of the memory-based optimal controller [8]. This reduction highlights the efficiency of the IRL approach, as it decreases actuator demand and the risk of saturation while maintaining effective chatter suppression.

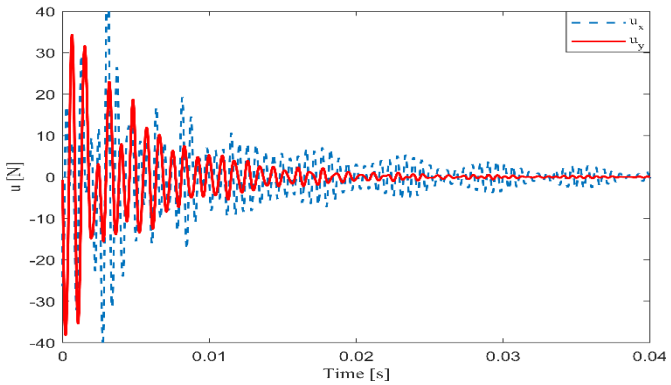

Figure 6. Control force signal for chatter suppression during the milling process.

4.2 Stability Lobe Diagram Analysis

Finally, the Stability Lobe Diagram for the proposed controller is depicted in Figure 7. The objective of the controller is to enlarge this stable region and the results reveal a substantial enlargement of the stable cutting region when the controller is implemented. Specifically, at the selected spindle speed, the allowable axial depth of cut increases by approximately 1 [mm] compared to the uncontrolled case, indicating enhanced tolerance to more aggressive cutting conditions without entering the chatter-unstable region.

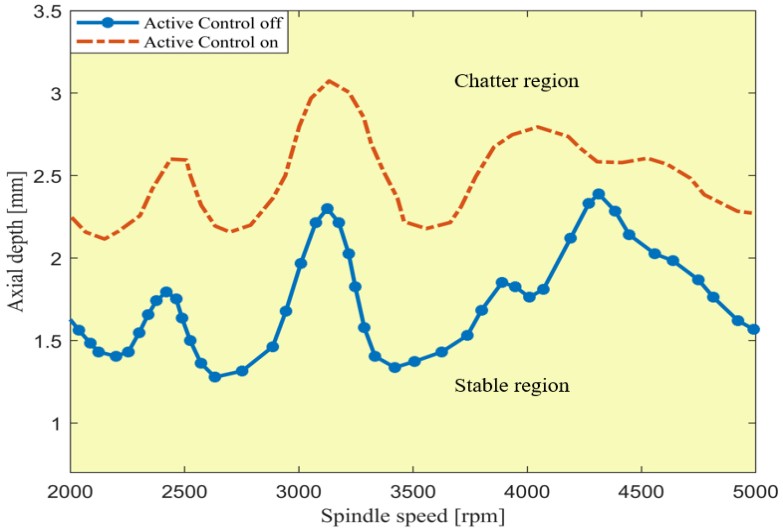

Figure 7. Stability Lobe Diagram of the milling process: red dashed line represents the system with the proposed controller, and blue line corresponds to the uncontrolled case.

5. Conclusion

This study presented an adaptive memoryless controller based on the Integral Reinforcement Learning (IRL) framework for chatter suppression in milling operations. Unlike conventional model-based controllers, the IRL-based approach eliminates the need for identifying system damping, stiffness, or cutting force coefficients and operates without requiring chip formation experiments or

cutting-force analysis. Through continuous interaction with the milling environment, the controller autonomously identifies the necessary parameters and effectively suppresses regenerative chatter. The analysis of the Stability Lobe Diagram indicates that the proposed controller expands the chatter-free stability region by approximately 1 [mm] in axial depth of cut, allowing milling to be performed at greater cutting depths with improved process stability and productivity.

REFERENCES

1. Caixu, Y. U. E., Haining, G. A. O., Xianli, L. I. U., Liang, S. Y., & Lihui, W. A. N. G. (2019). A review of chatter vibration research in milling. *Chinese Journal of Aeronautics*, *32*(2), 215-242.
2. Seguy, S., Dessein, G., Arnaud, L., & Insperger, T. (2009, September). Chatter control by spindle speed variation in high-speed milling. In *The International Conference on Structural Analysis of Advanced Materials*.
3. Wan, S., Li, X., Su, W., Yuan, J., & Hong, J. (2020). Active chatter suppression for milling process with sliding mode control and electromagnetic actuator. *Mechanical Systems and Signal Processing*, *136*, 106528.
4. Wan, S., Li, X., Su, W., Yuan, J., & Hong, J. (2020). Active chatter suppression for milling process with sliding mode control and electromagnetic actuator. *Mechanical Systems and Signal Processing*, *136*, 106528.
5. Guo, M., Xia, W., Liu, J., Guo, W., & Wu, C. (2023). Investigation on active vibration control to improve surface quality in precision milling process. *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*, 09544054231207422.
6. Paul, S., & Morales-Menendez, R. (2018). Active control of chatter in milling process using intelligent PD/PID control. *IEEE Access*, *6*, 72698-72713.
7. Moradi H, Bakhtiari-Nejad F, Movahhedy MR, Vossoughi G. Stability improvement and regenerative chatter suppression in nonlinear milling process via tunable vibration absorber. Journal of sound and vibration. 2012 Oct 8;331(21):4668-90.
8. Ahvazia PS, Mohammadia H, Mohammadia M. Optimal Memory-Based Feedback Control for Time-Delayed Systems with Application to Suppressing Regenerative Chatter. ISA transactions. 2025 Jul 21.
9. Moghadam R, Jagannathan S. Approximate optimal adaptive control of partially unknown linear continuous-time systems with state delay. In2019 IEEE 58th Conference on Decision and Control (CDC) 2019 Dec 11 (pp. 1985-1990). IEEE.
10. Wang G, Luo B, Xue S. Integral reinforcement learning-based optimal output feedback control for linear continuous-time systems with input delay. Neurocomputing. 2021 Oct 14;460:31-8.
11. Razumikhin BS. On the stability of systems with a delay. Prikl. Mat. Mekh. 1956 Jan;20(4):500-12.
12. Insperger T, Stépán G. Updated semi-discretization method for periodic delay-differential equations with discrete delay. International journal for numerical methods in engineering. 2004 Sep 7;61(1):117-41.
13. Altintas, Y., Engin, S., & Budak, E. (1998, November). Analytical stability prediction and design of variable pitch cutters. In *ASME International Mechanical Engineering Congress and Exposition* (Vol. 16066, pp. 141-148).
14. Moradi H, Vossoughi G, Movahhedy MR. Experimental dynamic modelling of peripheral milling with process damping, structural and cutting force nonlinearities. Journal of Sound and Vibration. 2013 Sep 16;332(19):4709-31.