Linear-quadratic flotation level control through reinforcement learning

Norlund, Frida; Tammia, Rasmus; Hägglund, Tore; Soltesz, Kristian

2023

Document Version:
Peer reviewed version (aka post-print)

Link to publication

Citation for published version (APA):
Norlund, F., Tammia, R., Hägglund, T., & Soltesz, K. (2023). *Linear-quadratic flotation level control through reinforcement learning*. Paper presented at 22nd IFAC World Congress, Yokohama, Japan.

Total number of authors:
4

Linear-Quadratic Flotation Level Control through Reinforcement Learning

F. Norlund [*] **R. Tammia** [*] **T. Hägglund** [**] **K. Soltesz** [**]

[*] *Boliden AB, Boliden, Sweden (email: {frida.norlund,rasmus.tammia}@boliden.com)*
[**] *Lund University, Dept. Automatic Control, Lund, Sweden (email: {kristian.soltesz,tore.hagglund}@control.lth.se)*

Abstract: In the mining industry, flotation is a commonly used process to separate valuable minerals from waste rock in a concentrator. The rougher flotation is the first stage of the process and in Boliden AB's concentrator at Aitik, it consists of two lines of four flotation cells. In this paper we consider one line of four cells and the buffer tank upstream of them. The operating conditions in the flotation process change as the ore quality varies. This is a challenge when modeling the process. We address this challenge by using a reinforcement learning (RL) algorithm to design a state feedback controller for level control, without the need of an explicit process model. Using simulations, we compare the performance of the resulting controller to that of the cascade coupled PI-control structure that operates the real plant today. The RL-based controller improves the performance and shows good potential. Convergence to an admissible control law requires careful hyper-parameter tuning. Industrial deployment thus requires further work to ensure the required reliability.

Keywords: Machine learning methods and applications, Advanced process control, Process optimisation

1. INTRODUCTION

In order to produce metals, the raw ore needs to be processed to concentrate the minerals that will later be smelted into metals. In the process of concentrating minerals, flotation is commonly used. In series of flotation cells, the differences in surface properties are used to separate the valuable minerals from waste rock. To do so, the milled ore is mixed with water to form a slurry to which chemical reagents are added. The reagents make the selected minerals water repellent, which allows them to attach to air bubbles generated at the bottom of the flotation cell and form a mineral froth on top of the slurry in the cell. To extract more of the minerals from the slurry, the tailing from one flotation cell is the feed to the next one. The froth is collected as it flows over the rim of the flotation cells and as stated in Bergh and Yianatos (2011), this makes good level control one of the foundations to having good overall recovery of the minerals.

Since the flotation cells are connected in series, the levels in the different cells form a strongly connected system. Therefore multivariable controllers and cascade coupling of SISO-loops are of interest for the level control. For example, LQ-control and a decoupling controller was investigated for level control by Stenlund and Medvedev (2002). Model predictive control (MPC) has also been implemented for flotation by Brooks and Koorts (2017), however, targeting the recovery of minerals instead of the level control.

All model-based control structures for flotation have one thing in common, the model is a key factor for how successful the controller will be. The quality of the incoming material will change over the ore deposit and this will change the optimal way to run the process. These variations are hard to capture in the model for measurement reasons. There may not be measurements of the quality in question or the frequency of the measurements may be too low. This becomes a challenge when the operating conditions drift far from the conditions under which the model of the process was created.

This paper investigates reinforcement learning for level control as an example of how a state feedback law can be designed without the need of an explicit process model. The considered process section will be the rougher flotation circuit at the concentrator plant in the Aitik mine, run by Boliden AB located in Gällivare, northern Sweden. The performance of the resulting controller will be compared to the existing control structure that controls the plant today. The comparisons will be performed in simulation.

2. PROCESS DESCRIPTION

The considered process consists of a buffer tank and four flotation cells. A schematic picture of the process is seen in Figure 1. Upstream from the considered process, there are two milling lines. Raw ore enters the milling lines, where it is mixed with water and ground to a fine sand, forming a slurry. This slurry enters the considered process section that starts with a buffer tank. Its volumetric inflow rate

[⋆] This work was partially supported by the Wallenberg AI, Autonomous Systems and Software Program (WASP) funded by the Knut and Alice Wallenberg Foundation and by the ELLIIT Strategic Research Area.

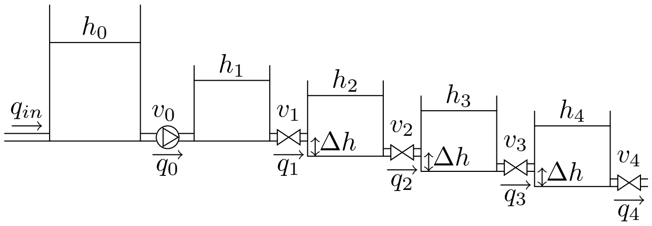

Fig. 1. Schematic overview of the considered process section: a buffer tank followed by four flotation cells in series. Slurry is actively pumped from the buffer tank to the first flotation cell. The slurry level in tank/cell i is h_i, and neighboring flotation cells are mounted at a height difference Δh. The flow q_i out of flotation cell i is moderated by the control signal v_i to a nonlinear valve.

is controlled by the milling line and hence considered as a disturbance. The slurry is actively pumped from the buffer tank to the first flotation cell. The flotation cells themselves are of equal size and mounted with Δh height difference between neighbors. Slurry flow between adjacent cells is driven by their difference in hydrostatic pressure, and moderated by a valve, as seen in Figure 1. Further details on the model is found in Norlund, F. (2022).

3. CONTROL

We investigate the use of a reinforcement learning (RL) algorithm to design a linear quadratic (LQ) controller. The algorithm is entirely data-driven and has no knowledge of the process model. This is of practical interest, since the process operating conditions change over time, and these changes are hard to accommodate for online within the scope of the model.

The underlying algorithm is based on Bradtke (1992) and Lewis et al. (2012). It relies on N samples, collected during closed-loop operation of the system, to update a state feedback controller based on a linear quadratic control cost function. This procedure is iterated until the control gain, $\mathbf{L}$ in the state feedback law

$$\boldsymbol{u}_k = -\mathbf{L}\boldsymbol{x}_k \tag{1}$$

converges to the gain that is optimal for the imposed cost function. Further details are found in Norlund, F. (2022). Notable is also that the RL-controller treats deviations from a linearization point instead of absolute states and control signals, such that

$$\begin{aligned} \boldsymbol{h} &= \boldsymbol{h}_{ref} + \boldsymbol{x}, \\ \boldsymbol{v} &= \boldsymbol{u}_{ref} + \boldsymbol{u}. \end{aligned} \tag{2}$$

4. SIMULATIONS

4.1 Designing a controller with the RL-algorithm

When using the RL-algorithm to design a controller, an initial feedback gain, $\mathbf{L}$, must be chosen to start the iterations. This gain must be stabilizing, but apart from that it has no other requirements. For our system, the identity matrix fulfills the requirements and it was chosen to be the initial feedback gain.

When running the algorithm, the control signal must perform a sufficient amount of exploration. Otherwise, the samples collected will not contain enough information about the dynamics of the system. To ensure sufficient excitation, a disturbance component $\boldsymbol{e}_k$ is added to each control signal sample, so that,

$$\boldsymbol{u}_k = -\mathbf{L}\boldsymbol{x}_k + \boldsymbol{e}_k. \tag{3}$$

For this system, the disturbance component, $\boldsymbol{e}_k$, consisted of Gaussian white noise and low frequency square waves. The period of the square wave was chosen to match the speed of the dynamics of the system components. The square waves are only present in one cell at a time, so that the effects of them are seen in the surrounding cells. The square waves are applied from left to right in the process, starting in the buffer tank and finishing in cell four. In Figure 2, the first tuning iteration is seen for the buffer tank and the second cell. The effects of the disturbances in the buffer tank are clearly visible in the second cell as well. The impacts of the square waves in the first cell is also clearly visible in the second cell, they take place between sample 400 and 800. The inflow of slurry to the buffer tank was assumed to be constant during the tuning process, and the level references for the cells were kept constant.

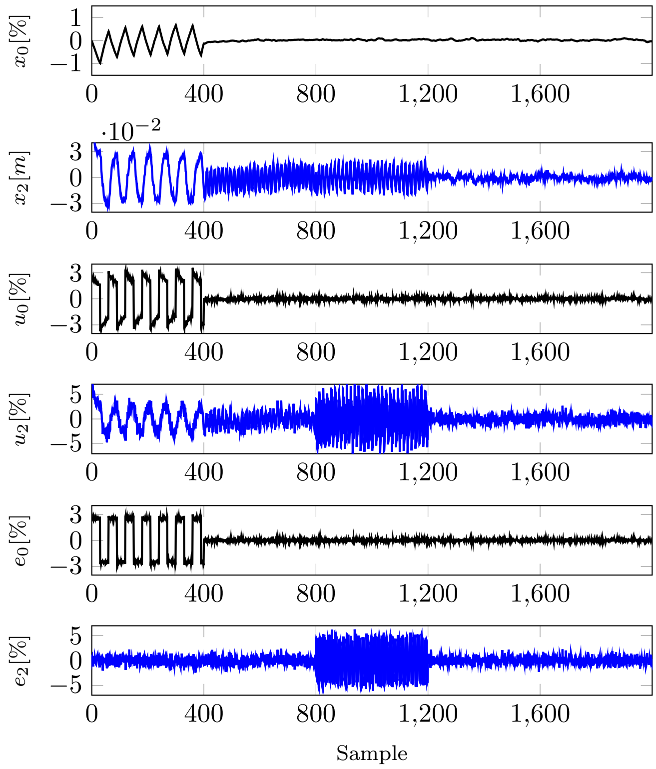

Fig. 2. The first tuning iteration of the RL-algorithm performed on the system consisting of a buffer tank and four flotation cells. There are square waves present in the disturbance term in equation (3) for one cell at a time, starting with the buffer tank. The buffer tank and the second cell are shown. The effects of the disturbances in the buffer tank on the flotation cell is visible, along with the effects of the first flotation cells disturbances on the second cell.

With five states and five control signals, the gain matrix will be of size 5×5. How the elements of its diagonal are updated over the first 15 iterations of the RL-algorithm is shown in Figure 3. With the control law in equation (1), the diagonal elements of $\mathbf{L}$ tells us how much the state, x_i,

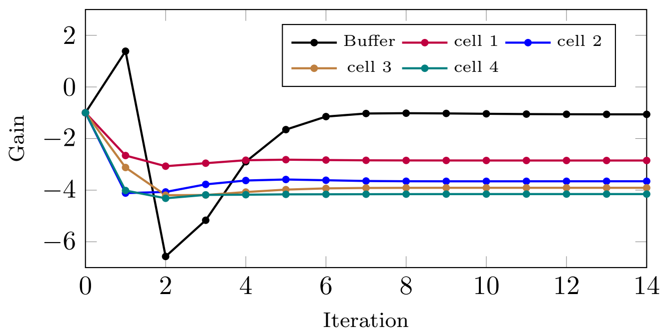

Fig. 3. The RL-algorithm is applied to the system consisting of a buffer tank and four flotation cells. The evolution—iteration by iteration—of the elements on the main diagonal of the gain matrix in the feedback law is visualized as they converge to stationary values.

contributes to the control signal u_i that controls x_i. Off-diagonal elements in the gain matrix also contribute to the control signals, but theses gains are, in our case, smaller. In the considered, and representative, example, it takes the algorithm seven iterations to converge to a feedback gain. After this, no notable improvements are achieved by the algorithm.

For the controller to be practically useful between tunings, it also needs reference tracking properties. The controller designed above will drive the states back to the references that were set when it was designed. To enable it to follow other references and correct stationary errors, without retuning it, integral action can be added to the controller. This can be achieved by extending the state vector with integral states of the tracking error. The feedback gain matrix must also be extended to include feedback gains for the integral states.

4.2 RL-based versus currently implemented controller

In the Aitik plant, variations in the slurry flow to the buffer tank is one of the biggest disturbances to the levels in the flotation cells. The biggest inflow disturbance occurs when one of the milling lines that supply flotation with slurry unexpectedly stops. This roughly halves the inflow of slurry to the flotation series. Process data from a real occurrence of this disturbance has been extracted and used as input to the model. In Figure 4, the RL-based controller's response to the abruptly changed slurry flow is visualized along with the slurry flow itself. The RL-based controller considered is the resulting controller from the previous section, after the tuning procedure has converged. There is no tuning active when the controller is tested in this section.

In the real plant, the level control is governed by cascade coupled PI-controllers. We have implemented a digital twin of the control system, including actual parameter values, in our simulation environment. This way, the control structure of the real plant and the RL-based controller can be exposed to the same disturbances and their performance can be compared. In Figure 5, the performance of the PI-controllers is shown when the system is exposed to the same inflow disturbance as in Figure 4.

The level in the buffer tank is not of interest as long as it does not overflow or become empty. It can be seen from Figures 4 and 5 that this requirement is met by both the RL-based controller and the PI-structure.

The effects on the levels in the flotation cells are of bigger interest. First observing the noise canceling properties of the controllers in steady state, it can be observed from Figures 4 and 5 that they both have good noise canceling properties. Looking at the root mean square error (RMSE) for a typical steady state sequence, the RL-based controller reduces it with roughly 50 % compared to the PI-controllers for the flotation cells. How ever, the noise canceling properties of the PI-controllers is already satisfactory.

When the big inflow disturbance occurs at time $t = 10$ in Figures 4 and 5, the levels in the cells are affected by this. It is desired that the amplitude of the level deviations due to the disturbance should be as small as possible and that the level should return to its reference fast. It is visible in the figures that the impact of the disturbance is bigger for the PI-controllers than for the RL-based controller, both when it comes to the amplitude of the deviations caused by the inflow disturbance, and the duration of the effects of it. Comparing the maximum amplitude deviations in the flotation cells, it is reduced with 42 % on average by the RL-based controller compared to the PI-controllers. When it comes to the time it takes for the levels to return to their references, a tolerance around the reference is chosen and the time it takes for the level to return to the tolerated area is measured. This time is on average reduced by 75 % by the RL-based controller compared to the PI-controllers.

5. DISCUSSION

As was demonstrated in the previous section, the RL-based feedback controller has overall better performance than the cascade-coupled PI-controllers. It has both better noise canceling properties and better disturbance rejection. This serves to show that the method has potential to improve operation under varying conditions. However, there are of course practical considerations that need to be thoroughly mapped out before attempting large scale deployment of the investigated method. Below we discuss some of the more important ones.

In our studied example, the feedback gain successfully converged over just a few iterations of the algorithm. This requires that the control signal perturbations, $\boldsymbol{e}_k$ of equation (3) provide sufficient exploration. Choosing this exploratory component in a proper way can be an issue if the knowledge of the system is poor. If the tuning process is to be performed on the real system, the choice of $\boldsymbol{e}_k$ will also be a trade off. They should be chosen as small as possible not to disturb production more than necessary, while being big enough to ensure that the control signal performs proper exploration. Too little exploration during the collection of data can lead to a poor update of the state feedback law. This can lead to that the updated controller performs worse than the previous one, or even destabilises the system.

To design a suitable cost function for the controller, a conception of the relative importance of tacking versus control signal activity is needed. So even though the algorithm does not require a process model, some knowledge of the process is important to have to be able to design a good

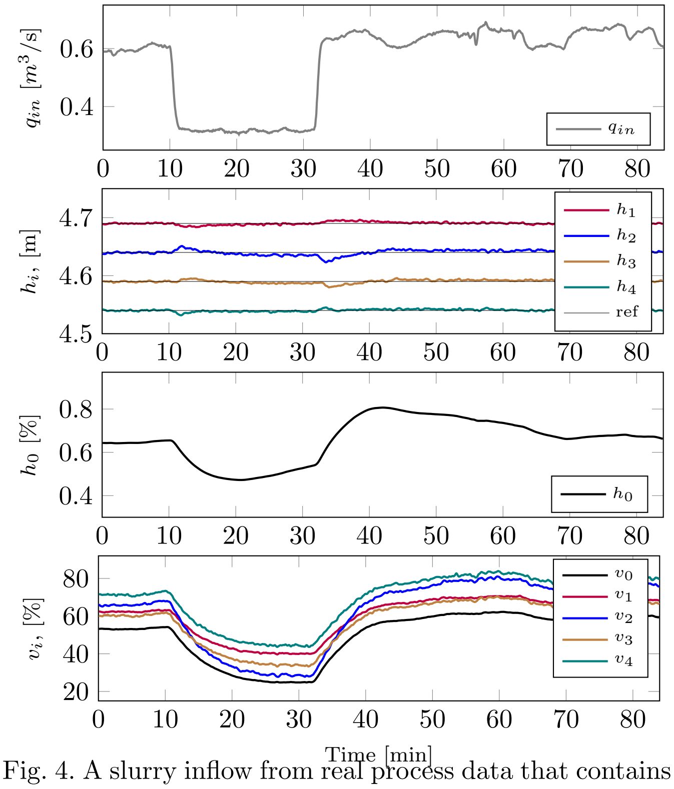

Fig. 4. A slurry inflow from real process data that contains typical disturbances is fed to the simulation model of the process consisting of a buffer tank and four flotation cells. The system is controlled by the model-free RL-based state feedback controller. The effect of the disturbances for the levels in the flotation cells and the buffer tank are shown along with the corresponding control signals.

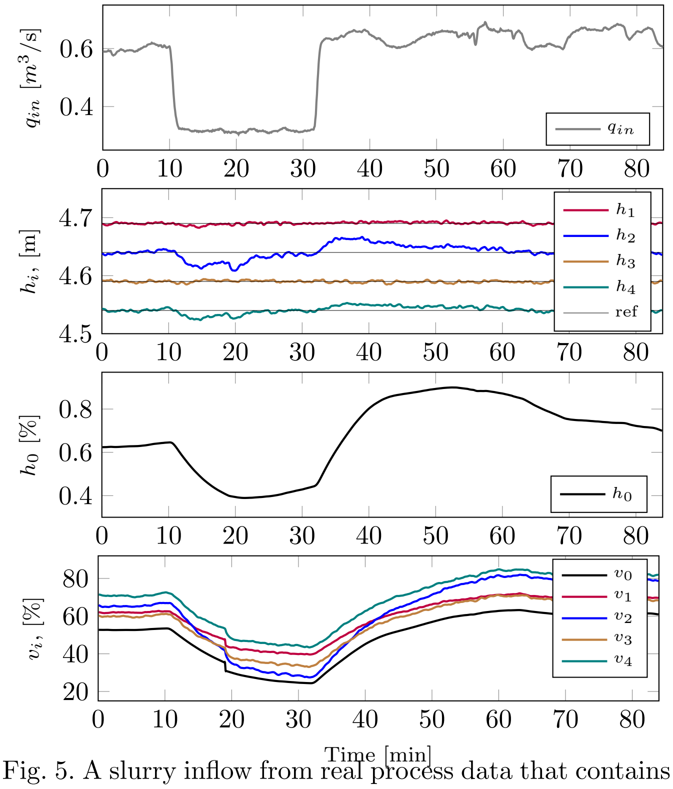

Fig. 5. A slurry inflow from real process data that contains typical disturbances is fed to the simulation model of the process consisting of a buffer tank and four flotation cells. The system is controlled by PI-controllers configured to be identical to those in the real plant. The effect of the disturbances for the levels in the flotation cells and the buffer tank are shown along with the corresponding control signals.

controller using the algorithm. For the initial choice of $\mathbf{L}$, it is also beneficial to have some previous process knowledge.

Looking at the data collection during the tuning procedure from a production point of view, the applied exploratory control signal has a large impact on production. To be practically feasible, further attention would need to be put on how to achieve adequate exploration, while maintaining adequate control performance.

Our study has shown that "model-free" RL-based control can achieve control performance that supersedes that of the currently implemented control system. However, this comes at the cost of initially exciting the dynamics more than would be practically admissible, to obtain sufficient model knowledge. Future work would therefore need to focus on improved experiment designs. Here methods that adjust the experiment online based on the system response, such as Berner and Soltesz (2017) could prove viable.

6. CONCLUSION

Summarizing the above observations, one could conclude that the RL-based control show good potential to design a state feedback controller without the need of an explicit process model. This makes it highly relevant for processes that are hard to model, where the process itself or its operating conditions change over time. However, the approach has a number of practical considerations that must be addressed when applying it to a real process. Particularly, an adequate balance between stable operation during exploration and sufficient excitation of the dynamics must be met.

REFERENCES

Bergh, L. and Yianatos, J. (2011). The long way toward multivariate predictive control of flotation processes. *Journal of Process Control*, 21(2), 226–234.

Berner, J. and Soltesz, K. (2017). Short and robust experiments in relay autotuners. In *2017 22nd IEEE International Conference on Emerging Technologies and Factory Automation (ETFA)*, 1–8. doi: 10.1109/ETFA.2017.8247696.

Bradtke, S. (1992). Reinforcement learning applied to linear quadratic regulation. *Advances in Neural Information Processing Systems*, 295–302.

Brooks, K. and Koorts, R. (2017). Model predictive control of a zinc flotation bank using online x-ray fluorescence analysers. *IFAC-PapersOnLine*, 50, 10214–10219.

Lewis, F., Vrabie, D., and Vamvoudakis, K. (2012). Reinforcement learning and feedback control: Using natural decision methods to design optimal adaptive controllers. *IEEE Control Systems Magazine*, 32(6), 76—105.

Norlund, F. (2022). Comparison of Level Control Strategies for a Flotation Series in the Mining Industry. Student Paper.

Stenlund, B. and Medvedev, A. (2002). Level control of cascade coupled flotation tanks. *Control Engineering Practice, Mechatronics*, 10(4), 443–448.