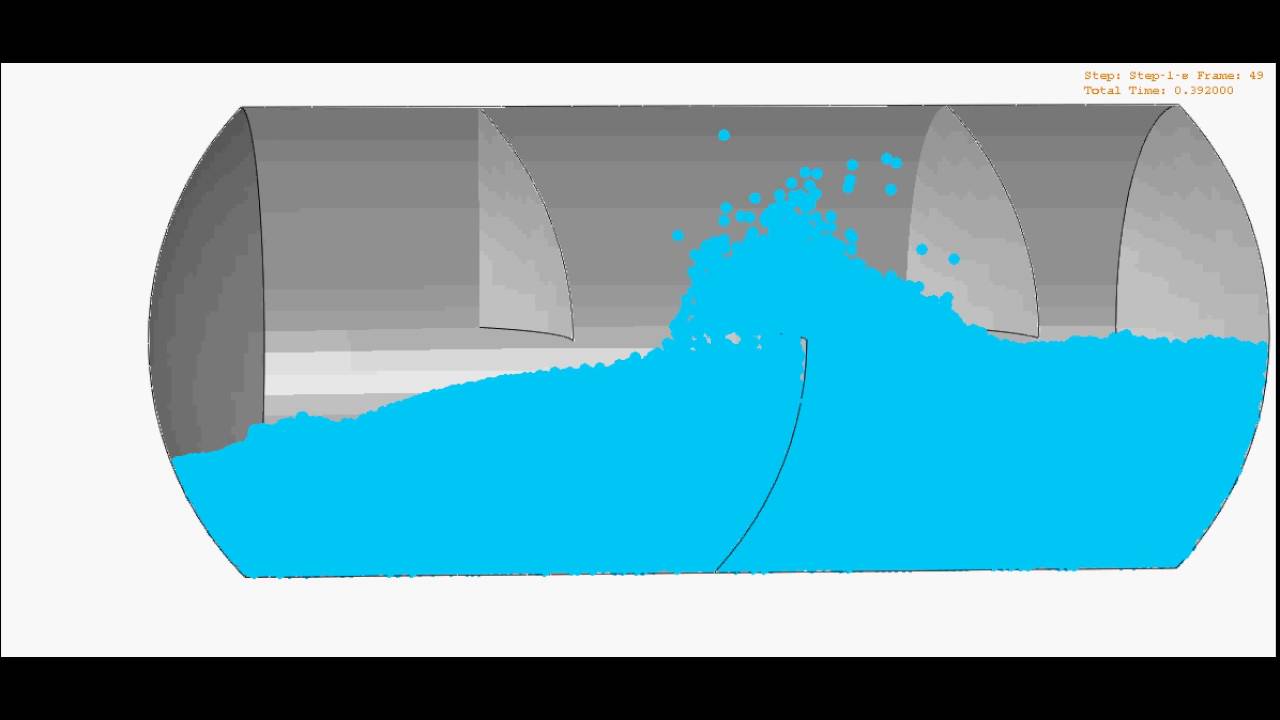
**ABSTRACT**

Sloshing is usually considered as the motion of a free liquid surface inside its containers. It is identified as a problem since it can significantly influence the system dynamics and response of a system. In a partially filled container, the sloshing problem is very common in many engineering applications such as oil storage tanks, liquefied natural gas storage tanks, liquid rocket fuel tanks, melted metal handling in steel plants, beverage industry, and attitude or trajectory maneuvers in spacecraft. In the presence of attitude or trajectory maneuvers of partially filled spacecraft, liquid sloshing often imposes significant effects on the motion of spacecraft and sometimes even induces instability**.** Therefore, the fuel slosh and its effect on the attitude of space vehicles have been studied extensively during the last few decades**.**

The interaction of the slosh with the container might lead to instabilities. The use of industrial robots to handle liquid containers in industrial sectors also encounters the sloshing problem that has attracted the attention of scientists and technologists. This problem is highly challenging due to the complexity of nonlinear fluid dynamics that arises within the vessel when it moves in a generic 3D trajectory. Thus, it is critical to design a controller that can effectively suppress slosh to maintain safe operations under all conditions.

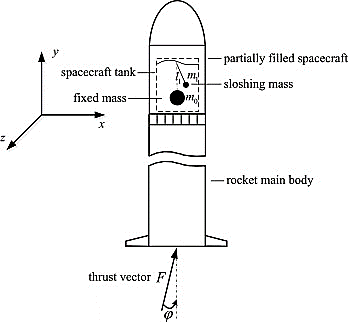
* + - 1. **Introduction**

Sloshing of liquid in a tank is critical in several areas, including launch vehicles carrying liquid fuel, satellites, industrial packaging of liquids, systems handling molten metal, and so on. Liquid tanks exist in many mechanical systems and play a dispensable role in modern industry and frontier science fields, such as oil storage tanks, liquefied natural gas storage tanks, and liquid rocket fuel tanks. Free-surface motions of liquid under gravity in tanks are of practical importance. The sloshing problem in a tank has an important application prospect, which interacts strongly with the container that limits its motion. It is known that partially filled tanks are prone to violent sloshing under certain motions. The large liquid movement creates highly localized impact pressure on tank walls. The interaction of the slosh with the container might lead to instabilities in some cases. Hence modeling, characterization, and control of nonlinear slosh phenomena are important in many applications.



**Fig. 1:** Slosh phenomenon in tanks

In the presence of attitude or trajectory maneuvers of partially filled spacecraft, liquid sloshing often imposes significant effects on the motion of spacecraft and sometimes even induces instability [1]. In particular, the attitude control problem for flexible spacecraft with fuel sloshing is the most critical problem of this area. In many space applications, it is necessary to place large and complex spacecraft (SC) in high altitude orbits. During translational or rotational accelerations of the SC, a large amount of fuel tends to move uncontrollably in the fuel tank, resulting in the sloshing phenomenon. The nonlinear coupling of the SC dynamics and the slosh dynamics can disturb the attitude of the SC from the desired position. We have already seen some failures due to fuel sloshing in the past [2].



**Fig. 2:** Fuel sloshing in spacecraft [3]

Hence, it is important to study and accurately characterize the phenomenon of sloshing and to develop, identify, and experimentally verify simple mathematical models of slosh that can be used for mission simulation and control development. Therefore, the fuel slosh and its effect on the dynamics of space vehicles have been studied extensively during the last few decades [4], [5], [6], [7]. Launch vehicles usually see the lateral and pitching excitations while in motion. Even if the liquid-filled spacecraft is modeled using a mechanical equivalent model, it is still a challenging problem for the control design for such a system. Apparently, it is an underactuated system since the sloshing mode is unable to be directly controlled. Controller for planar and three-dimensional spacecraft was designed in [7] and [8] respectively, based on underactuated system control design approach.

As the current industrial scenario becomes increasingly competitive, modern manufacturing systems must comply with strict, demanding functional requirements. Industrial automation plays a key role in optimizing manufacturing processes and maximizing production efficiency and flexibility while reducing waste and energy consumption. As a consequence, industrial robotics is pervading non-traditional industrial sectors to overcome the severe and unacceptable limitations due to the use of rigid automation systems. The resulting novel use of industrial robots in non-traditional industrial sectors has created new challenges that have attracted the attention of scientists and technologists.

In industrial automation, one of the key challenges is the use of industrial robots to handle liquid materials. A particularly interesting and promising case is the use of industrial robots in the food processing industry to accomplish complex tasks – i.e., moving, manipulating, pouring liquid ingredients, traditionally executed by human operators or by means of process control systems. In light of this, research in this area is gaining growing importance driven by the food and beverage industry needs to increase production performances, dependability and flexibility. As an example, the steel industry could address the problem of handling melted metal by means of industrial robots. These industrial cases can be solved by addressing the general problem of controlling a robotic manipulator to move a liquid-filled vessel without spilling.

Nonlinear fluid dynamics that arise within the vesselmake this problem highly challenging. Sloshing is extremely detrimental to industrial application performances and very complex to counteract.The use of industrial robots in industrial sectors also has the sloshing problem that has attracted the attention of scientists and technologists. It is critical to design a controller that can effectively suppress slosh to maintain safe operations.

**CHAPTER**

**2**

**Literature Survey & Proposed Objectives**

* + - 1. **Literature Survey & Proposed Objectives**

**2.1. Introduction**

This chapter contains the literature survey on the slosh control problem. Section 2.2 contains the slosh control strategies used till now. Reinforcement Learning introduction is given in section 2.3. Some basic definitions and concepts in reinforcement learning are discussed in section 2.4. The importance of reinforcement learning over conventional control methods is reported in section 2.5. Proposed objectives are given in section 2.6. Finally, conclusions are drawn based on the above sections and are reported in section 2.7.

**2.2. Slosh control strategies used till now**

Slosh is the motion of the free liquid surface in response to the force applied to the liquid directly or indirectly. The motion of the liquid occurs in different forms based on the nature of the applied force, its container geometry, etc. Accordingly, there exist various sloshing phenomena [9], [10], e.g., lateral, rotational, swirling, or even chaotic, quasi-periodic. Slosh induced forces and moments on container walls are detrimental in applications like molten metal pouring in foundry [11], liquid cargo carriers [12], space launch vehicles [9] or industrial packaging machines [13], and so on.

In modern launch vehicles, weight reduction is given more importance, so structural components are made up of lightweight materials that have increased the ratio of propellant to total vehicle mass. Hence large-amplitude slosh and complex slosh phenomenon would be inevitable. In launch vehicles, sloshing can be induced by the motion of the vehicle itself enroute. Few incidents of failure in space missions due to slosh are noted in [2]. These make the study, characterization, and control of slosh extremely important, especially for space launch vehicles.

Slosh dynamics, coupled with vehicle dynamics, form an underactuated system. Underactuated systems [14], [15], [16] are the systems that have more configuration variables to be controlled than the number of independent actuators to control them. Underactuated systems are found in various application areas such as robotics, aerospace, flexible structure systems, etc. There are various reasons due to which the systems reveal the underactuation property. Many times, systems are designed as underactuated to reduce the cost, weight and energy consumption like for aircrafts, spacecrafts, etc. Sometimes due to actuator failure also, system becomes underactuated, while some time, underactuation property is imposed deliberately to study and get insight of complex control problems and to exploit various properties of the systems like stabilisablity or controllability. Sometimes system dynamics are such that the system becomes underactuated, e.g., vehicle dynamics gets coupled with slosh. The Slosh-container system is an important example of underactuated systems.

The different sloshing phenomena can be characterized and modeled by the fluid dynamics approach as in [9], [10], and [17]. Petit and Rouchon [17] faced some control issues with the fluid dynamic approach. This approach involves complex equations like Navier-Stoke’s equations describing the motion of the fluid to be computed in real-time. This can be done by Computational Fluid Dynamics (CFD) approach, but it is computationally expensive and is not control amenable [9], [10]**.** A nonlinear and complicated model might have a more accurate prediction for the sloshing dynamics [12], [4]**,** [18]**,** [19]**,** but the dynamics become too challenging for the controller design.

This calls for the need for simplified mathematical models for slosh, to reduce computational time and cost, and offer control amenable models. In this case, equivalent mechanical models are quite useful since they provide simplified equations of motion of fluid by assuming oscillating point masses and rigid bodies, which make control design simpler. Pendulum and spring-mass damper are commonly used models where moving mass infers the sloshing mass. The pendulum model for slosh is discussed in [9] and [10]. They show that lateral slosh, a kind of linear and planar slosh, can be represented by a simple pendulum [9], [10], [20].

Representation of lateral sloshing by a simple pendulum model has been widely accepted and reported in the literature [9], [10]. Nonlinear sloshing arises from many reasons such as container geometry, external perturbations to the container, large-amplitude excitation near the resonance, etc. These would result in rotary, swirling, and chaotic types of sloshing phenomena. The spherical pendulum model is known to have several similar instabilities [9], [21], [22]. As mentioned earlier, rotary slosh is a kind of a nonlinear slosh that may arise by lateral harmonic excitation of a liquid-filled container. This generates a large amount of non-planar forces observed by [22], [23], which can disturb the stability of the system under consideration.

Numerous scientists have worked to provide solutions to the challenging problems posed by the sloshing dynamics. Different passive control techniques like baffles [9], [10] are reported to control the sloshing effects in launch vehicles specifically and in other applications alike. But it adds more weight to the system and hence increases the cost, making it less desirable. For the last two decades, there has been increased interest among researchers to explore active control strategies for slosh suppression [21], [24], [25], [26], [27], [28], [29]. In relation to the automatic machinery sector, many solutions have been proposed to avoid sloshing flows when moving liquid-filled containers. All these methodologies used in these studies rely on the assumption of a simplified linear model of the slosh dynamics, and sloshing is tackled as a model-based disturbance suppression.

Some have used control schemes such as PID control [30]**,** sliding-mode control [25],[31], [32], [33] **,** *H∞* control [34], [35], [36]**,** adaptive nonlinear dynamic inversion control [37]**,** linear quadratic regulator, linear quadratic Gaussian control [38]**,** Lyapunov-based feedback control [39], [40], [41]**.** As simple mechanical models are approximate models, they invariably leave us with model uncertainties and approximated nonlinearities, and hence call for the robust control design. Sliding Mode Control (SMC) is a well-developed and effective, robust control technique [42], [43], [44], [45], which offers robustness against matched uncertainties and disturbances and hence promising for the slosh control problem. SMC for the slosh problem was used previously in [25], [26], [27], [28] and [29].

It is important to stress that the majority of these studies consider simple one-dimensional motions, and, even when 3D motions are introduced, the slosh phenomenon is decoupled along the different motion axes and treated as a set of independent problems [46]. Finally, when multi-degrees-of-freedom robotic manipulators are taken into consideration, the possibility of changing both the position and the orientation of the vessel as described in [47] and [48] is generally ignored.

**2.6. Proposed Objective**

**Finding control actions using reinforcement learning for slosh control problem in various 2D trajectories.**

With respect to the motion path, traditionally, sloshing is analyzed for one-dimensional motion, where the sloshing is often modeled as a simple pendulum. The research work that addresses slosh suppression in higher dimensional space is relatively less. The aim here is to generate trajectories with concurrent consideration of motion time minimization and slosh suppression in a two-dimensional working space using reinforcement learning.

**2.7. Conclusion**

Reinforcement learning is becoming more popular today due to its broad applicability to solving problems relating to real-world scenarios. RL can be used in real-world problems where complex sequential decision-making is required. RL is quite useful where system models are unavailable or too difficult and expensive to build. It has found significant applications in fields such as game playing, robotics, vehicle navigation, intelligent transportation system, healthcare, recommender systems, and business management problems.

In the RL-based control systems, the big opportunity is to create new problem formulations where background foundational knowledge from control might be creatively mixed with RL to open new application domains or extend well beyond current performance objectives, especially for rapid changes in the environment and high levels of uncertainty. A key challenge here is to combine high-dimensional sensory inputs into learning control actions. The ability of the RL agent to achieve performance that exceeds all prior algorithms and a level that is comparable to a professional human player can be utilized in different control problems. These advancements in RL create an opportunity to examine the analysis and design of such artificial agents from a control theory perspective.

For the slosh control problem, a nonlinear and complicated model might have a more accurate prediction for the sloshing dynamics**,** but the dynamics become too challenging for the controller design. In previous slosh control approaches, simple mechanical models are preferred over complex ones and hence given approximated models invariably contain model uncertainties or mismatches and approximated nonlinearities. So for the slosh control problem, RL approach can be more effective than the conventional control strategies, as it can bypass the need to exactly model the complex liquid surface dynamics.

**REFERENCES**

[1] C. Nichkawde, P. M. Harish, and N. Ananthkrishnan, “Stability analysis of a multibody system model for coupled slosh-vehicle dynamics,” *J. Soundand Vib.*, vol. 275, pp. 1069–1083, 2004.

[2] J. P. Vreeburg, “Spacecraft maneuvers and slosh control,” *IEEE Control Syst. Mag.*, vol. 25, pp. 12–16, 2005.

[3] D. L. Cui, S. Z. Yan, X. S. Guo, and R. X. Gao, “Parametric resonance of liquid sloshing in partially filled spacecrafttanks during the powered-flight phase of rocket,” *Aerosp. Sci. Technol.*, vol. 35, pp. 93–105, 2014.

[4] L. D. Peterson, E. F. Crawley, and R. J. Hansman, “Nonlinear fluid slosh coupled to the dynamics of a spacecraft,” *AIAA*, vol. 27, pp. 1230–1240, 1989.

[5] D. E. Hill and J.R. Baumgarten, “Control of spin-stabilized spacecraft with sloshing fluid stores,” *Dyn. Syst. Meas. Control*, vol. 114, pp. 728–731, 1992.

[6] Y. Baozeng and Z. Lemei, “Hybrid control of liquid-filled spacecraft maneuvers by dynamic inversion and input shaping,” *AIAA*, vol. 52, pp. 618–626, 2014.

[7] Y. Baozeng, W. Wenjun, and Y. Yulong, “Modeling and coupling dynamics of the spacecraft with multiple propellant tanks,” *AIAA*, vol. 54, pp. 3608–3618, 2016.

[8] “Liquid sloshing in the automation industry.” [Online]. Available: https://www.vectorstock.com/royalty-free-vector/robotic-arm-working-on-conveyor-belt-with-bottles-vector-14924453.

[9] H. N. Abramson, “The dynamic behavior of liquids in moving containers,Technical Report, NASA, Washington, DC, USA,” 1966.

[10] R. A. Ibrahim, V. N. Pilipchuk, and T. Ikeda, “Recent Advances in Liquid Sloshing Dynamics,” *Appl. Mech. Rev.*, vol. 54, pp. 133–199, 2001.

[11] K. Yano, S. Higashikawa, and S. Terashima, “Motion control of liquid container considering an inclined transfer path,” *Control Eng. Pract.*, vol. 10, pp. 465–472, 2002.

[12] T. Acarman and ¨U. ¨Ozg¨uner, “Rollover prevention for heavy trucks using frequency shaped sliding mode control,” *Veh. Syst. Dynam- ics, Taylor Fr.*, vol. 44, pp. 737–762, 2003.

[13] M. Grundelius, “Iterative Optimal Control of Liquid Slosh in an Industrial Packaging Machine,” in *IEEE Conference on Decision and Control, Sydney, Australia,* 2000, pp. 3427–3432.

[14] M. W. Spong, “Underactuated mechanical Systems,” *Control Probl. Robot. Autom. Lect. Notes Control Inf. Sci.*, vol. 230, pp. 135–150, 1998.

[15] M. W. Spong, “Partial feedback linearization of underactuated mechanical systems,” in *International Conference on Intelligent Robots and Systems*, 1994, pp. 314–321.

[16] R. Olfati-Saber, “Nonlinear control of underactuated mechanical systems with application to robotics and aerospace vehicles, PhD Thesis,” Massachusetts Instititute of Technology, Cambridge, Massachusetts, United States, 2001.

[17] N. Petit and P. Rouchon, “Dynamics and solutions to some control problems for water-tank systems,” *IEEE Trans. Autom. Control*, vol. 47, pp. 594–609, 2002.

[18] D. D. Kana, “Validated spherical pendulum model for rotary liquid slosh,” *J. Spacecr. Rockets*, vol. 26, pp. 188–195, 1989.

[19] P. J. Aston, “Bifurcations of the horizontally forced spherical pendulum,” *Comp. Meth, Appl. Mech Engg*, vol. 170, pp. 343–353, 1999.

[20] R. A. Ibrahim, *Liquid Sloshing Dynamics Theory and Applications*. Cambridge University Press, NewYork, USA, 2005.

[21] P. S. Gandhi and A. Duggal, “Active stabilization of lateral and rotary slosh in cylindrical tanks,” in *IEEE International Conf. ICIT,* 2009, pp. 10–13.

[22] D. D. Kana, “A Model for Nonlinear Rotary Slosh in Propellant Tanks,” *J. Spacecr.*, vol. 24, pp. 169–177, 2002.

[23] K. R. Sreehari, “Modeling and Experiments on Large Amplitude Slosh in Cylindrical Tank,” M.Tech Thesis, IIT Bombay, 2010.

[24] R. Venugopal and D. S. Bernstein, “State space modeling and active control of slosh,” in *Proceedings of the 1996 IEEE International Conference on Control Applications*, 1996, pp. 1072–1077.

[25] B. Bandyopadhyay, P. S. Gandhi, and S. Kurode, “Sliding mode observer based sliding mode controller for slosh-free motion through PID scheme,” *IEEE Trans. Ind. Electron*, vol. 56, pp. 3432–3442, 2009.

[26] P. S. Thakar, B. Bandyopadhyay, P. S. Gandhi, and S. Kurode, “Robust control of rotary slosh using integral sliding modes,” in *12th International Workshop on Variable Structure Systems (VSS), Mumbai, India*, pp. 440–445.

[27] S. Kurode, “Sliding Mode Control for slosh-free motion using state and output feedback: An observer based and fast output sampling approach,” PhD Thesis, IIT Bombay, 2009.

[28] B. Bandyopadhyay, S. Kurode, and P. S. Gandhi, “Sliding mode control for slosh-free motion-a class of underactuated system,” *Int. J. Adv. Mechatron. Syst.*, vol. 1, pp. 203–213, 2009.

[29] P. S. Thakar, B. Bandyopadhyay, and P. S. Gandhi, “Sliding mode control for an underactuated slosh-container system using nonlinear model,” *Int. J. Adv. Mechatron. Syst. Inderscience Publ.*, vol. 5, pp. 335–344, 2013.

[30] H. Sira-Ramirez, “A flatness based generalized PI control approach to liquid sloshing regulation in a moving container,” in *in Proc. Amer. Control Conf., Anchorage, AK, USA*, 2002, pp. 2909–2914.

[31] S. Kurode, S. Spurgeon, B. Bandyopadhyay, and P. S. Gandhi, “Sliding mode control for slosh-free motion using a nonlinear sliding surface,” *IEEE/ASME Trans. Mechatronics*, vol. 18, pp. 714–724, 2013.

[32] S. Kurode, B. Bandyopadhyay, and P. S. Gandhi, “Sliding mode control for slosh-free motion of a container using partial feedback linearization,” in *in Proc. Int. Workshop Variable Struct. Syst., Antalya, Turkey*, 2008, pp. 367–372.

[33] H. Richter, “Motion control of a container with slosh: constrained sliding mode approach,” *J. Dyn. Syst., Meas., Control*, vol. 132, pp. 1–10, 2010.

[34] J. Kuang and A. Y. T. Leung, “Feedback for attitude control of liquid-filled spacecraft,” *Guid. Control. Dyn.*, vol. 24, pp. 46–53, 2001.

[35] K. Terashima and K. Yano, “Robust liquid container transfer control for complete sloshing suppression,” *IEEE Trans. Control Syst. Technol*, vol. 9, pp. 483–493, 2001.

[36] K. Terashima and G. Schmidt, “Motion control of a cart-based container considering suppression of liquid oscillations,” in *in Proc. IEEE Int. Symp. Ind. Electron., Santiago, Chile*, 1994, pp. 275–280.

[37] E. D. Weerdt, E. V. Kampen, D. V. Gemert, Q. P. Chu, and J. A. Mulder, “Adaptive nonlinear dynamic inversion for spacecraft attitude control with fuel sloshing,” in *AIAA Guidance, Navigation and Control Conference and Exhibit*, 2008, pp. 1–24.

[38] A. G. de Souza and L. C. G. de Souza, “Satellite attitude control system design taking into account the fuel slosh and flexible dynamics,” in *Math. Probl. Eng.*, 2014, pp. 1–8.

[39] M. Reyhanoglu and J. R. Hervas, “Nonlinear modeling and control of slosh in liquid container transfer via a PPR robot,” *Commun. Nonlinear Sci. Numer. Simul.*, vol. 18, no. 6, pp. 1481–1490, 2013.

[40] M. Reyhanoglu and J. R. Hervas, “Nonlinear dynamics and control of space vehicles with multiple fuel slosh modes,” *Control Eng. Pr.*, vol. 20, no. 9, pp. 912–918, 2012.

[41] M. Reyhanoglu and J. R. Hervas, “Robotically controlled sloshing suppression in point-to-point liquid container transfer,” *J. Vib. Control*, vol. 19, no. 14, pp. 2137–2144, 2013.

[42] R. A. Decarlo, S. H. Zak, and G. P. Matthews, “Variable Structure Control of Nonlinear Multivariable Systems: A Tutorial,” *Proc. IEEE*, vol. 76, pp. 212–232, 1988.

[43] V. I. Utkin, “Variable Structure System with Sliding Modes,” *IEEE Trans. Autom. Control.*, vol. 22, pp. 212–222, 1977.

[44] J. Y. Hung, W. Gao, and J. C. Hung, “Variable Structure Control: A Survey,” *IEEE Trans. Ind. Electron.*, vol. 40, pp. 2–22, 1993.

[45] C. Edwards and S. K. Spurgeon, *Sliding Mode Control: theory and application*. Taylor and Francis, UK, 1998.

[46] S. J. Chen, B. Hein, and H. H. Worn, “Using acceleration compen-sation to reduce liquid surface oscillation during a high speed transfer,” in *in Proceedings 2007 IEEE International Conference on Robotics and Au-tomation.*, 2007, pp. 2951–2956.

[47] W. Aribowo, T. Yamashita, and K. Terashima, “Integrated trajectory planning and sloshing suppression for three-dimensional motion of liquidcontainer transfer robot arm,” *ournal Robot.*, pp. 1–15, 2015.

[48] L. Moriello, L. Biagiotti, C. Melchiorri, and A. Paoli, “Control of liquid handling robotic systems: A feed-forward approach to suppress sloshing,” in *IEEE International Conference on Robotics and Automation (ICRA), Singapore*, 2017, pp. 4286–4291.

[49] R. S. Sutton and A. G. Barto, *Reinforcement Learning An Introduction*. Cambridge, MA, USA: MIT Press, 2018.

[50] “Block Diagram of reinforcement learning.” [Online]. Available: https://www.mathworks.com/help/reinforcement-learning/ug/create-agents-for-reinforcement-learning.html.

[51] V. Mnih *et al.*, “Human-level control through deep reinforcement learning,” *Nature*, vol. 518, pp. 529–533, 2015.

[52] D. Silver *et al.*, “Mastering the game of Go with deep neural networks and tree search,” *Nature*, vol. 529, pp. 484–489, 2016.

[53] D. Silver *et al.*, “Mastering the game of Go without human knowledge,” *Nature*, vol. 550, pp. 354–359, 2017.

[54] J. Kober, J. A. Bagnell, and J. Peters, “Reinforcement learning in robotics: A survey,” *Int. J. Rob. Res.*, vol. 32, pp. 1238–1274, 2013.

[55] A. Amini *et al.*, “Learning Robust Control Policies for End-to-End Autonomous Driving from Data-Driven Simulation,” *IEEE Robot. Autom. Lett*, vol. 5, pp. 1143–1150, 2020.

[56] H. Khadilkar, “A Scalable Reinforcement Learning Algorithm for Scheduling Railway Lines,” *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 2, pp. 353–362, 2019.

[57] S. Liu, K. C. See, K. Y. Ngiam, L. A. Celi, X. Sun, and M. Feng, “Reinforcement Learning for Clinical Decision Support in Critical Care: Comprehensive Review,” *J Med Internet Res.*, vol. 22, pp. 1–16, 2020.

[58] S. Choi, H. Ha, U. Hwang, C. Kim, J. Ha, and S. Yoon, “Reinforcement Learning-based Recommender System using Biclustering Technique,” in *arXiv:1801.05532*, 2018, pp. 1–12.

[59] E. L. Thorndike, *Animal intelligence*. Macmillan Company, New York, 1911.

[60] A. Turing, “Intelligent Machinery: Report for National Physical Laboratory Universal Turing Machine,” 1948.

[61] C. E. Shannon, “‘Theseus’ maze-solving mouse.” [Online]. Available: http://cyberneticzoo.com/mazesolvers/1952-–-theseus-maze-solving-mouse-–-claude-shannon-american/.

[62] I. P. Pavlov and G. V. Anrep, *Conditioned reflexes; an investigation of the physiological activity of the cerebral cortex*. Oxford University Press, London, 1927.

[63] M. Minsky, “Steps Toward Artificial Intelligence,” in *Proceedings of the IRE*, 1961, vol. 49, no. 1, pp. 8–30, doi: 10.1109/JRPROC.1961.287775.

[64] M. D. Waltz and K. S. Fu, “A heuristic approach to reinforcement learning control systems.,” *IEEE Trans. Automat. Contr.*, vol. 10, pp. 390–398, 1965.

[65] J. M. Mendel, “A survey of learning control systems,” *ISA Trans.*, vol. 5, pp. 297–303, 1966.

[66] K. S. Fu, “Learning control systems—Review and outlook,” *IEEE Trans. Automat. Contr.*, vol. 15, pp. 210–221, 1970.

[67] D. Michie and R. A. Chambers, “BOXES, An experiment in adaptive control,” *Mach. Intell.*, vol. 2, pp. 137–152, 1968.

[68] A. H. Klopf, “Brain Function and Adaptive Systems: A Heterostatic Theory,” 1972.

[69] A. H. Klopf, “A comparison of natural and artificial intelligence,” *ACM SIGART Bulletin*, pp. 11–13, 1975.

[70] A. H. Klopf, *The Hedonistic Neuron: A Theory of Memory, Learning, and Intelligence*. Washington: Hemisphere Pub. Corp, 1982.

[71] R. S. Sutton, “Single channel theory: A neuronal theory of learning,” *Brain Theory Newsl.*, vol. 4, pp. 72–75, 1978.

[72] R. S. Sutton and A. G. Barto, “Toward a modern theory of adaptive networks: Expectation and prediction,” *Psychol. Rev.*, vol. 88, no. 2, pp. 135–170, 1981, doi: 10.1037/0033-295X.88.2.135.

[73] R. Sutton and A. G. Barto, “An adaptive network that constructs and uses an internal model of its world,” *Cogn. Brain Theory*, vol. 4, no. 3, pp. 217–246, 1981.

[74] R. Bellman, “A Problem in the Sequential Design of Experiments,” *Indian J. Stat.*, vol. 16, no. 3/4, pp. 221–229, 1956.

[75] R. Bellman, “Dynamic programming and stochastic control processes,” *Inf. Control*, vol. 1, no. 3, pp. 228–239, 1958.

[76] R. Bellman, *Dynamic programming*. Princeton University Press, 1972.

[77] M. L. Puterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley & Sons, 1994.

[78] R. Bellman, “A Markovian Decision Process,” *Indiana Univ. Math. J.*, vol. 6, no. 4, pp. 679–684, 1957.

[79] R. Howard, *Dynamic Programming and Markov Processes*. MIT Press, Cambridge, MA, 1960.

[80] P. . Webros, “Advanced Forecasting Methods for Global Crisis Warning and Models of Intelligence,” *Gen. Syst.*, vol. XXII, no. January 1977, pp. 25–38, 1977.

[81] C. J. C. H. Watkins, “Learning from delayed rewards,” 1989.

[82] D. P. Bertsekas and J. N. Tsitsiklis, *Neuro-Dynamic Programming*. Athena Scientific, Belmont, MA, 1996.

[83] A. L. Samuel, “Some Studies in Machine Learning,” *IBM J. Res. Dev.*, vol. 3, no. 3, pp. 210–229, 1959.

[84] C. E. Shannon, “XXII. Programming a computer for playing chess,” *London, Edinburgh, Dublin Philos. Mag. J. Sci.*, vol. 41, no. 314, pp. 256–275, 1950, doi: 10.1080/14786445008521796.

[85] I. H. Witten, “An adaptive optimal controller for discrete-time Markov environments,” *Inf. Control*, vol. 34, no. 4, pp. 286–295, 1977.

[86] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” in *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, 2015, pp. 1–14.

[87] A. Karpathy, G. Toderici, S. Shetty, T. Leung, R. Sukthankar, and F. F. Li, “Large-scale video classification with convolutional neural networks,” in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2014, pp. 1725–1732.

[88] V. Mnih *et al.*, “Playing Atari with Deep Reinforcement Learning,” in *arXiv preprint arXiv: 1312.5602*, 2013, pp. 1–9.

[89] V. Mnih *et al.*, “Human-level control through deep reinforcement learning,” *Nature*, vol. 518, no. 7540, pp. 529–533, 2015.

[90] D. Silver *et al.*, “Mastering the game of Go with deep neural networks and tree search,” *Nature*, vol. 529, no. 7587, pp. 484–489, Jan. 2016.

[91] D. Silver *et al.*, “Mastering the game of Go without human knowledge,” *Nature*, vol. 550, no. 7676, pp. 354–359, 2017.

[92] M. Campbell, A. J. Hoane, and F. H. Hsu, “Deep Blue,” *Artif. Intell.*, vol. 134, no. 1–2, pp. 57–83, 2002.

[93] K. Arulkumaran, M. P. Deisenroth, M. Brundage, and A. A. Bharath, “Deep reinforcement learning: A brief survey,” *IEEE Signal Processing Magazine*, vol. 34, no. 6, pp. 26–38, 2017.

[94] J. Kober, J. A. Bagnell, and J. Peters, “Reinforcement learning in robotics: A survey,” *Int. J. Rob. Res.*, vol. 32, no. 11, pp. 1238–1274, 2013.

[95] A. S. Polydoros and L. Nalpantidis, “Survey of Model-Based Reinforcement Learning: Applications on Robotics,” *J. Intell. Robot. Syst. Theory Appl.*, vol. 86, no. 2, pp. 153–173, 2017.

[96] R. S. Sutton and A. G. Barto, *Reinforcement Learning An Introduction*. Cambridge, MA, USA: MIT Press, 2018.

[97] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature*, vol. 521, pp. 436–444, 2015.

[98] D. Hassabis, D. Kumaran, C. Summerfield, and M. Botvinick, “Neuroscience inspired artificial intelligence,” *Neuron*, vol. 95, pp. 245–258, 2017.

[99] “Relation between AI, ML and DL.” [Online]. Available: https://www.ospreydata.com/2020/02/24/ai-ml-models-101-what-is-a-model/.