# Stick slip vibrations in drilling: modeling, estimation, avoidance

Jean Auriol, Nasser Kazemi Nojadeh

*Proceedings of the 6th International Colloquium on Nonlinear Dynamics and Control of Deep Drilling Systems,*
*T.G. Ritto, D.A. Castello, B. Balachandran, E. Detournay and N. Van de Wouw (Editors).*
*Universidade Federal do Rio de Janeiro (UFRJ), Rio de Janeiro, RJ, Brazil, July 1st to 3rd, 2024*

# Stick slip vibrations in drilling: modeling, estimation, avoidance

**Jean Auriol**[1] **and Nasser Kazemi**[2]

[1] Université Paris-Saclay, CNRS, CentraleSupélec, L2S, Gif-sur-Yvette, France.

[2] Université du Qubéc à Montréal, Department of Earth and Atmospheric Sciences, Montreal, Canada.

*Abstract: This writeup presents a field-validated torsional model for the drill string dynamics and different algorithms for estimating the friction factors along the drill string and bottom hole rotational velocity. These friction terms characterize the interaction between the drill pipe and the wellbore walls (Coulomb source terms) within the curving wellbore. This information is essential to designing the next generation of stick-slip mitigation controllers, developing real-time wellbore monitoring tools, and enabling effective toolface control for directional drilling.*

***Keywords: drill-string vibrations, stick-slip, observer design, transformers***

## INTRODUCTION

Extraction of resources at depths greater than a few hundred meters in the earth's subsurface (oil, gas, minerals, and thermal energy) necessitates the drilling of long slender boreholes (from the surface to the subsurface target) that may induce deviated well paths with extensive horizontal sections (Sveinbjornsson and Thorhallsson, 2014; Wei *et al.*, 2022). Advanced drilling control strategies are essential for designing such complex wells, as undesirable vibrations can occur during transient off-bottom phases. Notably, torsional oscillations, which cause damage and reduce the rate of penetration, must be mitigated. This phenomenon, known as *stick-slip*, results from significant friction between the drill string and the borehole (Aarsnes *et al.*, 2018). To prevent such oscillations, advanced control approaches have been developed, incorporating the underlying distributed dynamics into controller design (Auriol *et al.*, 2022a). These control strategies have demonstrated faster convergence to a reference trajectory compared to state-of-the-art controllers while preventing stick-slip oscillations. They are based on the field-validated model presented in (Aarsnes and Shor, 2018) that accurately describes the evolution of angular velocity and torque $(\omega(t,x), \tau(t,x))$ along the drilling device using coupled hyperbolic Partial Differential Equations (PDEs). However, implementing such control laws requires knowledge of the distributed state along the drill string. In the field, available measurements are mainly surface data, such as surface RPM or motor torque. Additionally, these methods require all system parameters, particularly subsurface physical properties, to be known (Auriol *et al.*, 2019, 2021).

In this abstract, we present the high-fidelity model introduced by Aarsnes and Shor (2018) to describe torsional oscillations in the drilling device. We then show how this model can be used to design appropriate near real-time state observers. In particular, we present an adaptive observer (adjusted from (Aarsnes *et al.*, 2019)) and an innovative *dual architecture of physics-informed transformer-based neural networks* (Vaswani *et al.*, 2017). Indeed, recent advances have shown that deep neural networks can be used to learn distributed dynamics from measurements (Lu *et al.*, 2021). Finally, we briefly show how these estimations can be used to control the downhole behavior of the drill string while avoiding undesired oscillations.

## TORSIONAL DYNAMICS OF THE DRILL STRING

This section introduces a hyperbolic PDEs model incorporating distributed friction terms to describe the torsional motion of drill string dynamics. The proposed high-fidelity model, detailed in (Aarsnes and Shor, 2018), has been validated against field data. Its computational simplicity makes it suitable for control and estimation applications. We assume that the torsional motion of the drill string is the dominating dynamic behavior (i.e., there is no distributed axial dynamics). Moreover, the effects of along-string cuttings distribution on the friction is assumed to be homogeneous and the transition from static to dynamic Coulomb friction is a jump, i.e., the Stribeck curve is assumed negligible. Finally, the bit is also assumed to be off-bottom, meaning there is no bit-rock interaction, which occurs during transient phases when the bit is not in contact with the rock, such as when adding a new pipe section to the drilling system or when removing the drill string from the borehole to address a failure. For $(t,x) \in [0,T] \times [0,L]$ (where $T > 0$ is the chosen time window and $L > 0$ is the length of the drilling device), we denote $\omega(t,x)$ the angular velocity and $\tau(t,x)$ the torque at any point of the drill string. The states satisfy the following equations

$$\frac{\partial \tau(t,x)}{\partial t} + JG\frac{\partial \omega(t,x)}{\partial x} = 0, \qquad J\rho\frac{\partial \omega(t,x)}{\partial t} + \frac{\partial \tau(t,x)}{\partial x} = \mathscr{S}(t,x), \tag{1}$$

where $J$ is the polar moment for inertia, $G$ is the shear modulus, and $\rho$ is the mass density, averaged for a drill string section and supposedly known. The source terms $\mathscr{S}$ is due to frictional contact with the borehole. It verifies

$$\mathscr{S}(t,x) = -k_t \rho J \omega(t,x) - \mathscr{F}(t,x), \tag{2}$$

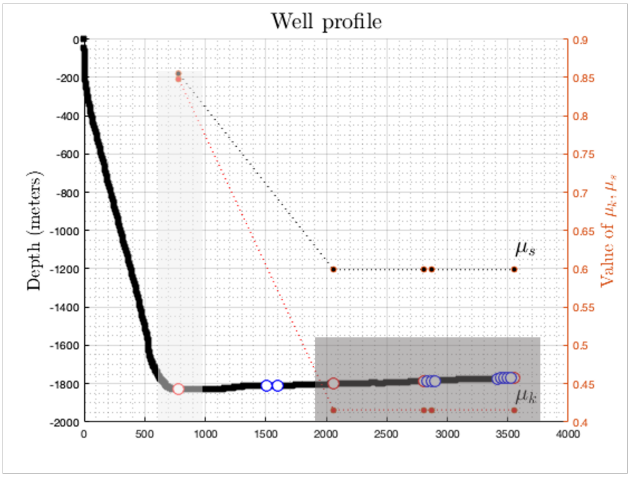


Figure 1: Schematic indicating the distributed drill string lying in deviated bore-hole.

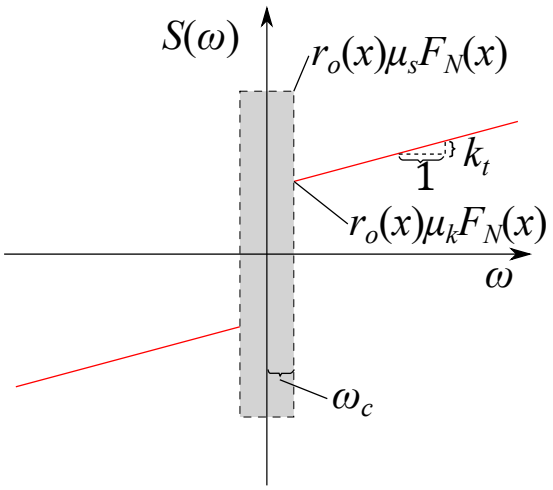


Figure 2: Friction source term $S(\omega,x)$. The red curve indicates the dynamic torque as a function of angular velocity.

where the damping constant $k_t$ is the viscous shear stress and $\mathscr{F}$ is a differential inclusion that corresponds to the Coulomb friction between the drill string and the borehole, also known as the side force:

$$\begin{cases} \mathscr{F}(t,x) = \text{sign}(\omega(t,x))F_k(x), & |\omega(t,x)| > \omega_c, \\ \mathscr{F}(t,x) \in \pm r_o(x)\mu_s F_N(x), & |\omega(t,x)| \leqslant \omega_c, \end{cases} \tag{3}$$

where the function $\mathscr{F}(\omega) = -\frac{\partial \tau(t,x)}{\partial x} \in \pm r_o(x)\mu_s F_N(x)$ is the differential inclusion. The functions $F_k(x) \doteq r_o(x)\mu_k F_N(x)$ (resp. $F_s(x) \doteq r_o(x)\mu_s F_N(x)$) corresponds to the dynamic (resp. static) Coulomb torque. The expression of the normal force acting between the drill string and the borehole wall $F_N(x)$ depends on the well geometry (Sheppard *et al.*, 1987). While the outer drill string radius $r_o(x)$ is known, the static and kinetic friction coefficients $(\mu_s, \mu_k)$ and the angular velocity threshold $\omega_c$ are not. At the surface level, the top drive is actuated by an electrical motor that imparts torque to the drill string. The evolution of the angular velocity at the surface level and the downhole torque follows

$$\frac{d}{dt}\omega(t,0) = \frac{1}{I_{TD}}(\tau_m(t) - \tau(t,0)), \quad \tau(t,L) = 0 \tag{4}$$

where $I_{TD}$ corresponds to the top-drive inertia and $\tau_m$ to the motor torque. We denote $\omega_0(t) = \omega(t,0)$ the top-drive angular velocity. The motor torque corresponds to the control input and may be expressed as a PI feedback (Åström and Murray, 2010): i.e. $\tau_m = k_p(\omega_{SP} - \omega_{t,0}) + k\int_0^t(\omega_{SP}(\xi) - \omega_{t,0}(\xi))d\xi$, where $\omega_{SP}$ is a desired reference set-point, $k_p$ is a proportional gain and $k$ an integral gain. It can also be defined through more advanced control strategies (as shown in (Aarsnes *et al.*, 2018; Auriol *et al.*, 2022a)). However, such advanced mitigation laws require knowledge of the entire distributed angular state. In the field, we typically have access only to the top-drive angular velocity and the motor torque $(\omega(t,0), \tau_m(t))$ measured with a frequency of $1Hz$. Therefore, it is necessary to reconstruct the entire state from these measurements, as proposed in (Aarsnes *et al.*, 2019). Reconstructing the state may also require knowledge of unknown physical parameters, such as static and kinetic friction coefficients. In the following, we denote $Y = (P, \{(\omega(t,0), \tau_m(t))_{i\in[1,N]}\}) \in \mathscr{Y} = \mathbb{R}^{N_p} \times \mathbb{R}^{2\times N}$ an input data containing a set of $N_p$ known physical parameters (depth of bit $L$, collar and pipes mechanical properties) concatenated in the vector $P$ and $N$ sequential pairs of surface measurements.

## STATE ESTIMATION

The design of the next generation of stick-slip mitigation controllers requires reliable estimations of the state and friction parameters. More precisely, our objective is twofold: Estimate the physical parameters $(\mu_s, \mu_k)$ and the distributed state $(\omega(t,x), \tau(t,x))$ from the surface data. In this section, we present different methods (adjusted from Auriol *et al.* (2022b)) to estimate all (or part) of the friction parameters.

### First approach: Adaptive observer

The first approach we propose is adapted from the adaptive state observer introduced in (Aarsnes *et al.*, 2019). This observer integrates measurements from physical sensors, particularly the top-drive angular velocity $\omega_0$, with the proposed system dynamics model. It provides reliable estimates of the torque and RPM states as well as the side forces friction parameters when the bit is off-bottom (Aarsnes *et al.*, 2019). Here, we summarize the main ideas of this observer. Denote

with the $\hat{}$ superscript the estimated states and $e(t) = \hat{\omega}_0(t) - \omega_0(t)$ the measured estimation error of the top-drive angular velocity. The observer equations given in (Aarsnes *et al.*, 2019) correspond to a copy of the original dynamics expressed in the Riemann coordinates, to which are added output correction terms. The Riemann invariants are the states corresponding to a transformation of the system which has a diagonalized transport matrix. We define them as $\alpha = \omega + \frac{1}{J\sqrt{G\rho}}\tau$ and $\beta = \omega - \frac{1}{J\sqrt{G\rho}}\tau$. The observer equations read as follows

$$\dot{\hat{\omega}}_0 = a_0\left(\hat{\beta}_p(t,0) - \hat{\omega}_0\right) + \frac{1}{I_{TD}}\tau_m - p_0 e(t), \tag{5}$$

$$\frac{\partial \hat{\alpha}}{\partial t}(t,x) + c_t\frac{\partial \hat{\alpha}}{\partial x}(t,x) = \hat{\mathscr{S}}(t,x) - p_\alpha(t,x)e(t), \qquad \frac{\partial \hat{\beta}}{\partial t}(t,x) - c_t\frac{\partial \hat{\beta}}{\partial x}(t,x) = \hat{\mathscr{S}}(t,x) - p_\beta(t,x)e(t), \tag{6}$$

where $c_t = \sqrt{\frac{G}{\rho}}$. The source term in each section is computed from the estimated states and friction factors $\hat{\mathscr{S}}(t,x) = k_t(\hat{\alpha}(t,x) + \hat{\beta}(t,x)) + \frac{1}{J\rho}\hat{\mathscr{F}}(t,x)$ where $\hat{\mathscr{F}}$ has an expression identical to (3), the different variables being replaced by their estimates. The different boundary conditions are now expressed as $\hat{\alpha}(t,0) = 2\hat{\omega}_0(t) - \hat{\beta}(t,0) - P_0 e(t)$, $\hat{\beta}(t,L) = \hat{\alpha}(t,L)$, and the estimates of the friction factor are updated according to

$$\dot{\hat{\mu}}_s(t) = \begin{cases} -l_s e(t), & |\min_x \hat{\omega}(t,x)| \leqslant \omega_c, \\ 0, & |\min_x \hat{\omega}(t,x)| > \omega_c, \end{cases} \quad \dot{\hat{\mu}}_k(t) = \begin{cases} 0, & |\min_x \hat{\omega}(t,x)| \leqslant \omega_c, \\ l_k e(t), & |\min_x \hat{\omega}(t,x)| > \omega_c, \end{cases} \tag{7}$$

Finally, we use a saturation to improve the robustness of the approach: $\hat{\mu}_s = \max(\hat{\mu}_s, \hat{\mu}_k)$. The different constants and observer gains $a_0, p_\alpha, p_\beta, p_0, p_1, P_0, P_1, l_s, l_k$ can be found in (Aarsnes *et al.*, 2019). As shown in (Aarsnes *et al.*, 2019), the convergence is guaranteed in the absence of friction term. Moreover, the proposed procedure is robust to uncertainties and delays and provides correct estimations of $\mu_k$ and $\mu_s$ when tested in simulations and against field data. However, it requires the knowledge of $\omega_c$, and there is no proof of convergence for the adaptive part.

## Second approach: transformer-based dual architecture network

We now present a learning methodology based on a *dual architecture*, originally introduced in Redaud *et al.* (2024). A first transformer-based neural network (using $Y$ as input) estimates the physical parameters $(\mu_s, \mu_k)$. We opted for transformers (Vaswani *et al.*, 2017) instead of recurrent neural networks to enhance speed and reduce the computational cost of training. The estimated physical parameters $(\mu_s, \mu_k)$ and the input $Y$ are then fed into a second transformer-based neural network, which aims to provide the distributed state $(\omega(t,x), \tau(t,x))$. Following (Raissi *et al.*, 2019; Sun *et al.*, 2021), we incorporate physical laws into the training loss functions to improve the performance of the estimation algorithms. The output of the first network is required to compute loss functions. Both networks are trained simultaneously using an extensive simulated training dataset. The proposed network is schematically illustrated in Figure 5.

### *Estimation of physical parameters*

The first neural network aims at estimating the physical parameters $M = (\mu_k, \mu_s) \in \mathbb{R}^2$. It is a transformer (Vaswani *et al.*, 2017) characterized by trainable parameters $\theta$, and denoted $T_\theta(.)$. It aggregates the sequence of inputs $Y \in \mathbb{R}^{N_p+2N_t}$ and gives as an output $\hat{M} \in \mathbb{R}^2$. We obtain $\theta$ by minimizing the $L_2-$ loss function $\mathscr{L}_{L_2}(\theta) = \frac{1}{N_b}\sum_{i=1}^{N_b} \|M - T_\theta(Y)\|_2^2$.

### *Estimation of the distributed state*

Inspired by (Lu *et al.*, 2021), we tackle the second objective using a two-branch architecture. Denote $\mathscr{X} = [0, L]$ as the spatial domain and $\mathscr{T} = [0, T]$ as the temporal domain where the state is defined. We aim to select the most appropriate state representation concerning physical constraints using sets of discrete inputs $Y$. Therefore, we approximate

$$S: \begin{array}{ccc} \mathscr{Y} & \longrightarrow & \mathscr{C}^\infty(\mathscr{T}\times\mathscr{X}, \mathbb{R}^2) \\ Y & \mapsto & S_{Y,\Theta}(\cdot,\cdot) \end{array},$$

where $\mathscr{C}^\infty$ is the set of infinitely differentiable functions from $\mathscr{T}\times\mathscr{X}$ to $\mathbb{R}^2$ (the regularity could be leveraged). This defines a class of parametric functions $S_{Y,\Theta}$, which we use for approximating the real states $(\omega(t,x), \tau(t,x))$. The corresponding approximations is denoted as $(\hat{\omega}(t,x), \hat{\tau}(t,x))$. In our design, the first branch of the proposed architecture relies on a transformer encoder to aggregate the input sequences $Y$ augmented with $\hat{M}$. This produces an abstract representation of the system when combined with an abstract representation of the requested coordinates. Inspired by Fourier Neural Operators (Li *et al.*, 2021), this abstract representation is then used to output the intensity, frequency, and phase of the Fourier decomposition representing the distributed state $X(t,x) \doteq (\omega(t,x), \tau(t,x))$ along the drill string. The second branch constructs the spatiotemporal grid mesh $(t,x) \in \mathscr{T}\times\mathscr{X}$, where the estimation is evaluated.

### *Physic-informed neural networks*

Following (Sun *et al.*, 2021; Wang *et al.*, 2021), we define the composite loss $\mathscr{L}(\Theta) = \mathscr{L}_{\mathscr{D}}(\Theta) + \mathscr{L}_{\text{PDE}}(\Theta) + \mathscr{L}_{\text{BC}}(\Theta)$. The term $\mathscr{L}_{\mathscr{D}}$ is the usual error term when training a neural network on a dataset. It corresponds to the

state estimation residual in the squared $L_2-$norm. The term $\mathscr{L}_{\text{PDE}}$ ensures that (1) is satisfied. Finally, the last term $\mathscr{L}_{\text{BC}}$ guarantees that the solution meets the boundary condition. More precisely, we have

$$\mathscr{L}_{\mathscr{D}}(\Theta) = \frac{1}{N_x}\frac{1}{N_t}\frac{1}{N_b}\sum_{i=1}^{N_b}\sum_{j=1}^{N_x}\sum_{k=1}^{N_t}\|X(t_k,x_j) - S_{Y,\Theta}(t_k,x_j)\|_2^2,$$

$$\mathscr{L}_{\text{PDE}}(\Theta) = \frac{1}{N_x}\frac{1}{N_t}\frac{1}{N_b}\sum_{i=1}^{N_b}\sum_{j=1}^{N_x}\sum_{k=1}^{N_t}\|\mathscr{O}(f_{Y,\Theta}(t_k,x_j))\|_2^2, \quad \mathscr{L}_{\text{BC}}(\Theta) = \frac{1}{N_t}\frac{1}{N_b}\sum_{i=1}^{N_b}\sum_{k=1}^{N_t}\|\mathscr{B}(S_{Y,\Theta}(t_k))\|_2^2 + |\hat{\tau}(t_k,L)|^2,$$

with $\mathscr{B}(S_{Y,\Theta}(t)) = \frac{\partial\hat{\omega}}{\partial t}(t,0) - \frac{1}{I_{TD}}(\tau_m(t) - \hat{\tau}(t,0))$ and $\mathscr{O}(S_{Y,\Theta}(t,x)) = \begin{pmatrix} \frac{\partial\hat{\tau}}{\partial t}(t,x) + JG\frac{\partial\hat{\omega}}{\partial x}(t,x) \\ \frac{\partial\hat{\tau}}{\partial x}(t,x) + J\rho\frac{\partial\hat{\omega}}{\partial t}(t,x) - \mathscr{S}(\hat{\omega},x) \end{pmatrix}$. This machine-learning approach is reliable and easy to implement. It consists of a fast algorithm (once properly trained) that provides good estimations of the friction parameters and the state. The data and physics-driven losses add *a priori* knowledge of the underlying dynamics (1)-(4) during training. However, the proposed approach requires thousands of training points and may lack generalizability as it depends on the model and the initial condition of the system.

## CONTROLLER DESIGN

During the drilling process, the operator usually wants to control the downhole behavior of the drill string and optimize the Rate Of Penetration (ROP) while avoiding undesired oscillations. More precisely, the main control objective is to regulate the downhole angular velocity at the start-up of a drilling operation (e.g., after a connection) to avoid entering a stick-slip limit cycle. Most of the controllers applied in industrial applications correspond to high-gain PI control laws, such as the SoftSpeed and SoftTorque approaches. These approaches are easy to implement and analyze since the gains are tuned to obtain a certain reduction in the proximal reflection coefficient over a limited frequency range. However, they may present several fundamental limitations, such as possible poor inherent robustness margins (which can be overcome using impedance matching controllers such as ZTorque (Dwars, 2015)) or the generation of significant oscillations when changing the set-points. Advanced control methods have been proposed in (Auriol *et al.*, 2022a), leading to better performance but requiring real-time estimation. Below, we summarize the main ideas behind the approaches proposed in (Auriol *et al.*, 2022a) and detail their corresponding requirements, limitations, and performance. In these control strategies, the effect of the friction terms is seen as a distributed source term $d(t)$ that needs to be compensated.

1. **ZTorque and feed-forward controller.** A feedforward controller that can be seamlessly integrated with standard industry ZTorque impedance matching feedback controllers without affecting the closed-loop behavior has been proposed in (Aarsnes *et al.*, 2018). This control law adheres to the 3DOF controller architecture (Åström and Murray, 2010), as it has three components: a feedback term, a feedforward term (leveraging the model's differential flatness), and a disturbance cancellation term. This approach is a state-of-the-art controller and requires the knowledge of top-drive torque and estimation of the disturbance term. The feedforward component and disturbance compensation term reduce residual oscillations but may induce overshoots and long convergence times.

2. **Multiplicity Induced Dominancy Controller.** This approach involves analyzing the transfer function between the downhole velocity and the actuator. This transfer function represents a time-delay system for which we can design a control law that ensures the placement of the dominant root in the complex plane, thereby guaranteeing the stability of the closed-loop system. This approach requires the estimation of downhole velocity and disturbance terms. It is robust with respect to parameter uncertainties and exhibits slight overshoot and oscillations with a low computational effort. However, it may be sensitive to noise.

3. **Recursive dynamics interconnection framework.** This procedure is well-suited for multi-sectional drilling devices. For each section, we find the virtual input that stabilizes the section and guarantees that the output converges towards the virtual input of the next subsystem. The virtual input of the last subsystem (downhole ODE) is chosen to guarantee the tracking of the reference signal. It requires a reliable estimation of the disturbance term and predictions of the states. As it consists of a direct compensation of the disturbance term, it can guarantee a fast convergence without residual oscillations at the cost of an instantaneous high-control effort. All in all, it is a complex control algorithm with an extensive computational cost.

## EXPERIMENTAL SETTING

In this section, we illustrate the performance of the proposed transformer-based observer and of the ZTorque and feedforward controller with simulations.

### Generation of dataset and Training parameters

To train and validate our estimation algorithm, we generate a wide dataset following the numerical scheme presented in (Aarsnes and Shor, 2018). To obtain representative data, we use the real well geometry (J1), illustrated in Figure 1.

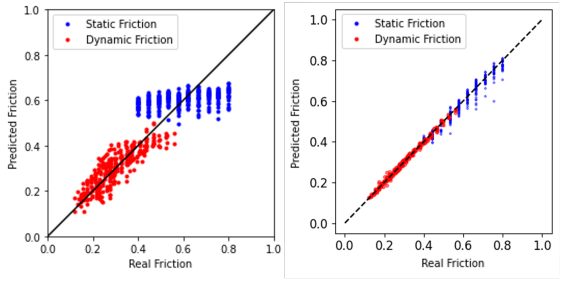


Figure 3: Estimation of $(\mu_k, \mu_s)$ using (Auriol *et al.*, 2022b) (left) and the transformer approach (right).

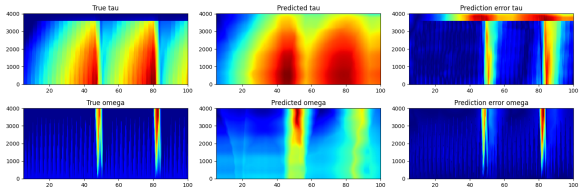


Figure 4: Example of a prediction for a sample of the validation set after training.

We generated 1000 sequences of 100s of 20Hz surface measurements (motor torque and surface angular velocity), for $L \in [2500, 4000]$m, $\mu_s \in [0.2, 0.8]$, $\mu_k \in [0.06, 0.72]$. The reference trajectory is constant (60 RPM with slope). The implemented control input is ZTorque with a feed-forward component. The dataset is split $80\% - 20\%$ as training and validation datasets. To obtain the parameters $(\theta, \Theta)$ minimizing the losses $\mathscr{L}_{L_2}(\theta), \mathscr{L}(\Theta)$, we use AdamW (Loshchilov and Hutter, 2017) with an initial learning rate of $10^{-3}$. The training is done on 100 epochs, with a batch size $N_b = 16$. We use automatic differentiation techniques (Baydin *et al.*, 2015) to compute the derivatives.

## Simulation results

Both networks are trained simultaneously on a comprehensive dataset of simulated data based on the proposed well geometry. The trained neural networks' performance is then evaluated using a separate validation dataset, as detailed in (Auriol *et al.*, 2022b). We compare these results with the estimations obtained from the convolutional neural network-based approach presented in (Auriol *et al.*, 2022b). The friction coefficients were estimated on the validation dataset with an average relative error of $\delta(\mu_k) = 2.3\%$ (resp. $\delta(\mu_s) = 3.3\%$) and a standard deviation of $5.2e^{-3}$(resp. $1.8e^{-2}$) after 2500 steps. As illustrated in Figure 3, our proposed method outperformed the existing estimation methods presented in (Auriol *et al.*, 2022b), as the standard deviation is reduced, and the average estimated value is closer to the true value.

We also obtained promising results for the state estimation with an average absolute error of 2.3 on 200 validation examples. We have plotted in Figure 4 the obtained estimation for $\tau$ and $\omega$ and compared them with their real values. We used color plots to picture these 2D data (red corresponding to higher values, the horizontal axis being time, and the vertical axis being the curvilinear abscissa). As expected, the solutions predicted by the proposed networks are consistent with the physics. Finally, we illustrate in Figure 6 how this observer can be combined with the ZTorque feedback controller with a flatness trajectory planning feed-forward component to stabilize the system around different set-points. In the considered scenario, the lowest part of the drilling device (collar and BHA) is almost horizontal, and the stationary drill string is initially kept in place by the Coulomb friction until enough torque is built up to overcome it. The different physical parameters can be found in Auriol *et al.* (2022a).

## CONCLUSIONS

We presented a field-validated torsional model for drill string dynamics. We then introduced two algorithms to estimate the friction factors along the drill string, providing an estimate of the bottom hole rotational velocity. These observers were then used to design the next generation of stick-slip mitigation controllers. Future improvements include reducing the numerical complexity of the different control strategies and a better online adaptive estimation of the disturbance term. The current architecture for the machine learning observer includes hyperparameters that can be optimized to enhance performance, e.g., by balancing the empirical loss of model predictions, model complexity, and physical loss.

## REFERENCES

Aarsnes, U.J.F., Auriol, J., Di Meglio, F. and Shor, R., 2019. "Estimating friction factors while drilling". *Journal of Petroleum Science and Engineering*, Vol. 179, pp. 80–91.

Aarsnes, U.J.F., Di Meglio, F. and Shor, R.J., 2018. "Avoiding stick slip vibrations in drilling through startup trajectory design". *Journal of Process Control*, Vol. 70, pp. 24–35.

Aarsnes, U.J.F. and Shor, R.J., 2018. "Torsional vibrations with bit off bottom: Modeling, characterization and field data validation". *Journal of Petroleum Science and Engineering*, Vol. 163, pp. 712–721.

Åström, K.J. and Murray, R.M., 2010. *Feedback systems: an introduction for scientists and engineers*. Princeton university press, 2nd edition. ISBN 9780691135762.

Auriol, J., Boussaada, I., Shor, R.J., Mounier, H. and Niculescu, S.I., 2022a. "Comparing advanced control strategies to eliminate stick-slip oscillations in drillstrings". *IEEE access*, Vol. 10, pp. 10949–10969. ISSN 2169-3536.

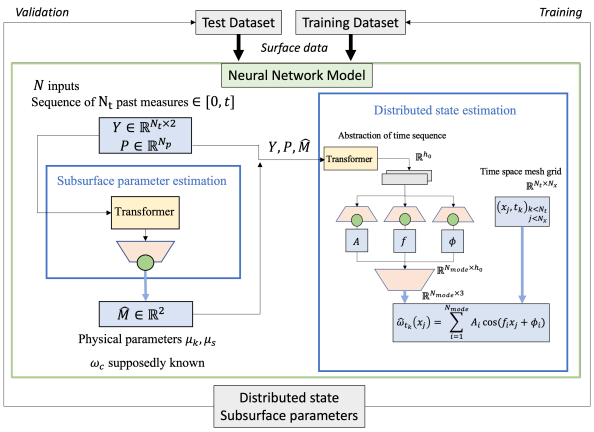


Figure 5: Detailed view of the neural network architecture.

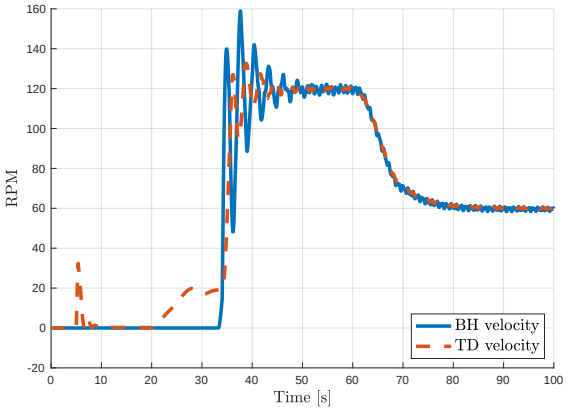


Figure 6: Response to velocity set-point changes using a ZTorque feedback controller with a flatness trajectory planning feed-forward component (depth=1800m).

Auriol, J., Shor, R., Niculescu, S.I. and Kazemi, N., 2022b. "Estimating drill string friction with model-based and data-driven methods". In *2022 American Control Conference (ACC)*. pp. 3464–3469.

Auriol, J., Kazemi, N. and Niculescu, S.I., 2021. "Sensing and computational frameworks for improving drill-string dynamics estimation". *Mechanical Systems and Signal Processing*, Vol. 160, p. 107836.

Auriol, J., Kazemi, N., Shor, R.J., Innanen, K.A. and Gates, I.D., 2019. "A sensing and computational framework for estimating the seismic velocities of rocks interacting with the drill bit". *IEEE Transactions on geoscience and remote sensing*, Vol. 58, No. 5, pp. 3178–3189.

Baydin, A., Pearlmutter, B., Radul, A. and Siskind, J.M., 2015. "Automatic differentiation in machine learning: a survey".

Dwars, S., 2015. "Recent Advances in Soft Torque Rotary Systems". In *Proceedings of 2015 SPE/IADC Drilling Conference*. March, pp. 17–19.

Li, Z., Kovachki, N., Azizzadenesheli, K., Liu, B., Bhattacharya, K. and Stuart, A., 2021. "Fourier neural operators for parametric partial differential equations". p. 16.

Loshchilov, I. and Hutter, F., 2017. "Decoupled weight decay regularization". doi:10.48550/ARXIV.1711.05101.

Lu, L., Jin, P. and Karniadakis, G.E., 2021. "DeepONet: Learning nonlinear operators for identifying differential equations based on the universal approximation theorem of operators". *Nature Machine Intelligence*, Vol. 3, No. 3, pp. 218–229.

Raissi, M., Perdikaris, P. and Karniadakis, G.E., 2019. "Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations". *Journal of Computational physics*, Vol. 378, pp. 686–707.

Redaud, J., Darrin, M., Kazemi, N. and Auriol, J., 2024. "Physics-informed dual architecture neural networks for enhanced estimation of drilling dynamics". In *IGARSS, IEEE International Geoscience and Remote Sensing Symposium*. Vol. to appear.

Sheppard, M.C., Wick, C. and Burgess, T., 1987. "Designing Well Paths To Reduce Drag and Torque". *SPE Drilling Engineering*, Vol. 2, pp. 344–350.

Sun, J., Innanen, K.A. and Huang, C., 2021. "Physics-guided deep learning for seismic inversion with hybrid training and uncertainty analysis". *GEOPHYSICS*, Vol. 86, No. 3, pp. R303–R317.

Sveinbjornsson, B.M. and Thorhallsson, S., 2014. "Drilling performance, injectivity and productivity of geothermal wells". *Geothermics*, Vol. 50, pp. 76–84.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L. and Polosukhin, I., 2017. "Attention is all you need". doi:10.48550/ARXIV.1706.03762.

Wang, S., Wang, H. and Perdikaris, P., 2021. "Learning the solution operator of parametric partial differential equations with physics-informed DeepONets". *Science Advances*, Vol. 7, No. 40, p. 8605.

Wei, C., Mao, L., Yao, C. and Yu, G., 2022. "Heat transfer investigation between wellbore and formation in U-shaped geothermal wells with long horizontal section". *Renewable energy*, Vol. 195, pp. 972–989. ISSN 0960-1481.