SELF-LEARNING CONTROLLERS IN CONTRAST TO ADAPTIVE CONTROLLERS

TRIAL LECTURE

Adriana Reyes Lúa

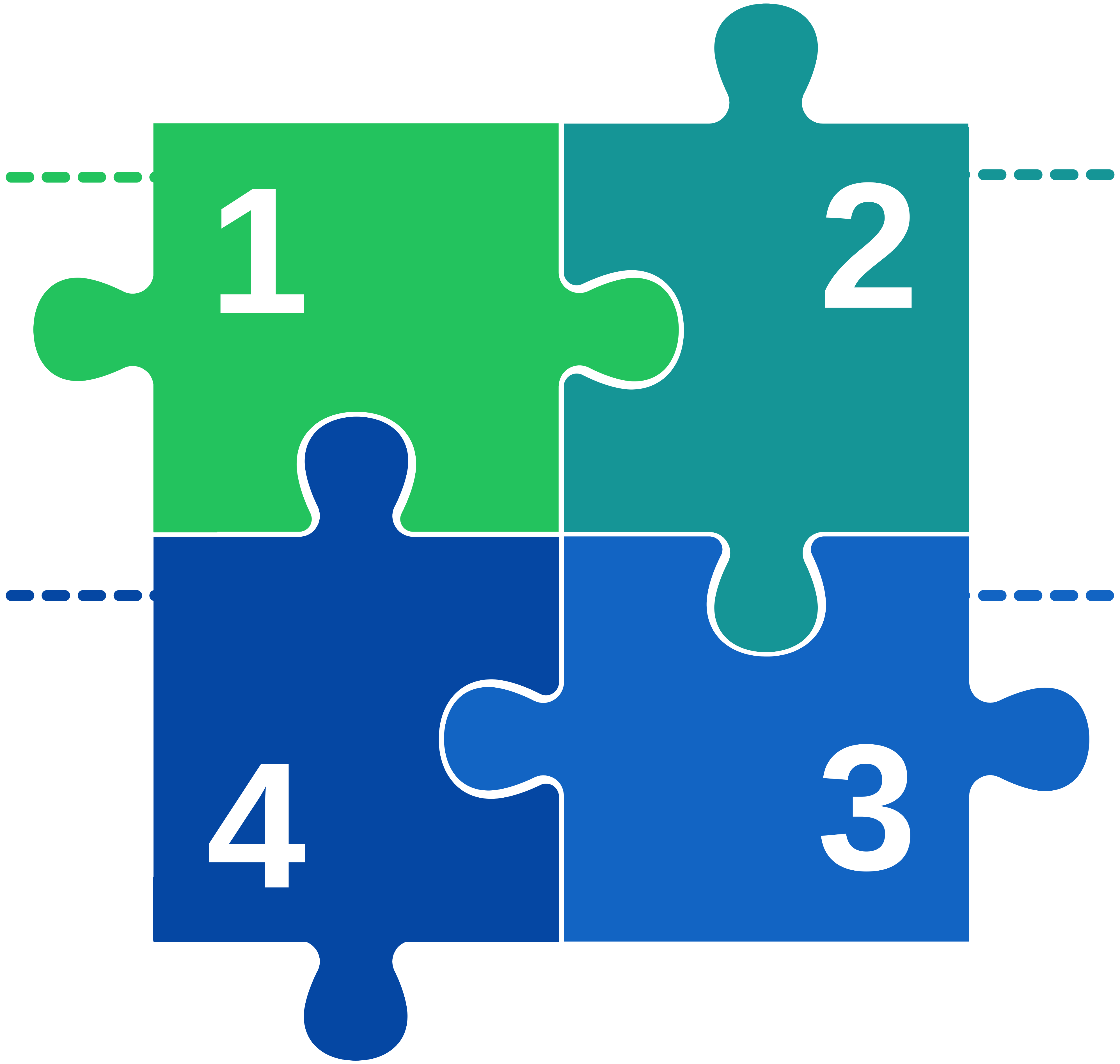
February 28th 2020

AGENDA

SELF LEARNING CONTROLLERS IN CONTRAST TO ADAPTIVE CONTROLLERS

BACKGROUND

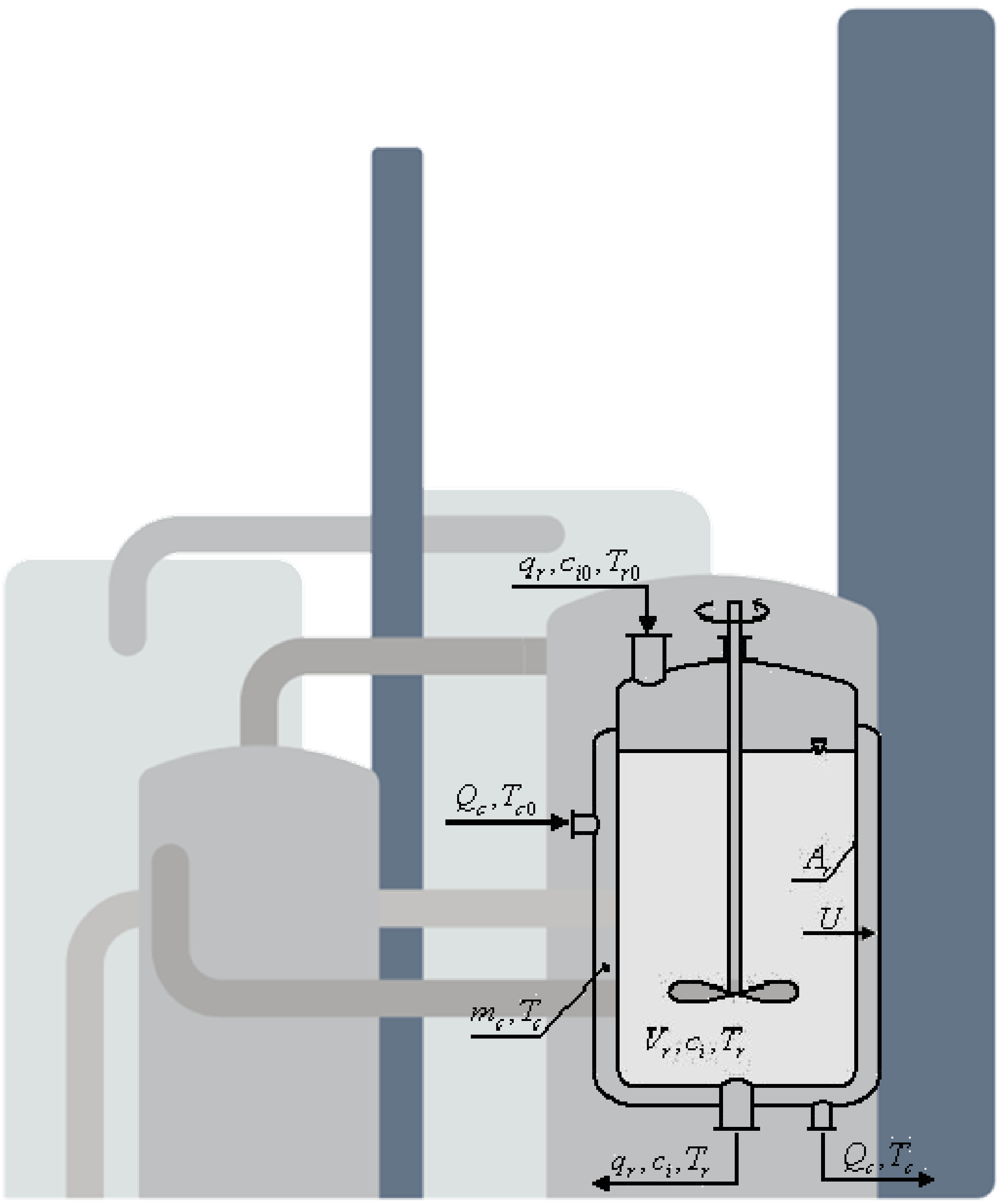
ADAPTIVE CONTROL

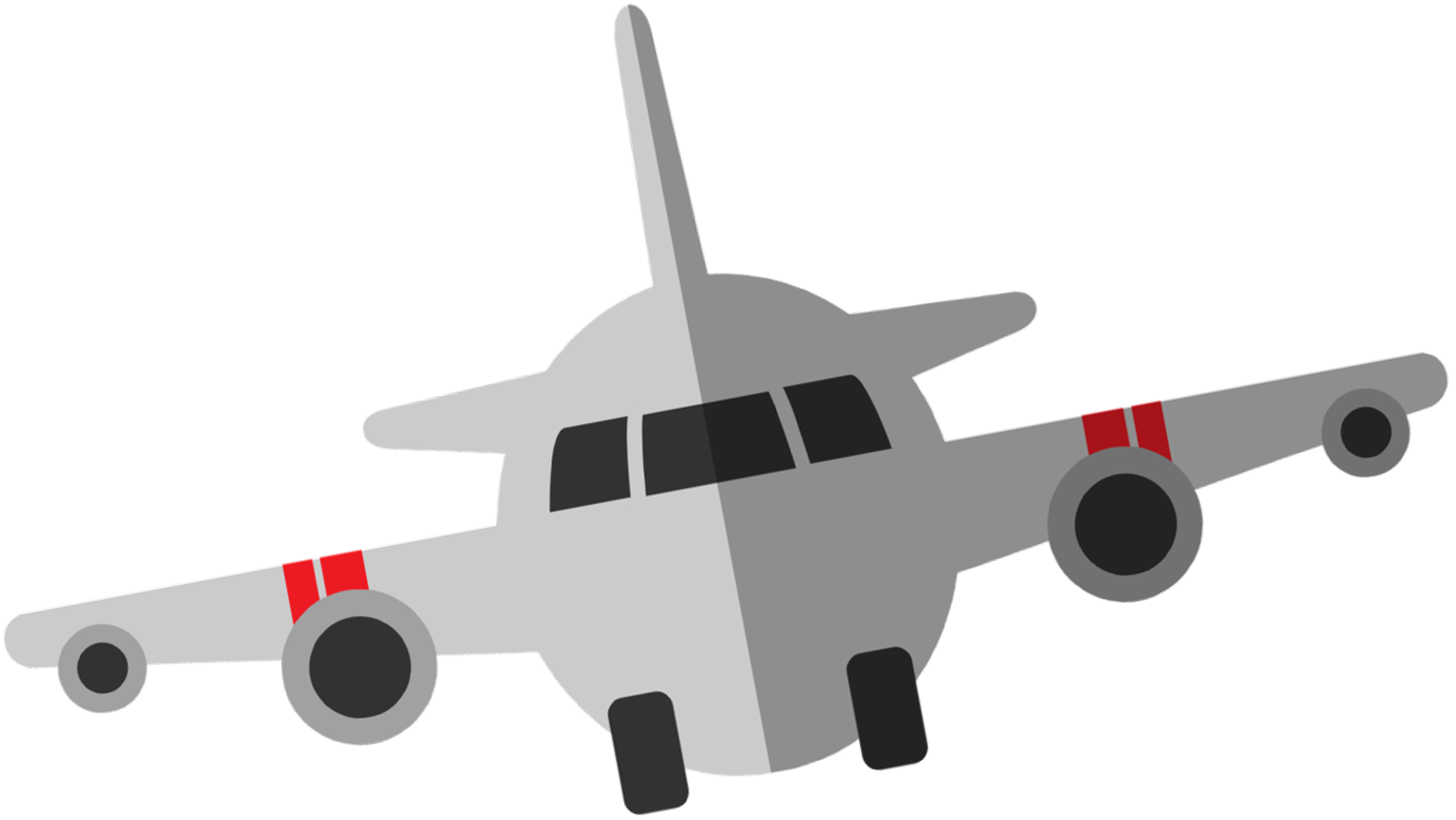


CONTRAST

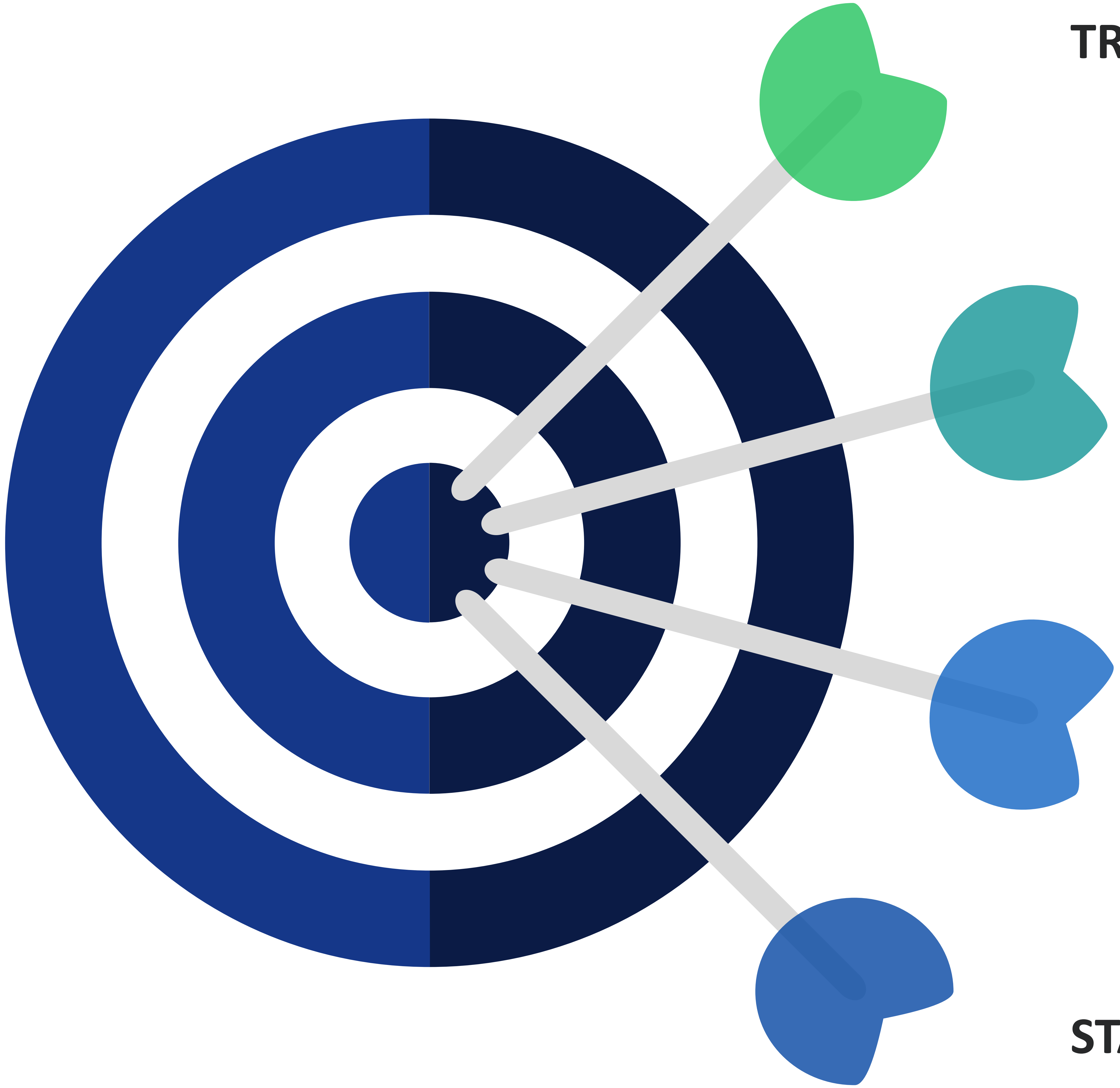
SELF-LEARNING CONTROL

PROCESS PLANT





CONTROL OBJECTIVES



TRACK REFERENCE/OPTIMIZE

COMPENSATE FOR UNCERTAINTY

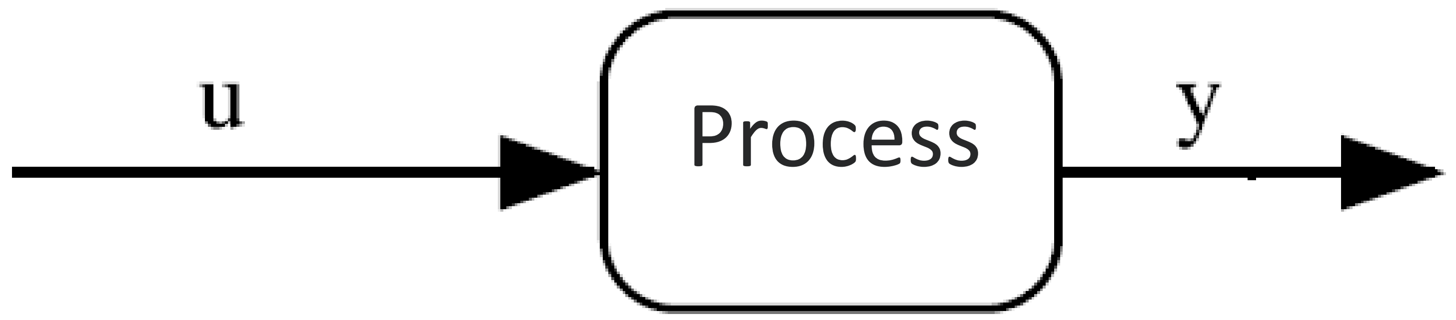
REJECT DISTURBANCES AND ATTENUATE NOISE

STABILIZE THE PROCESS

THE FEEDBACK LOOP

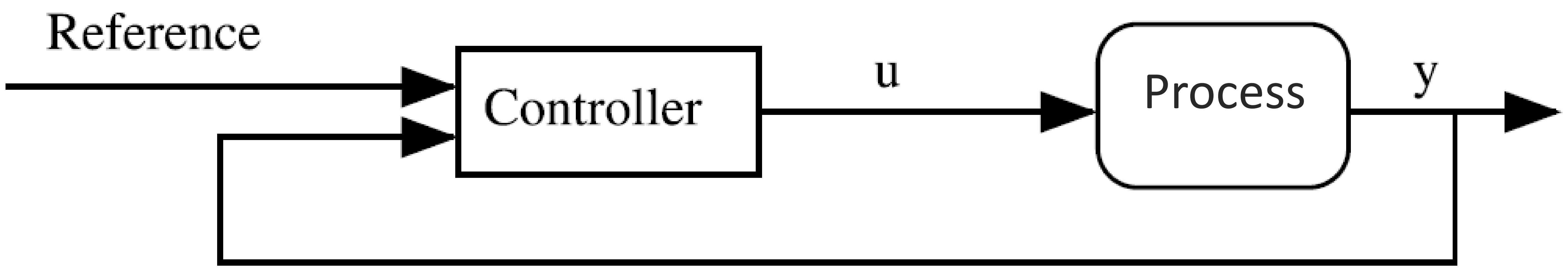
y = controlled variable

u = manipulated variable



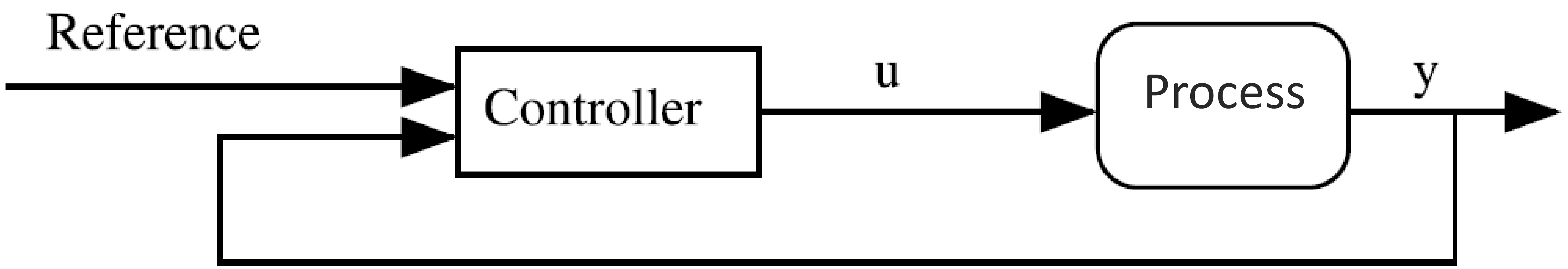
THE FEEDBACK LOOP

y = controlled variable
u = manipulated variable



TYPICAL CONTROLLER DESIGN

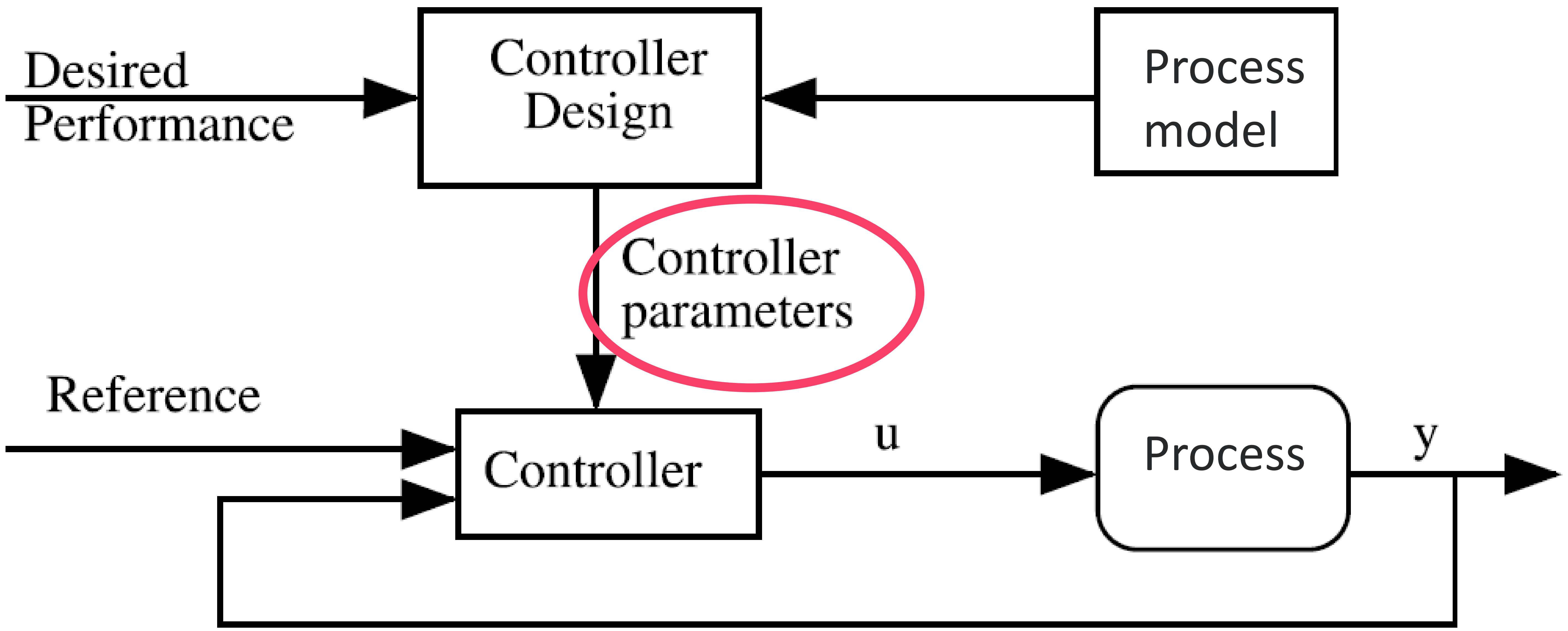
y = controlled variable
u = manipulated variable



PID-controller:

$$u(t) = u^0 + K_C\left(e(t) + \frac{1}{\tau_I}\int_0^t e(t) + \tau_d \frac{de(t)}{dt}\right)$$

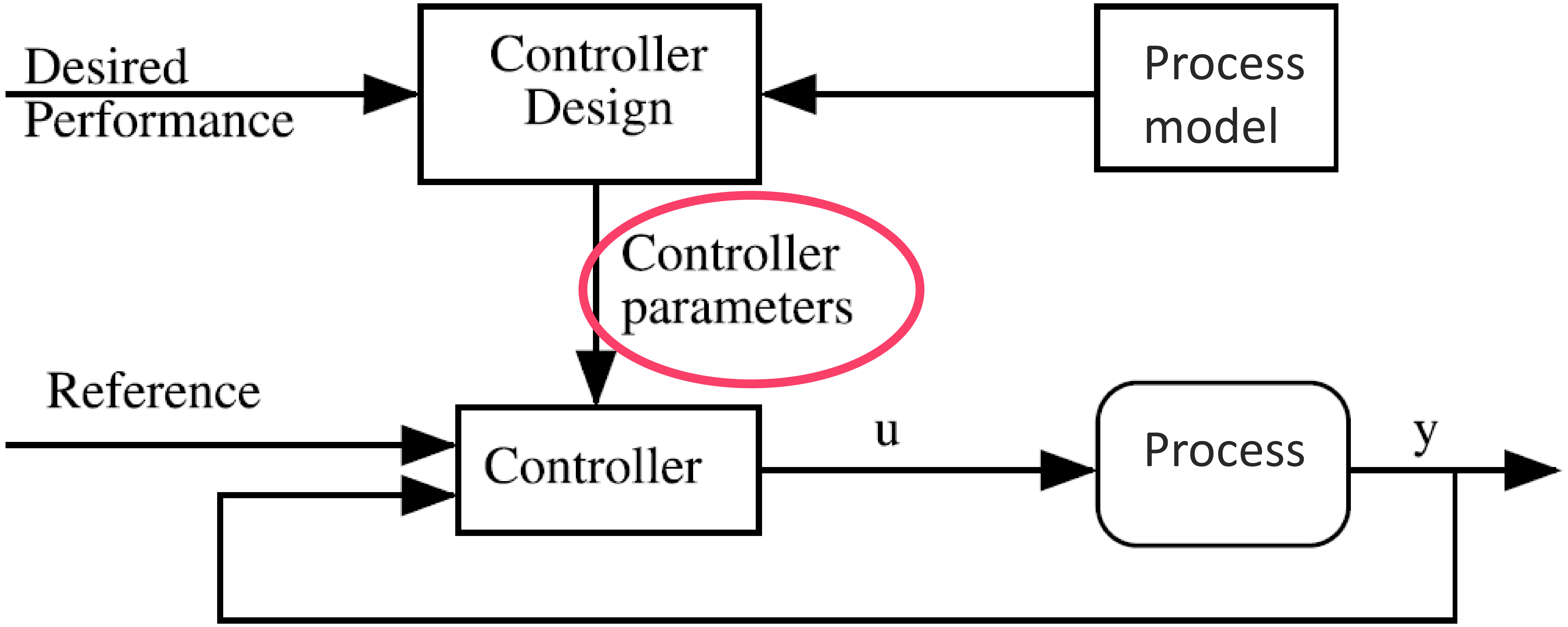
TYPICAL CONTROLLER DESIGN



PID-controller:

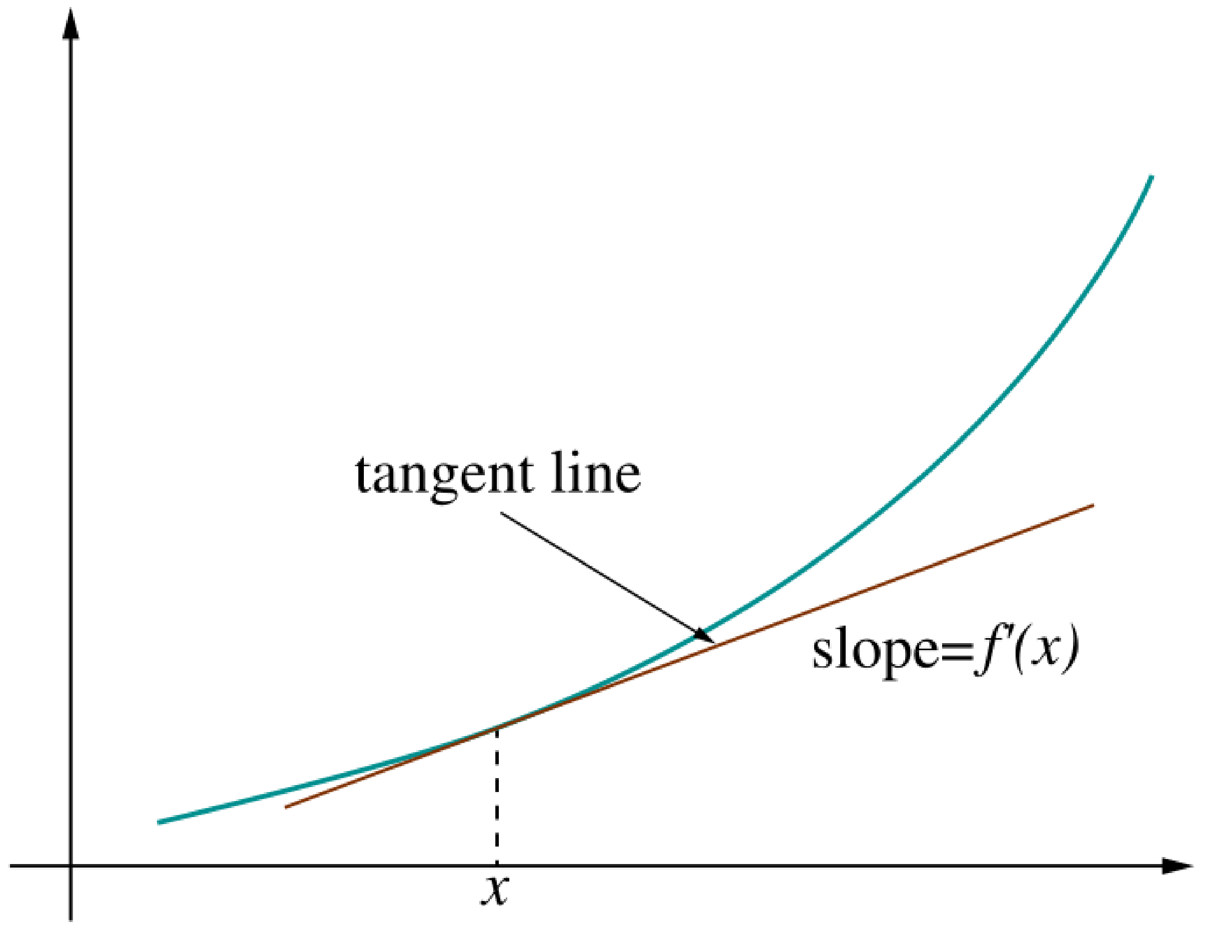
$$u(t) = u^0 + K_C\left(e(t) + \frac{1}{\tau_I}\int_0^t e(t) + \tau_d \frac{de(t)}{dt}\right)$$

TYPICAL CONTROLLER DESIGN

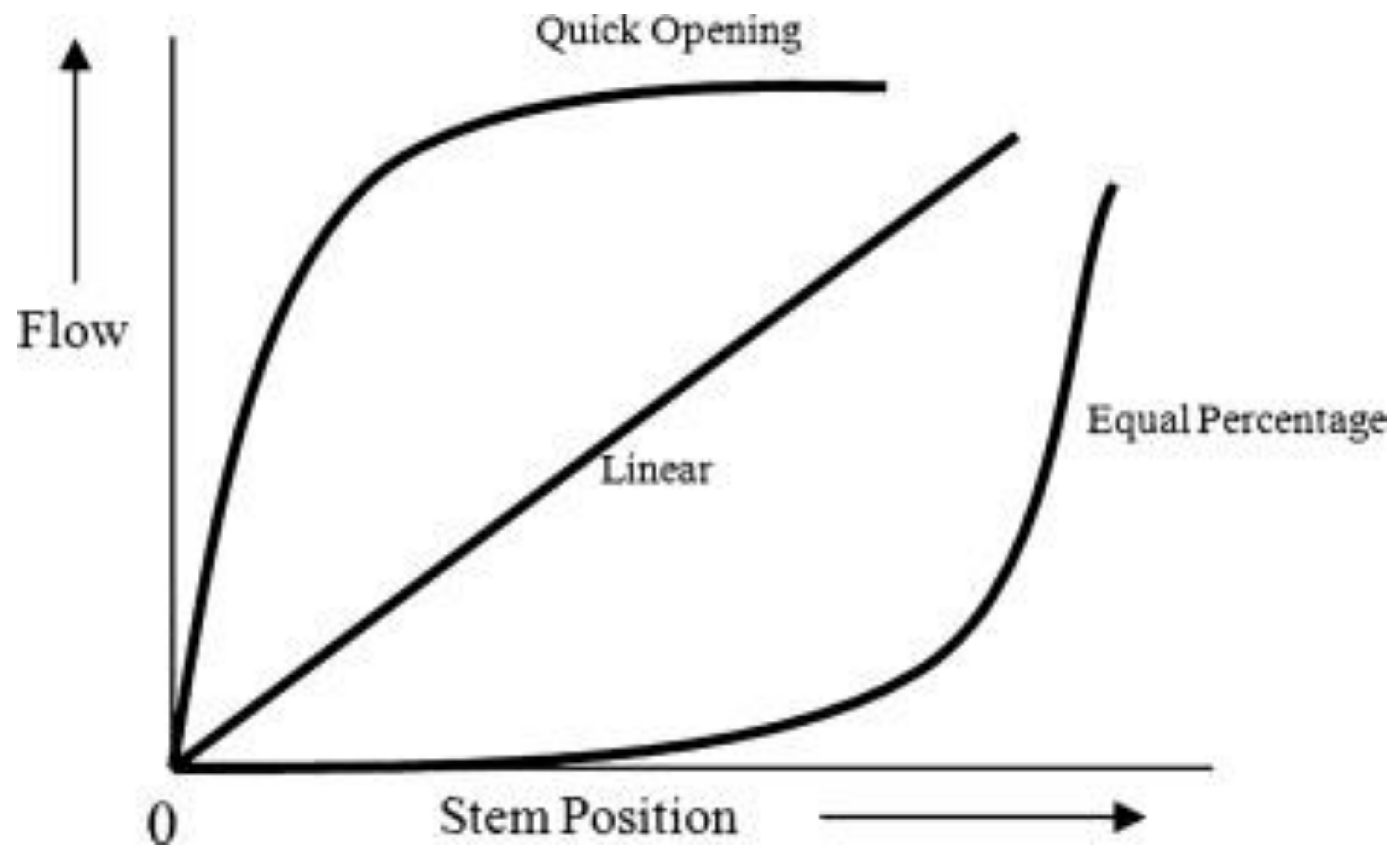


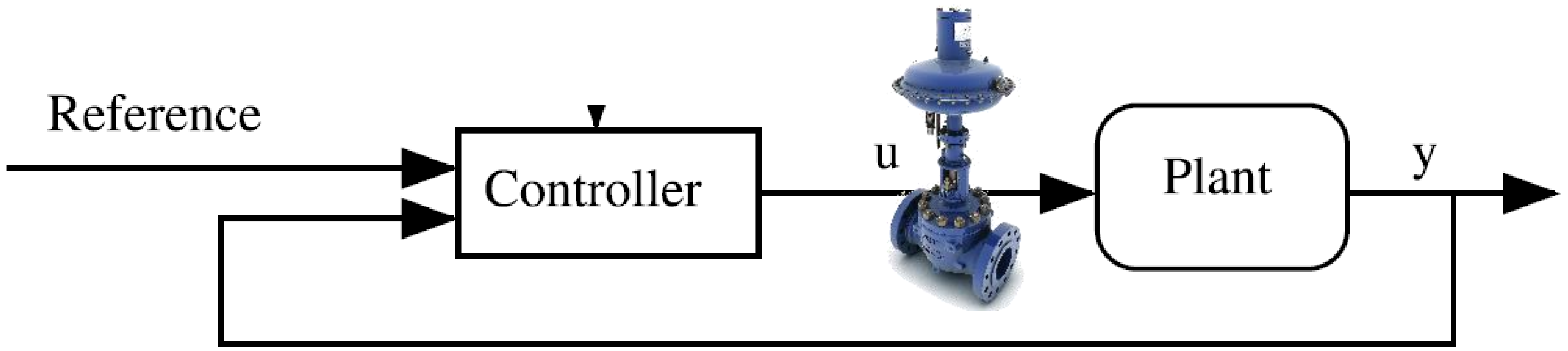
PID-controller:

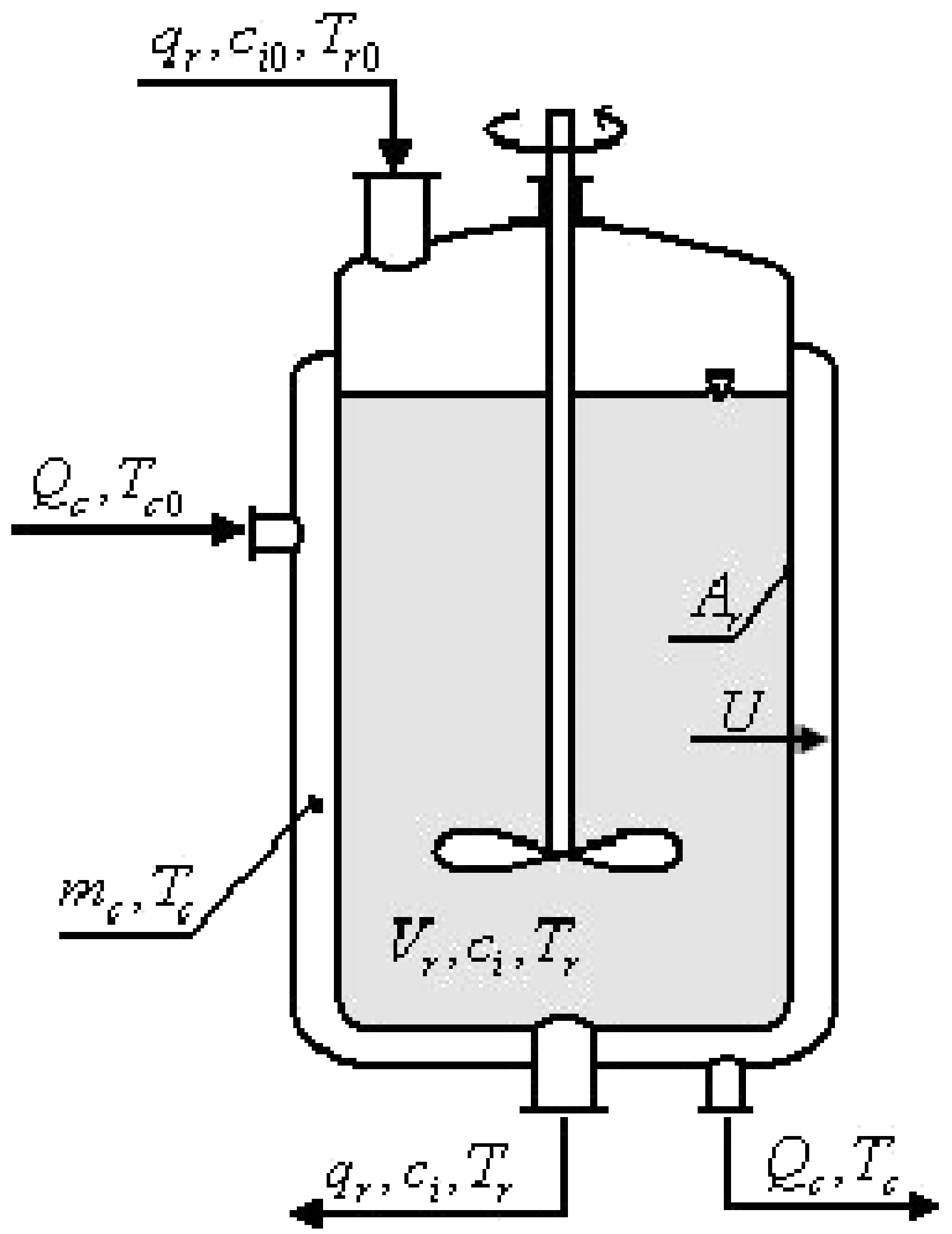
$$u(t) = u^0 + K_C\left(e(t) + \frac{1}{\tau_I}\int_0^t e(t) + \tau_d \frac{de(t)}{dt}\right)$$



VARIATIONS IN PARAMETERS



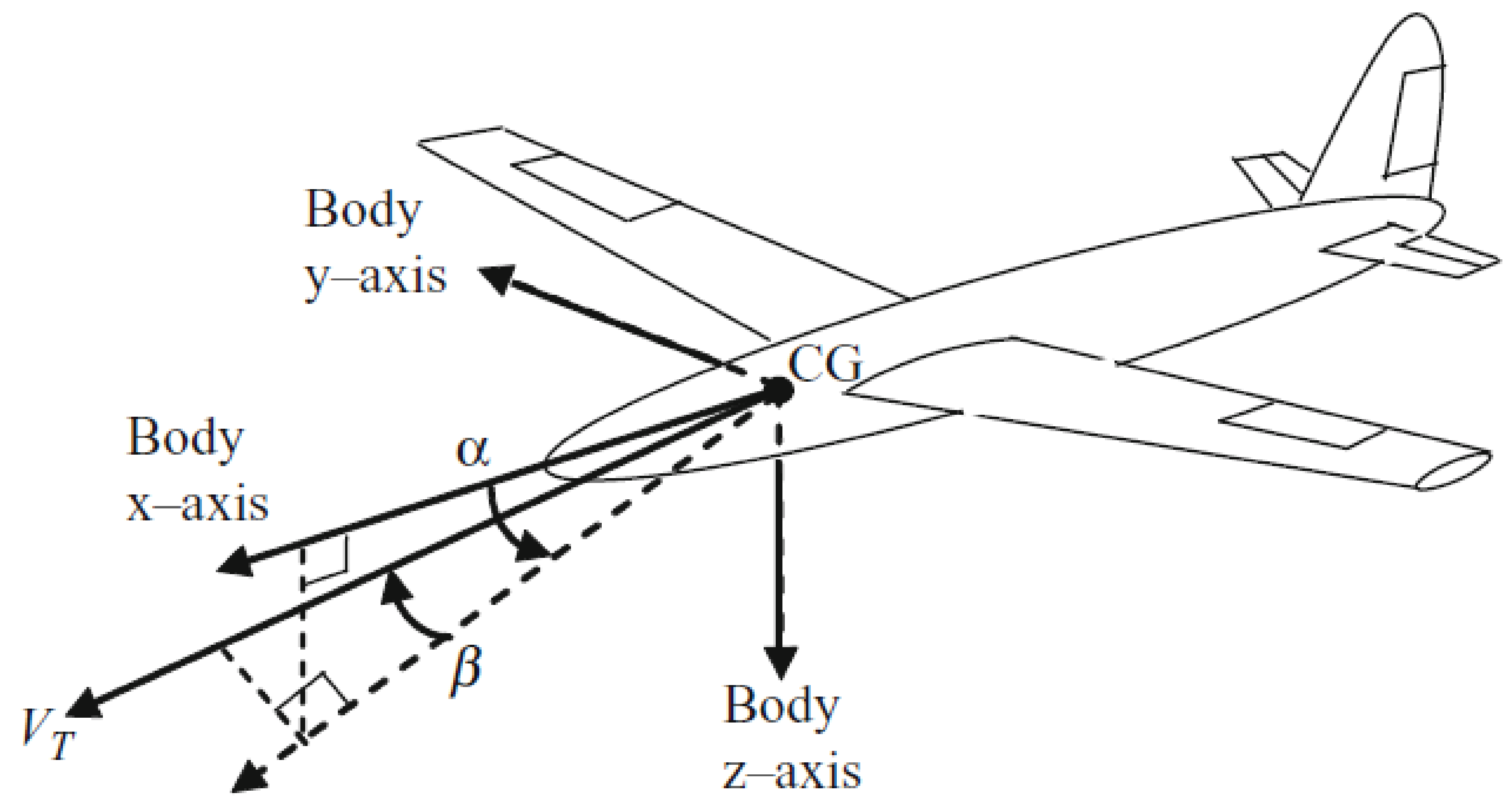


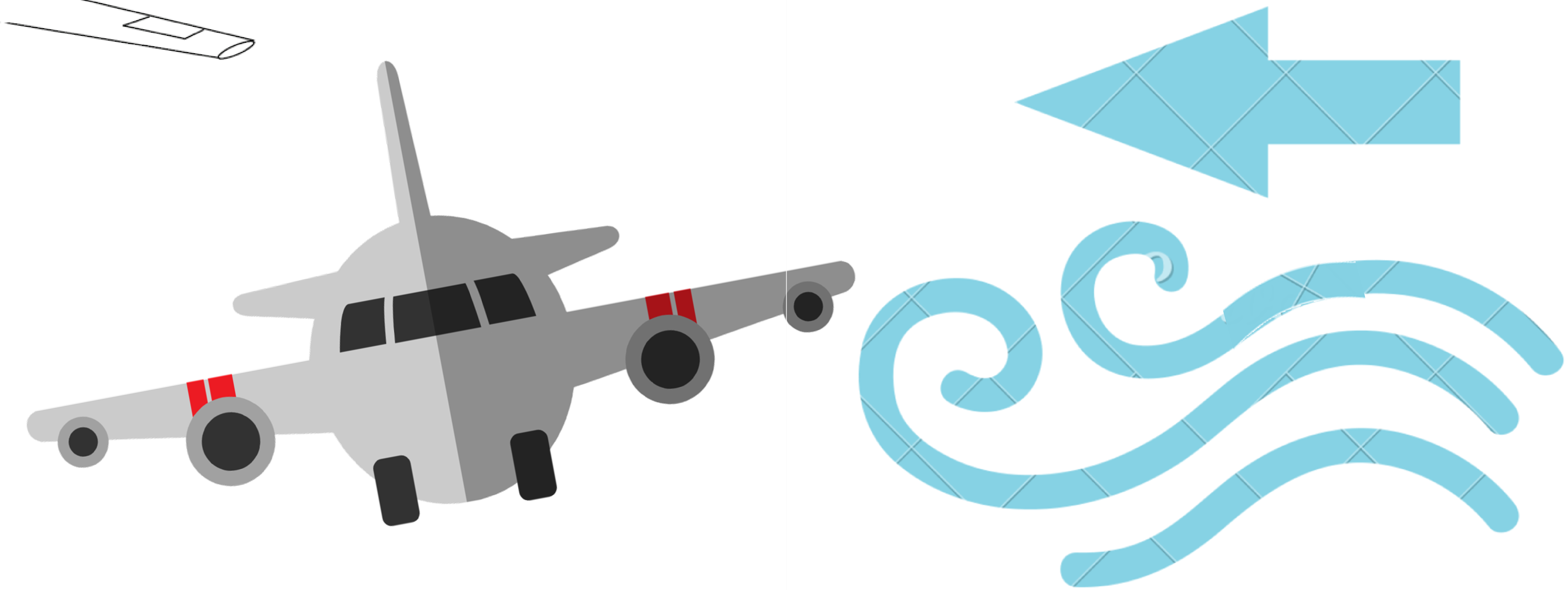


$$\tau = \frac{V}{q(t)} \qquad g(s) = \frac{e^{-\theta s}}{\tau s+1}$$

Grace, A., & Frawley, P. (2011). Experimental parametric equation for the prediction of valve coefficient (Cv) for choke valve trims. *International Journal of Pressure Vessels and Piping*, *88*(2–3), 109–118

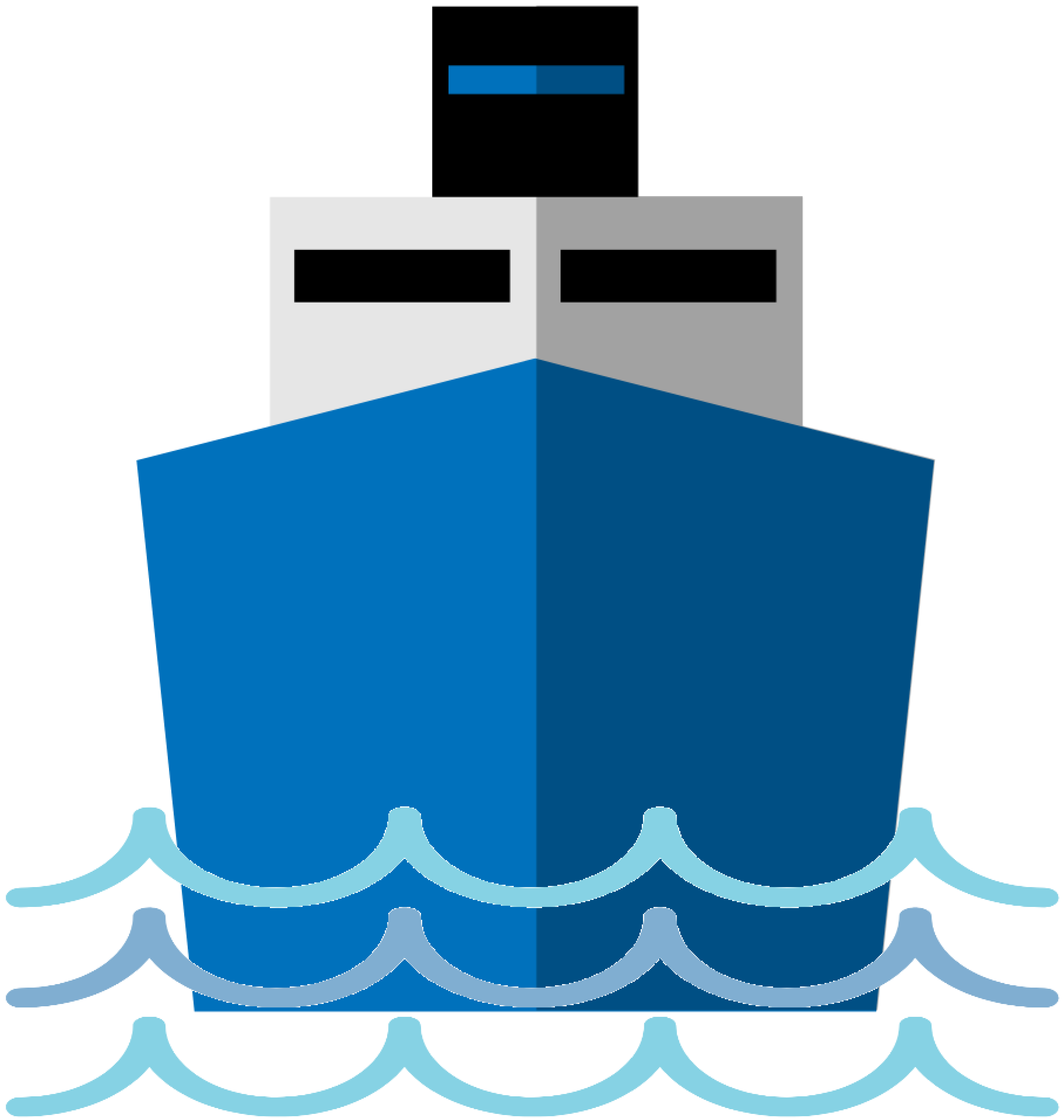
VARIATIONS IN PARAMETERS

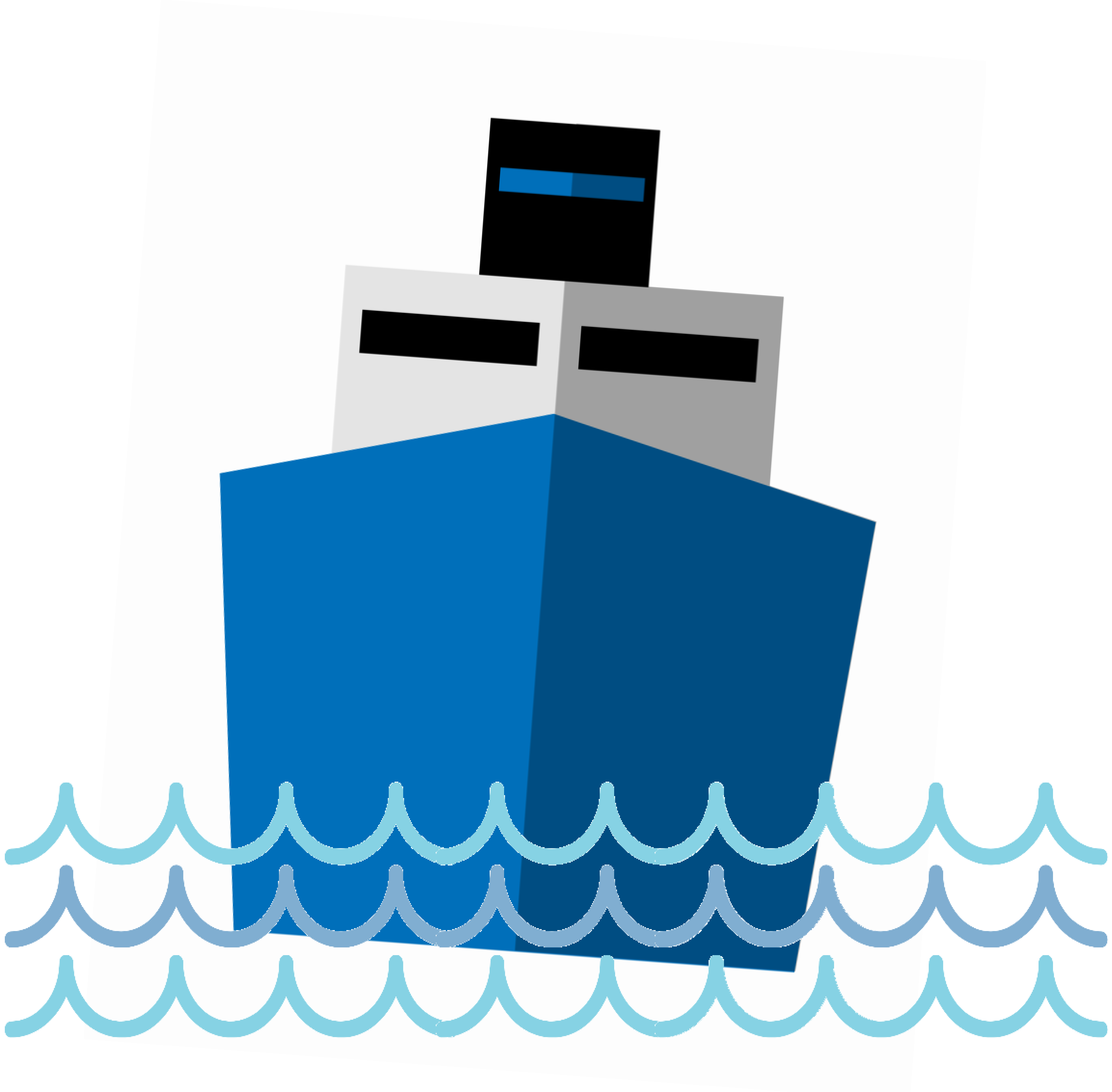




Åström, Karl J. and Wittenmark, Björn. (1995) Adaptive Control. Second Edition
Lavretsky, E. and Wise, K. (2013) Robust and adaptive control with aeropsace appllications

VARIATIONS IN PARAMETERS





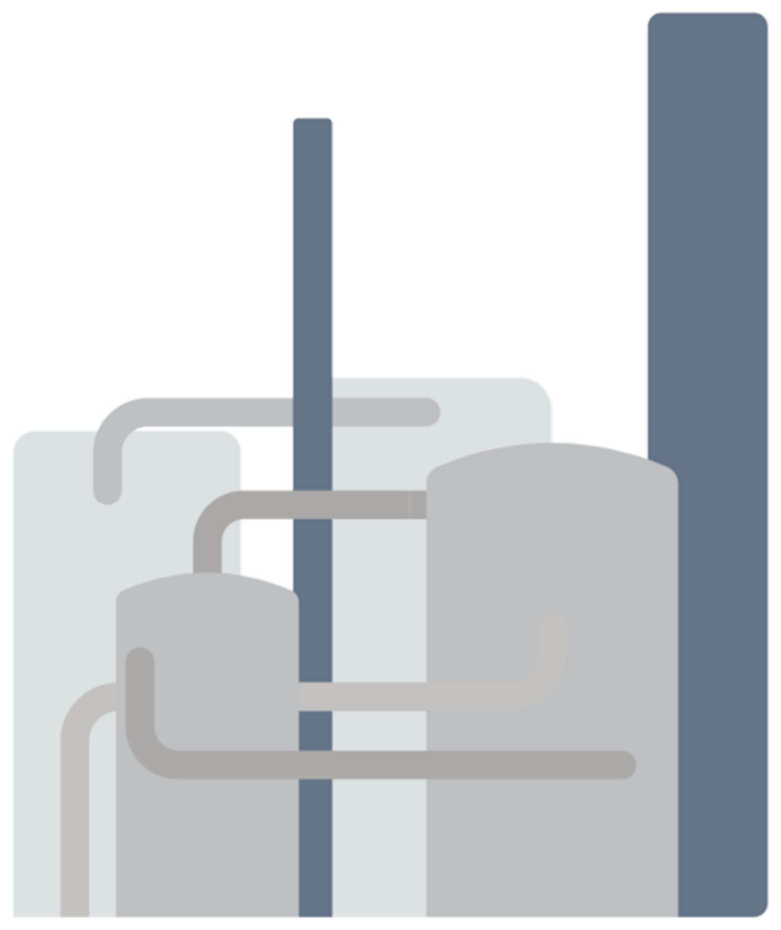
VARIATIONS IN PARAMETERS

Short-term changes:

Fast adaptation

Long term changes:

Slow adaptation



Åström, K. J., Hägglund, T., Hang, C., & Ho, W. K. (1992). Automatic Tuning and Adaptation for PID Controllers - A Survey. *IFAC Proceedings Volumes*, *25*(4).

ADAPTIVE CONTROLLERS

An adaptive controller is a controller with adjustable parameters and a mechanism for adjusting the parameters

Åström and Wittenmark (1995)

An adaptive controller is a combination of an online parameter estimator with a control law that is derived from the known parameter case

Petros and Sun (2012)

The parameters of an adaptive controller are continuously adjusted to accomodate changes in process dynamics and disturbances

Åström, K. J., Hägglund, T., Hang, C., & Ho, W. K. (1992)

Åström, K. J., Hägglund, T., Hang, C. c., & Ho, W. K. (1992). Automatic Tuning and Adaptation for PID Controllers - A Survey. *IFAC Proceedings Volumes*, *25*(4).
Åström, Karl J. and Wittenmark, Björn. Adaptive Control. Second Edition (1995)
Ioannou, Petros, A. and Sun, Jing. Robust Adaptive Control. (2012)

ADAPTIVE CONTROLLERS

An adaptive controller is a controller with adjustable parameters and a mechanism for adjusting the parameters

Åström and Wittenmark (1995)

An adaptive controller is a combination of an online parameter estimator with a **control law that is derived from the known parameter case**

Petros and Sun (2012)

The parameters of an adaptive controller are continuously adjusted to accomodate changes in process dynamics and disturbances

Åström, K. J., Hägglund, T., Hang, C., & Ho, W. K. (1992)

Åström, K. J., Hägglund, T., Hang, C. c., & Ho, W. K. (1992). Automatic Tuning and Adaptation for PID Controllers - A Survey. *IFAC Proceedings Volumes*, *25*(4).
Åström, Karl J. and Wittenmark, Björn. Adaptive Control. Second Edition (1995)
Ioannou, Petros, A. and Sun, Jing. Robust Adaptive Control. (2012)

ADAPTIVE CONTROLLERS

An adaptive controller is a controller with adjustable parameters and a mechanism for adjusting the parameters

Åström and Wittenmark (1995)

An adaptive controller is a combination of an online parameter estimator with a **control law that is derived from the known parameter case**

Petros and Sun (2012)

The parameters of an adaptive controller are continuously adjusted to accomodate **changes in process dynamics and disturbances**

Åström, K. J., Hägglund, T., Hang, C. , & Ho, W. K. (1992)

Åström, K. J., Hägglund, T., Hang, C. c., & Ho, W. K. (1992). Automatic Tuning and Adaptation for PID Controllers - A Survey. *IFAC Proceedings Volumes*, *25*(4).
Åström, Karl J. and Wittenmark, Björn. Adaptive Control. Second Edition (1995)
Ioannou, Petros, A. and Sun, Jing. Robust Adaptive Control. (2012)

ADAPTIVE CONTROLLERS

An adaptive controller is a controller with **adjustable parameters and a mechanism for adjusting the parameters**

Åström and Wittenmark (1995)

An adaptive controller is a combination of an online parameter estimator with a **control law that is derived from the known parameter case**

Petros and Sun (2012)

The parameters of an adaptive controller are continuously adjusted to accomodate **changes in process dynamics and disturbances**

Åström, K. J., Hägglund, T., Hang, C. , & Ho, W. K. (1992)

Åström, K. J., Hägglund, T., Hang, C. c., & Ho, W. K. (1992). Automatic Tuning and Adaptation for PID Controllers - A Survey. *IFAC Proceedings Volumes*, *25*(4).
Åström, Karl J. and Wittenmark, Björn. Adaptive Control. Second Edition (1995)
Ioannou, Petros, A. and Sun, Jing. Robust Adaptive Control. (2012)

ADAPTIVE CONTROLLERS

An adaptive controller is a controller with **adjustable parameters and a mechanism for adjusting the parameters**

Åström and Wittenmark (1995)

An adaptive controller is a combination of an **online parameter estimator** with a **control law that is derived from the known parameter case**

Petros and Sun (2012)

The parameters of an adaptive controller are continuously adjusted to accomodate **changes in process dynamics and disturbances**

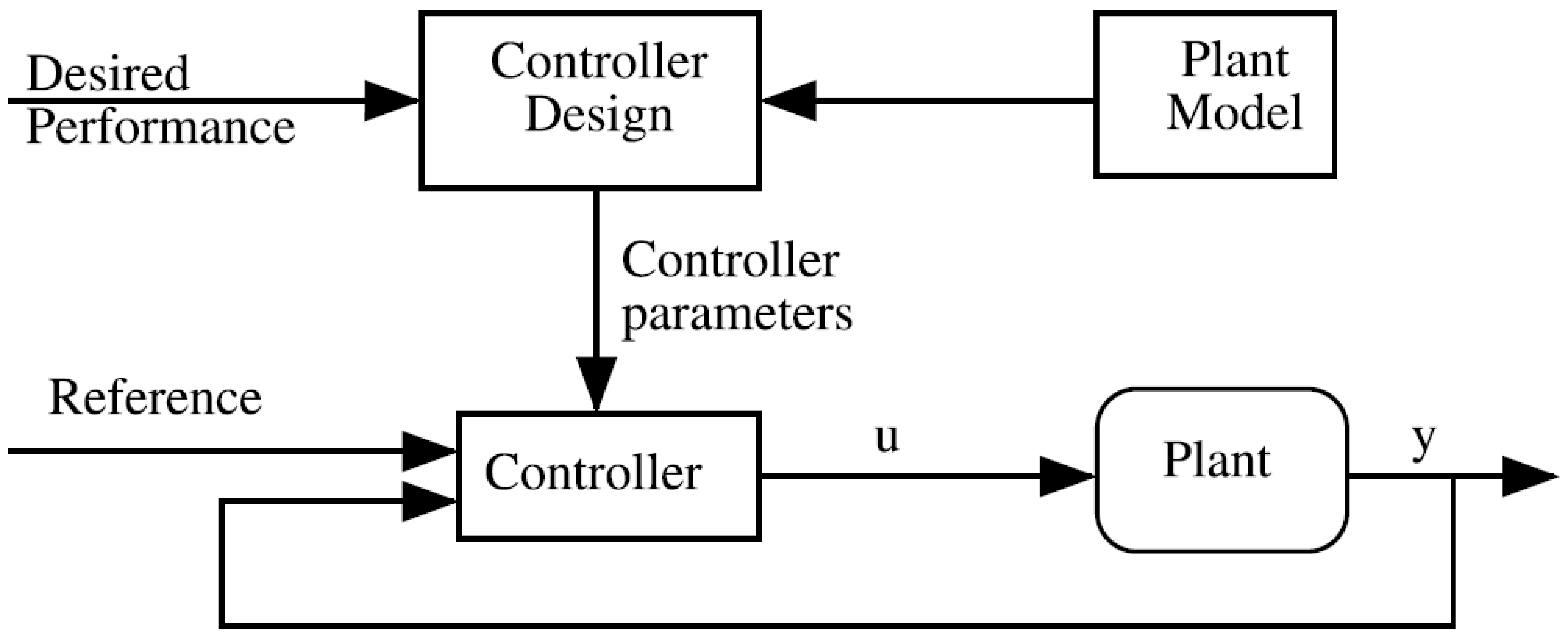
Åström, K. J., Hägglund, T., Hang, C. , & Ho, W. K. (1992)

Åström, K. J., Hägglund, T., Hang, C. c., & Ho, W. K. (1992). Automatic Tuning and Adaptation for PID Controllers - A Survey. *IFAC Proceedings Volumes*, *25*(4).
Åström, Karl J. and Wittenmark, Björn. Adaptive Control. Second Edition (1995)
Ioannou, Petros, A. and Sun, Jing. Robust Adaptive Control. (2012)

ADAPTIVE CONTROLLERS VS TYPICAL CONTROLLERS

DESIGN

Typical controller design

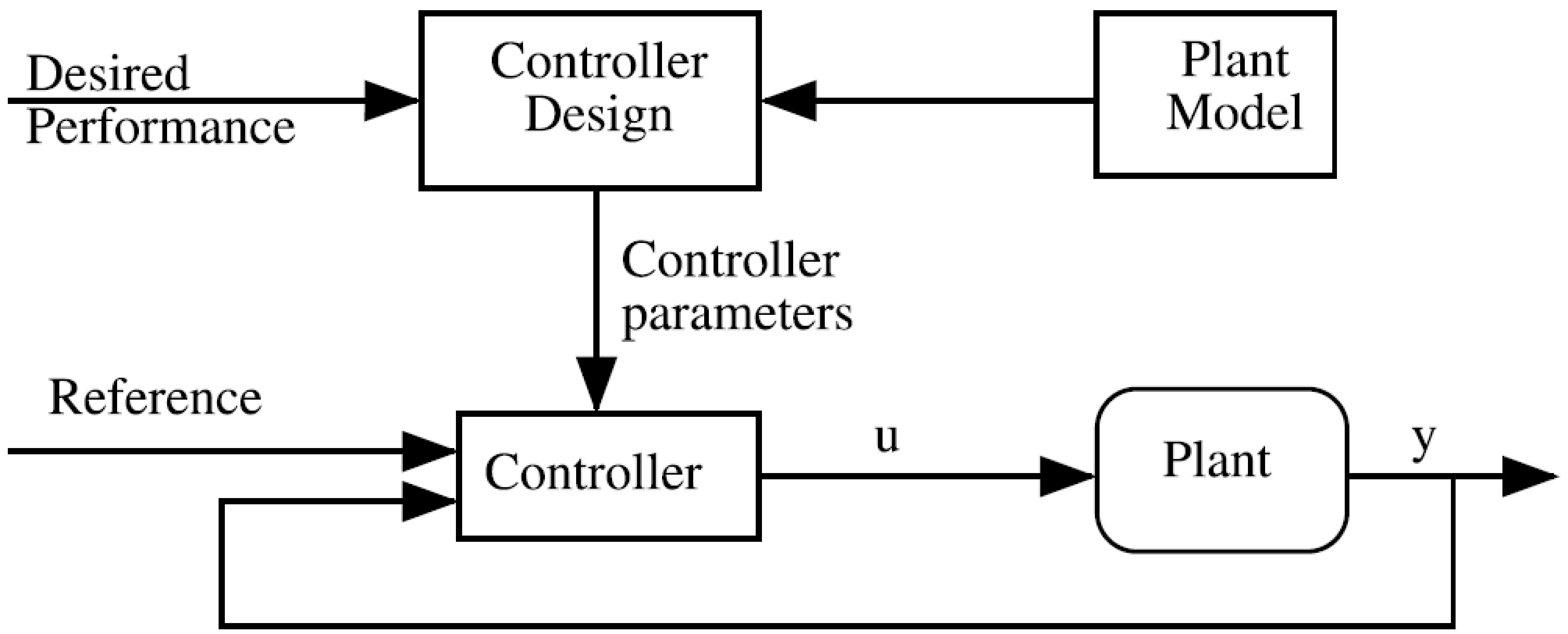


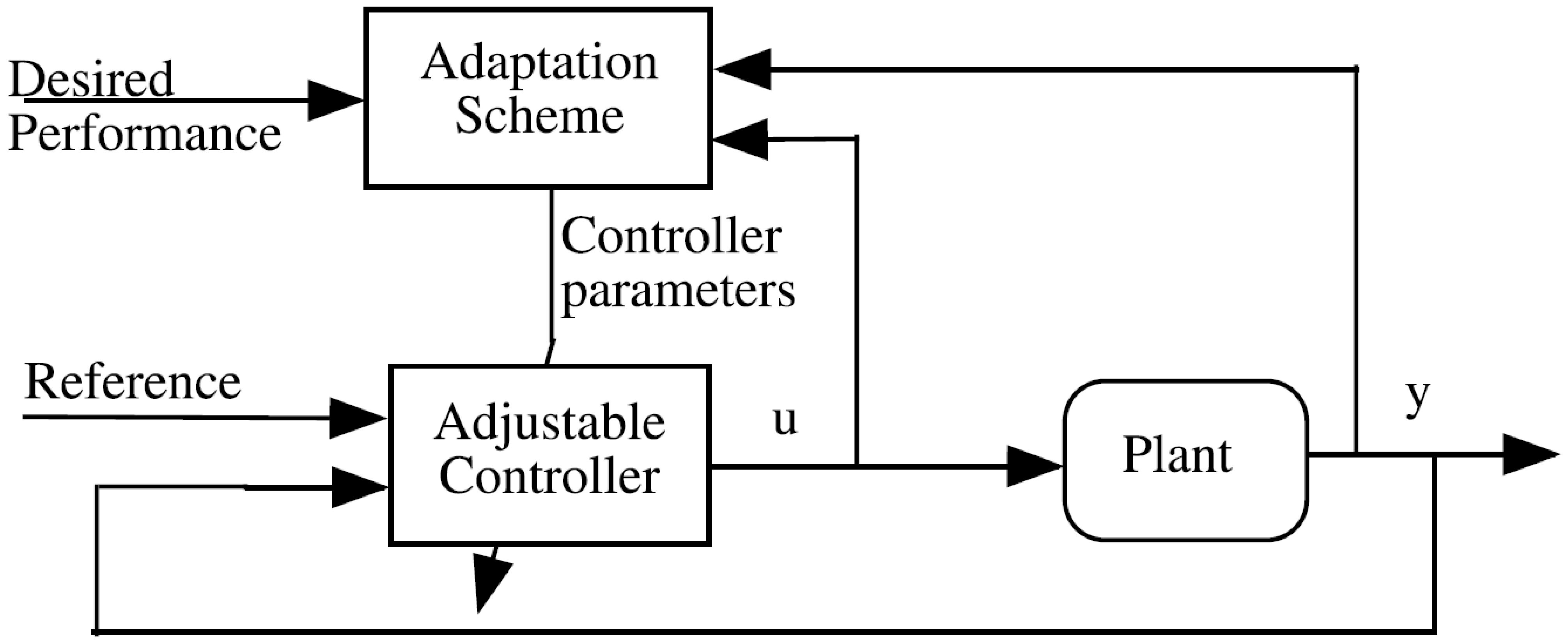
Landau, I.D. , Lozano, R. , M'Saad, M. and Karimi, A. (2011) Adaptive Control. Communications and Control Engineering. Springer

ADAPTIVE CONTROLLERS VS TYPICAL CONTROLLERS

DESIGN

Typical controller design





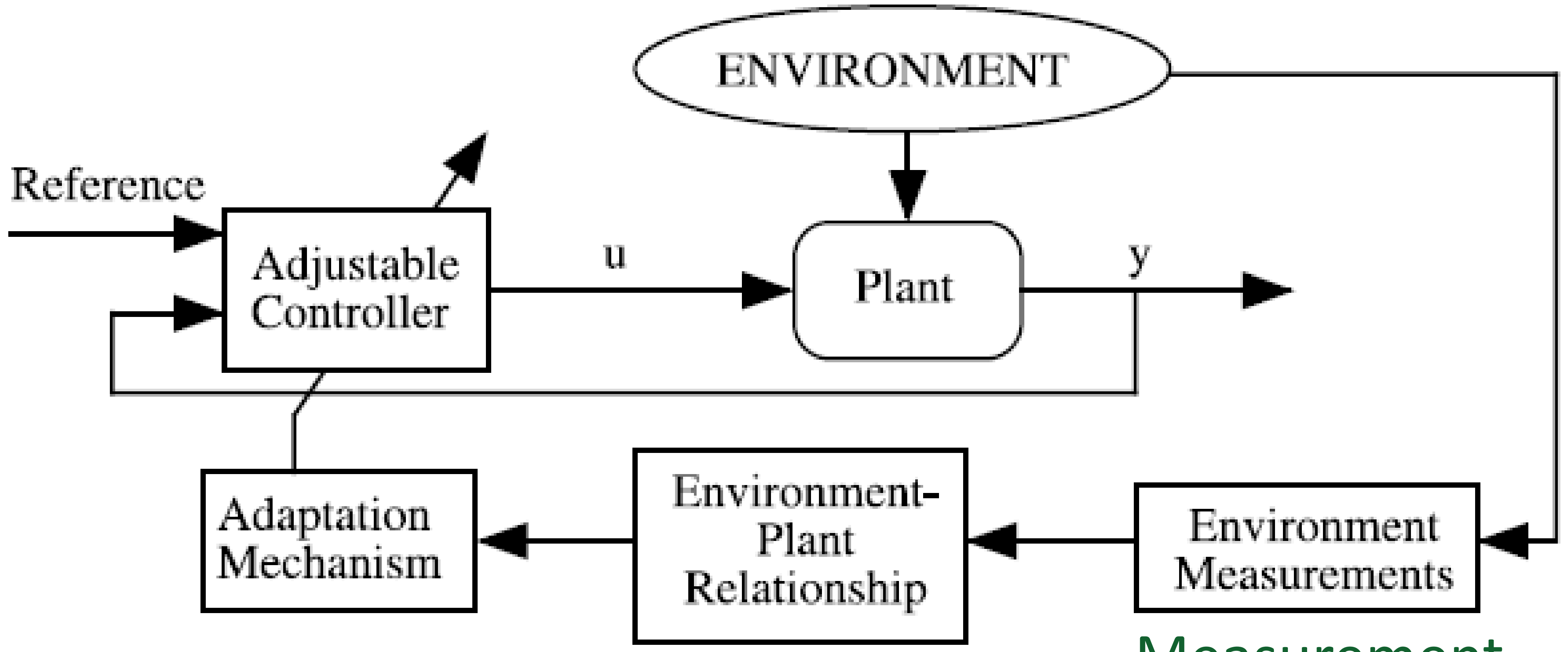
An adaptive control system

Landau, I.D. , Lozano, R. , M'Saad, M. and Karimi, A. (2011) Adaptive Control. Communications and Control Engineering. Springer

GAIN SCHEDULING: OPEN LOOP ADAPTATION

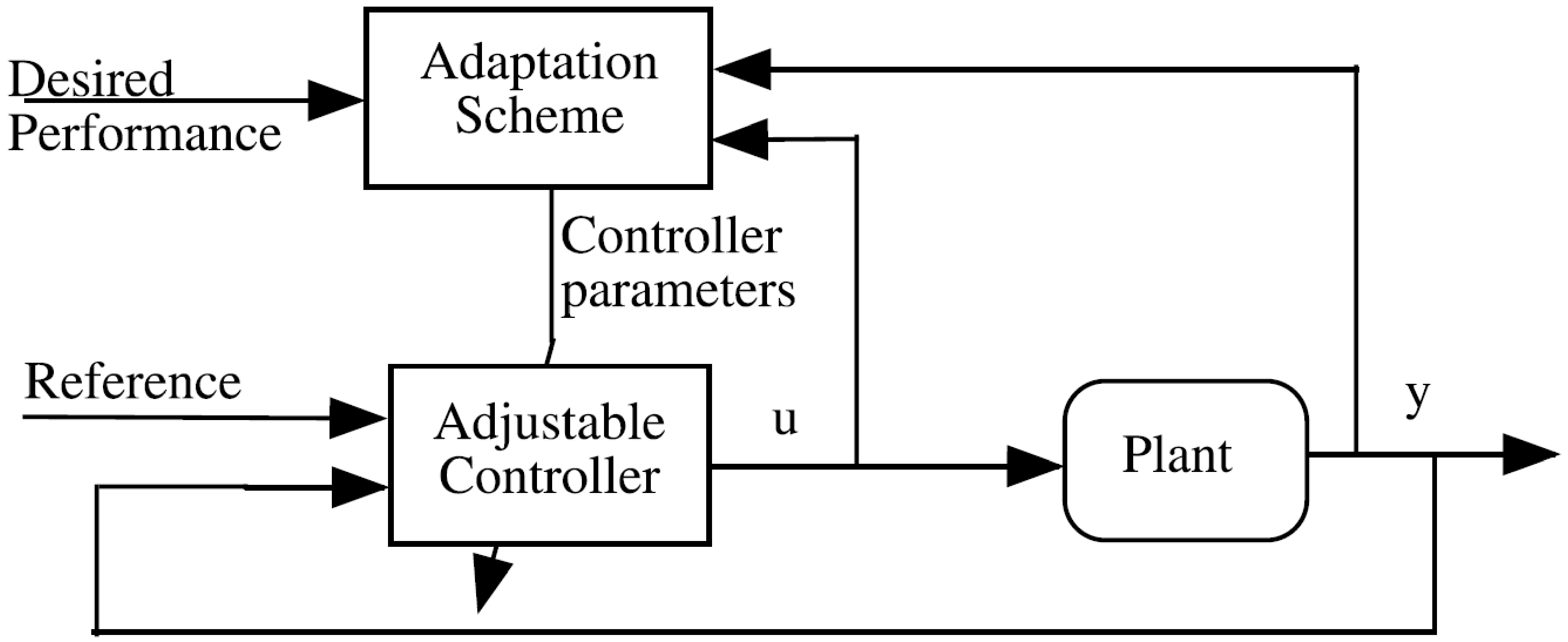
Gain scheduling:

- Linear controller
- Parameters are changed as a function of operating conditions in a pre-programmed way.



Feedforward compensation

Measurement, controller output, external signal



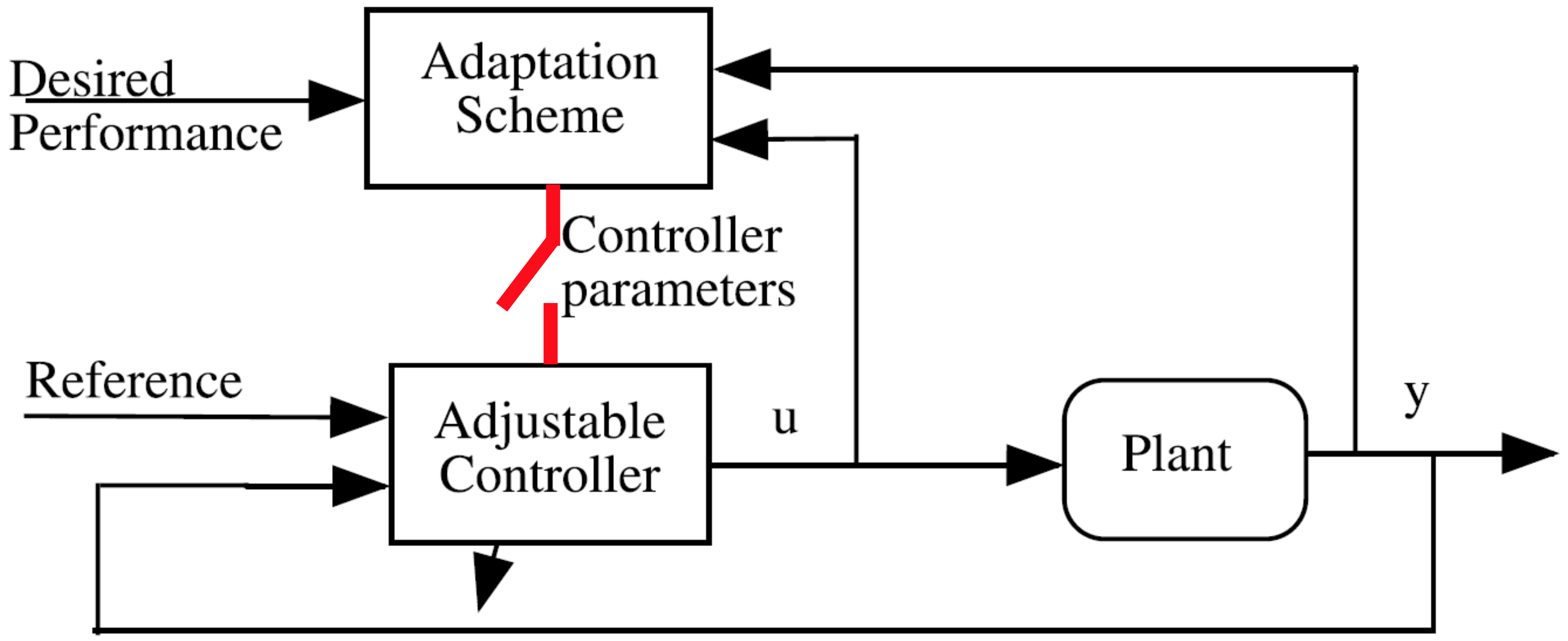
An adaptive control system

Landau, I.D. , Lozano, R. , M'Saad, M. and Karimi, A. (2011) Adaptive Control. Communications and Control Engineering. Springer
Åström, Karl J. and Wittenmark, Björn. (1995) Adaptive Control. Second Edition
Åström, K. J., Hägglund, T., Hang, C. , & Ho, W. K. (1992). Automatic Tuning and Adaptation for PID Controllers - A Survey. *IFAC Proceedings Volumes*, *25*(4).

AUTO-TUNING: "ONE SHOT" ADAPTATION

Auto-tuning

Controller parameters are tuned automatically on demand from an operator or external signal



Adaptation

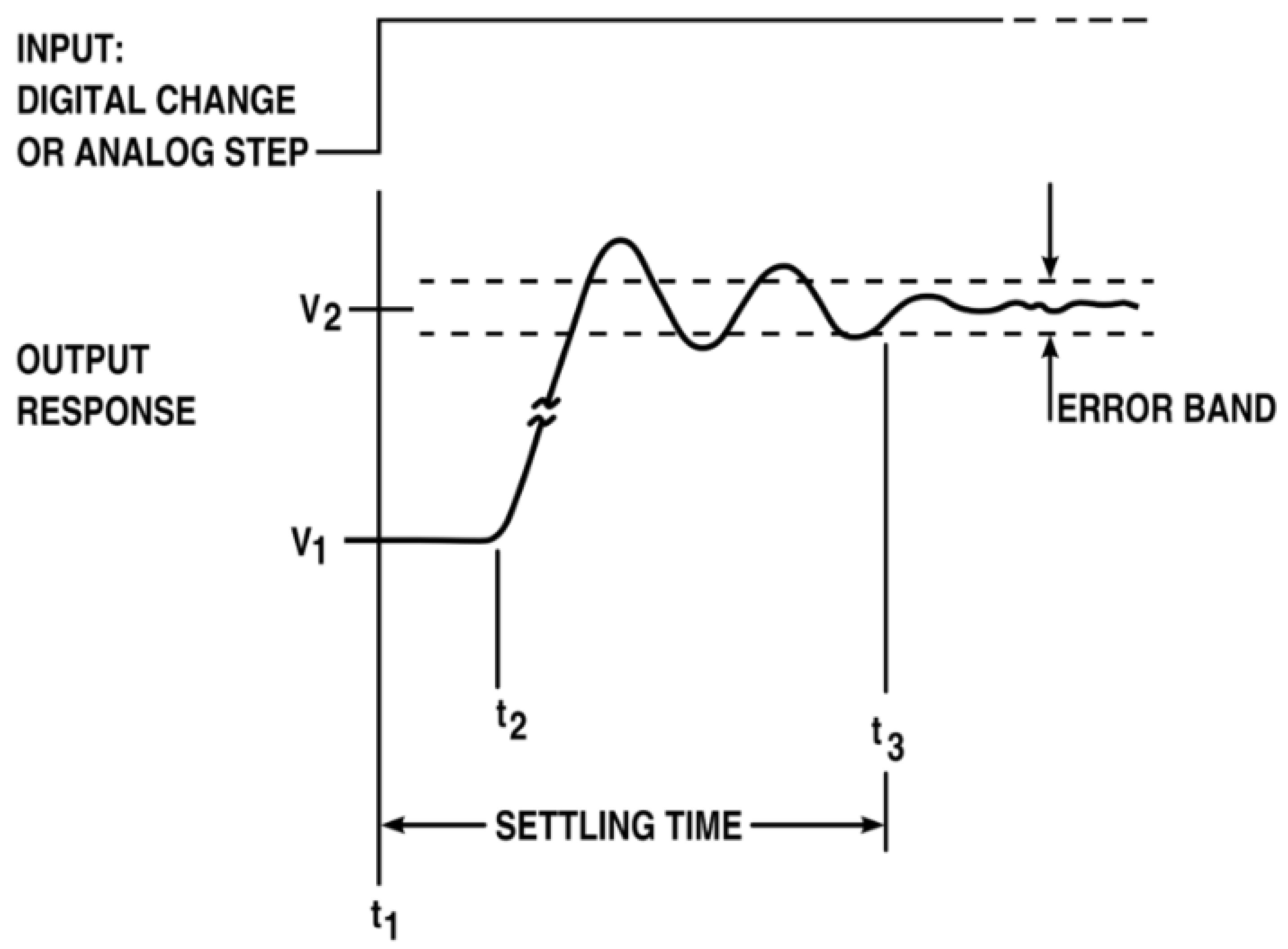
The parameters of a controller are continuously updated

Åström, Karl J. and Hägglund, Tore. Automatic tuning of PID controllers (1988)
Åström, K. J., Hägglund, T., Hang, C. c., & Ho, W. K. (1992). Automatic Tuning and Adaptation for PID Controllers - A Survey. *IFAC Proceedings Volumes*, *25*(4).

IDENTIFICATION

Open loop

Step or pulse

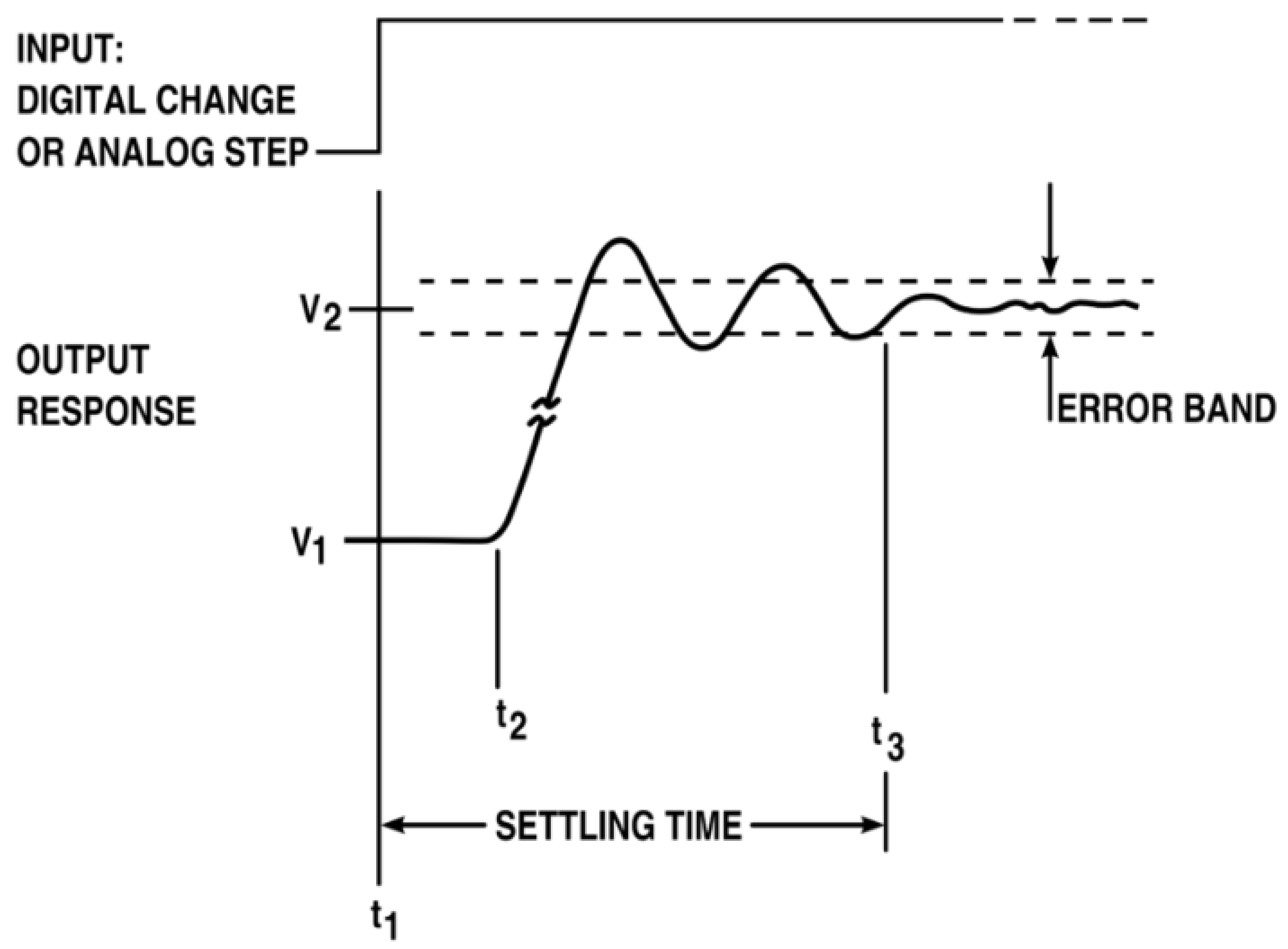


Åström, Karl J. and Hägglund, Tore. Automatic tuning of PID controllers (1988)
Åström, K. J., Hägglund, T., Hang, C. c., & Ho, W. K. (1992). Automatic Tuning and Adaptation for PID Controllers - A Survey. *IFAC Proceedings Volumes*, *25*(4).

IDENTIFICATION

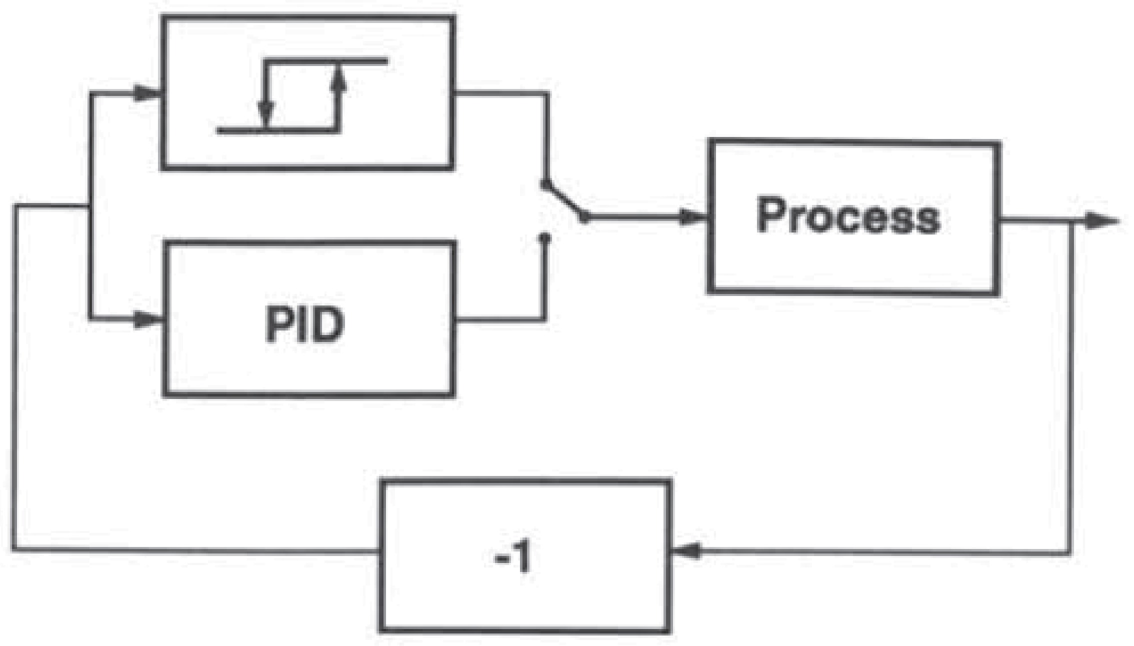
Open loop

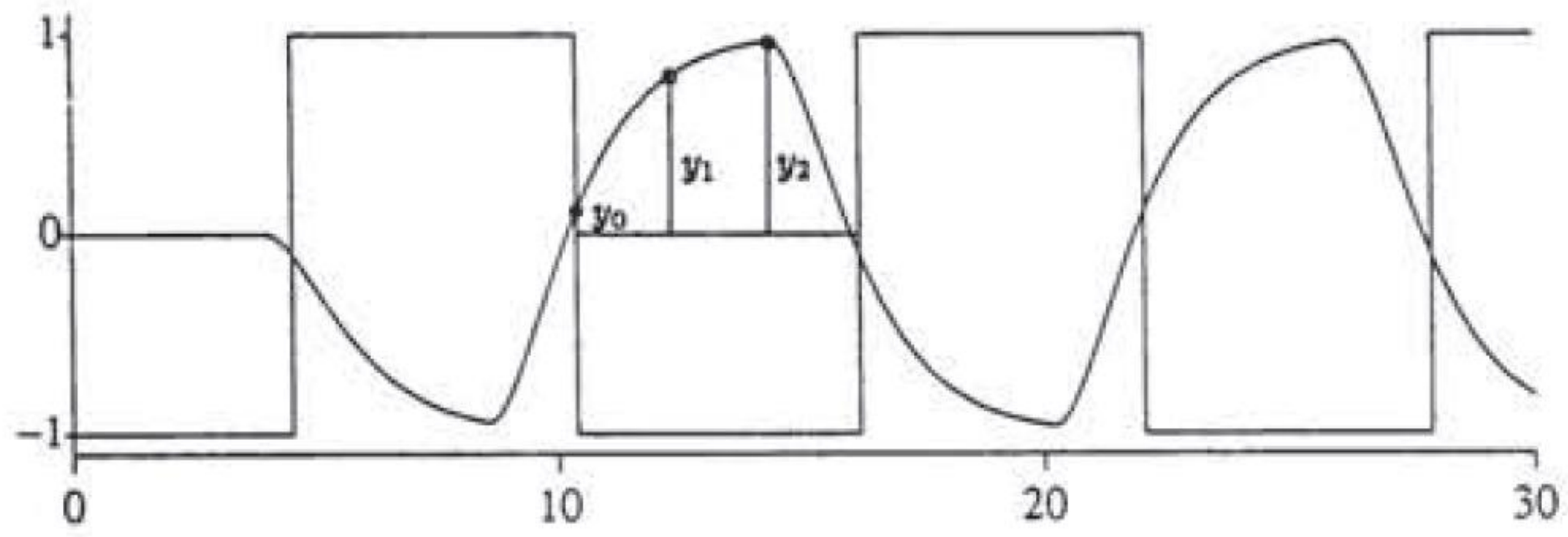
Step or pulse



Closed loop (online)

Known disturbance (e.g. relay feedback) to get frequency response information used to tune.



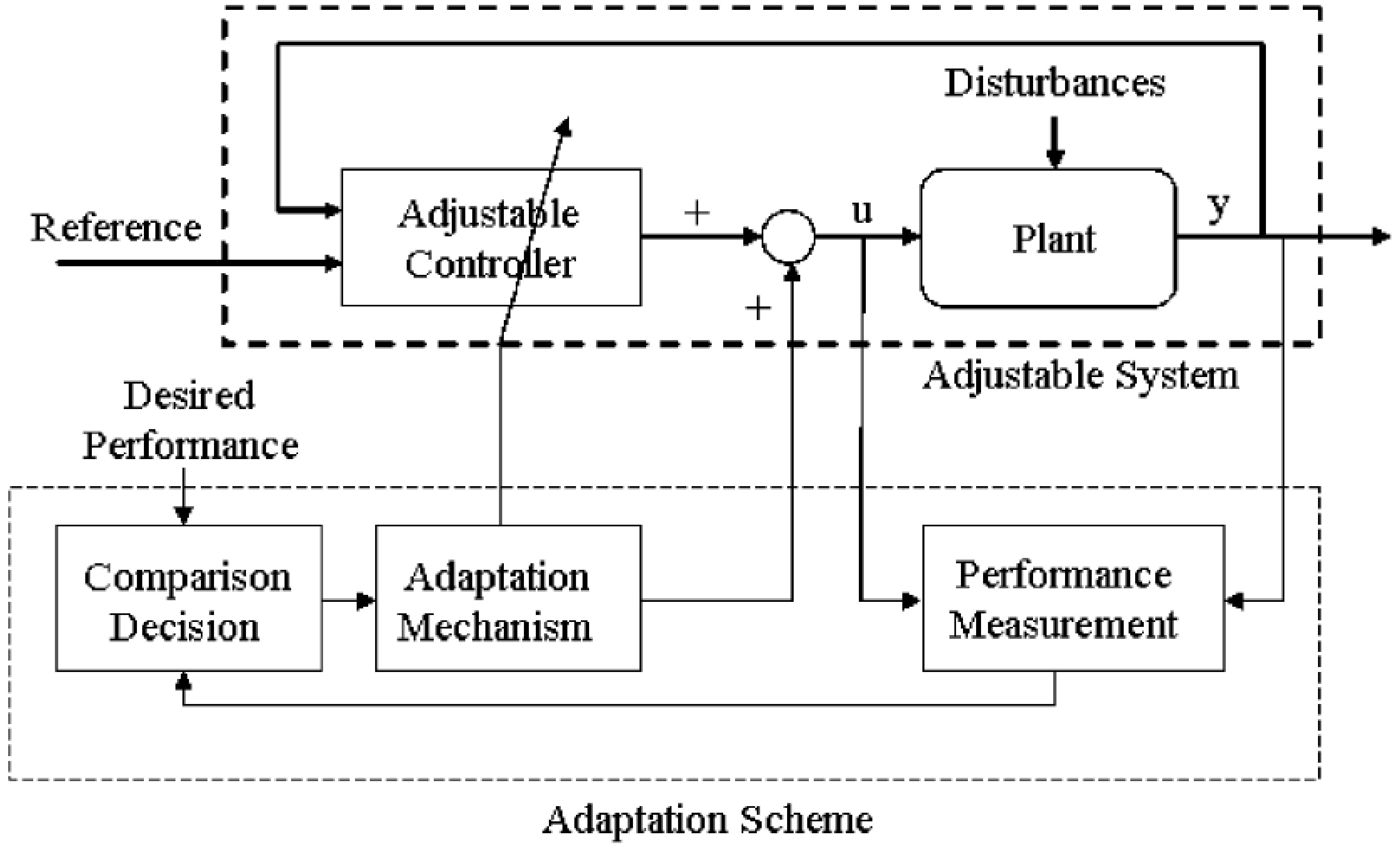


Åström, Karl J. and Hägglund, Tore. Automatic tuning of PID controllers (1988)
Åström, K. J., Hägglund, T., Hang, C. c., & Ho, W. K. (1992). Automatic Tuning and Adaptation for PID Controllers - A Survey. *IFAC Proceedings Volumes*, *25*(4).

ADAPTIVE CONTROLLERS

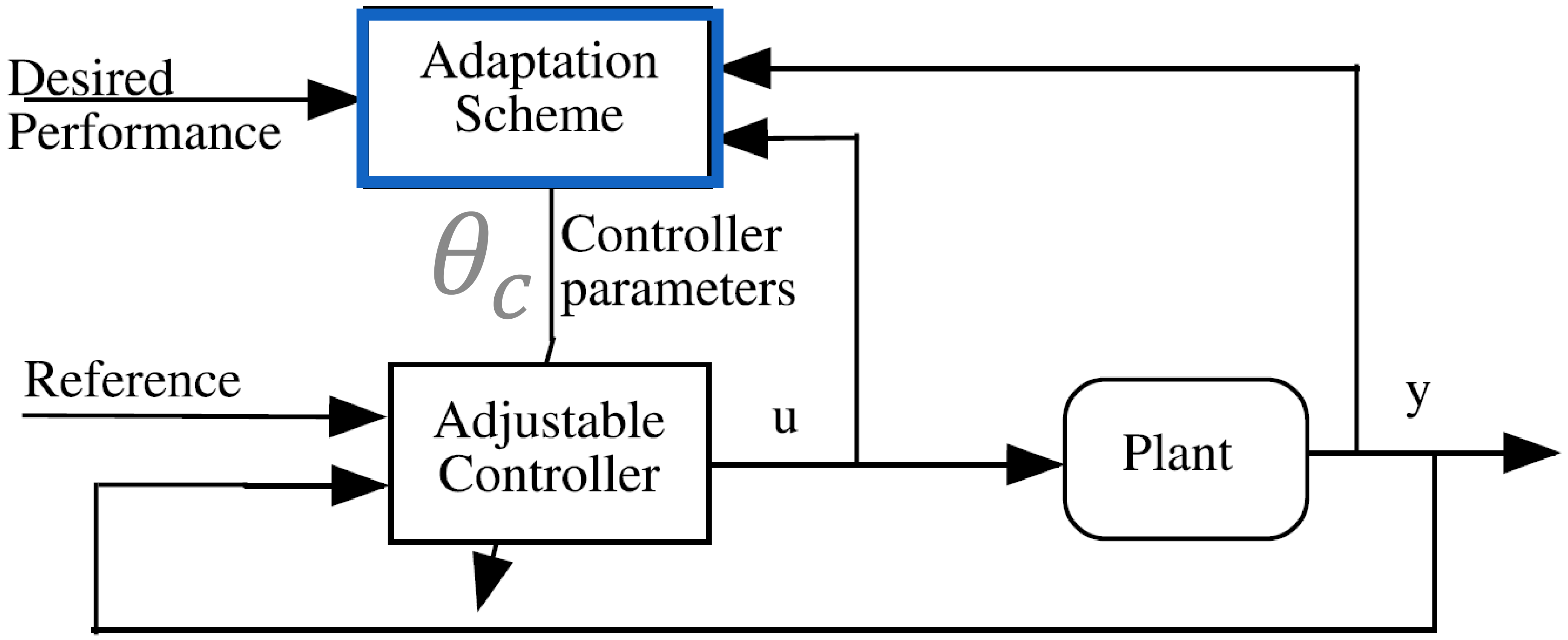
SOME COMMON CHARACTERISTICS

1. Controller with fixed structure and complexity
2. *A priori* information about structure of plant model
3. Specified performances can be achieved with appropriate values of controller parameters
4. Closed loop control of a certain performance index



Landau, I.D. , Lozano, R. , M'Saad, M. and Karimi, A. (2011) Adaptive Control. Communications and Control Engineering. Springer

ADAPTIVE CONTROLLERS



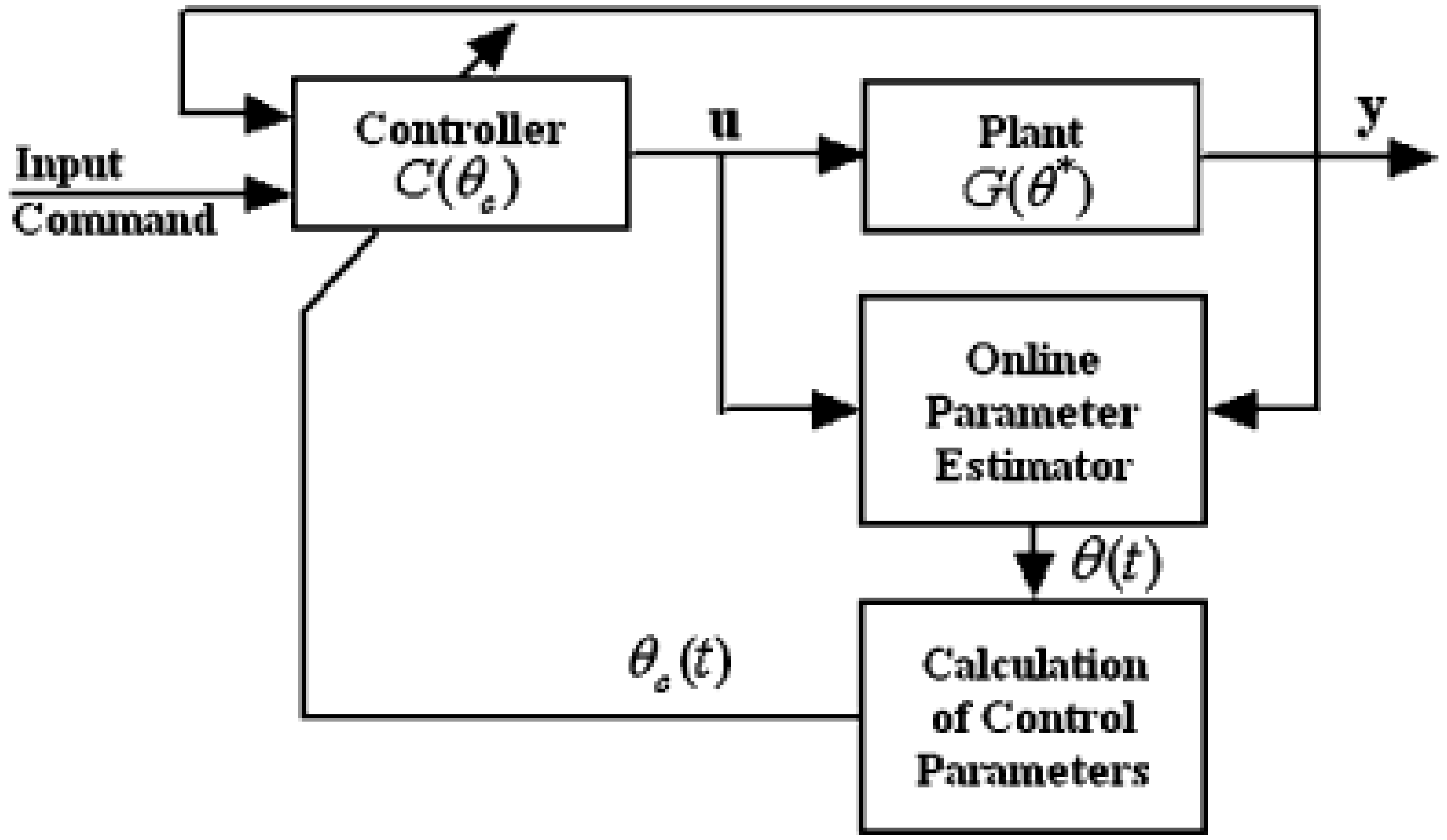
- Adaptation scheme
- Parmeter estimator
- Adaptive law
- Update law
- Adjustment mechanism

Landau, I.D. , Lozano, R. , M'Saad, M. and Karimi, A. (2011) Adaptive Control. Communications and Control Engineering. Springer
Åström, Karl J. and Wittenmark, Björn. Adaptive Control. Second Edition (1995)
Ioannou, Petros, A. and Sun, Jing. Robust Adaptive Control. (2012)
Anderson, B. (2005). Failures of adaptive control theory and their resolution. *Communications in Information and Systems*, *5*(1), 1–20.

ADAPTIVE CONTROLLERS

DIRECT AND INDIRECT IMPLEMENTATIONS

INDIRECT or EXPLICIT



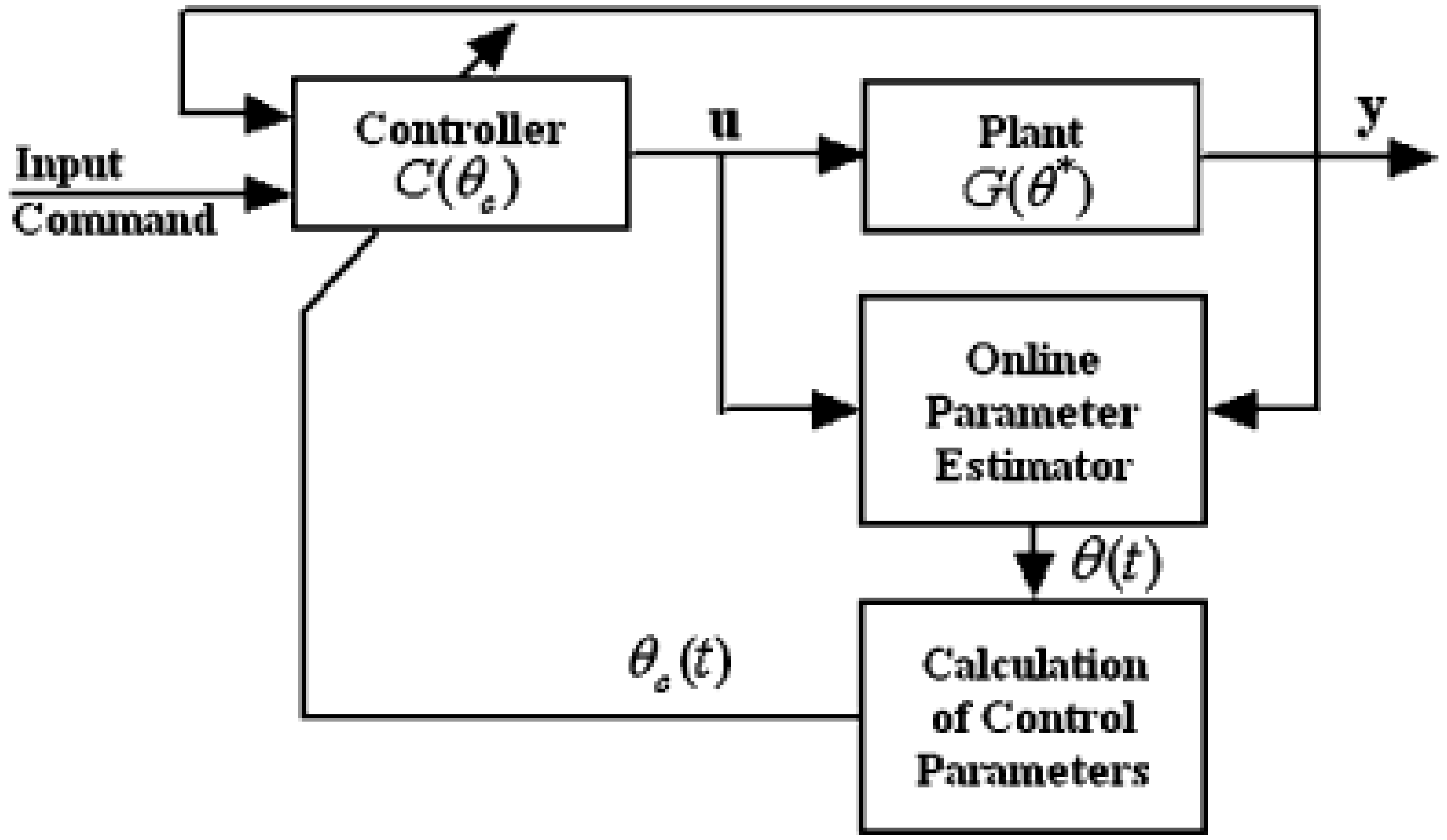
Performance specified in terms of the desired plant model

Åström, Karl J. and Wittenmark, Björn. (1995) Adaptive Control. Second Edition
Landau, I.D. , Lozano, R. , M'Saad, M. and Karimi, A. (2011) Adaptive Control. Communications and Control Engineering. Springer
Mahmoud, M, Xia, Y. (2012) Applied Control Sytems Design

ADAPTIVE CONTROLLERS

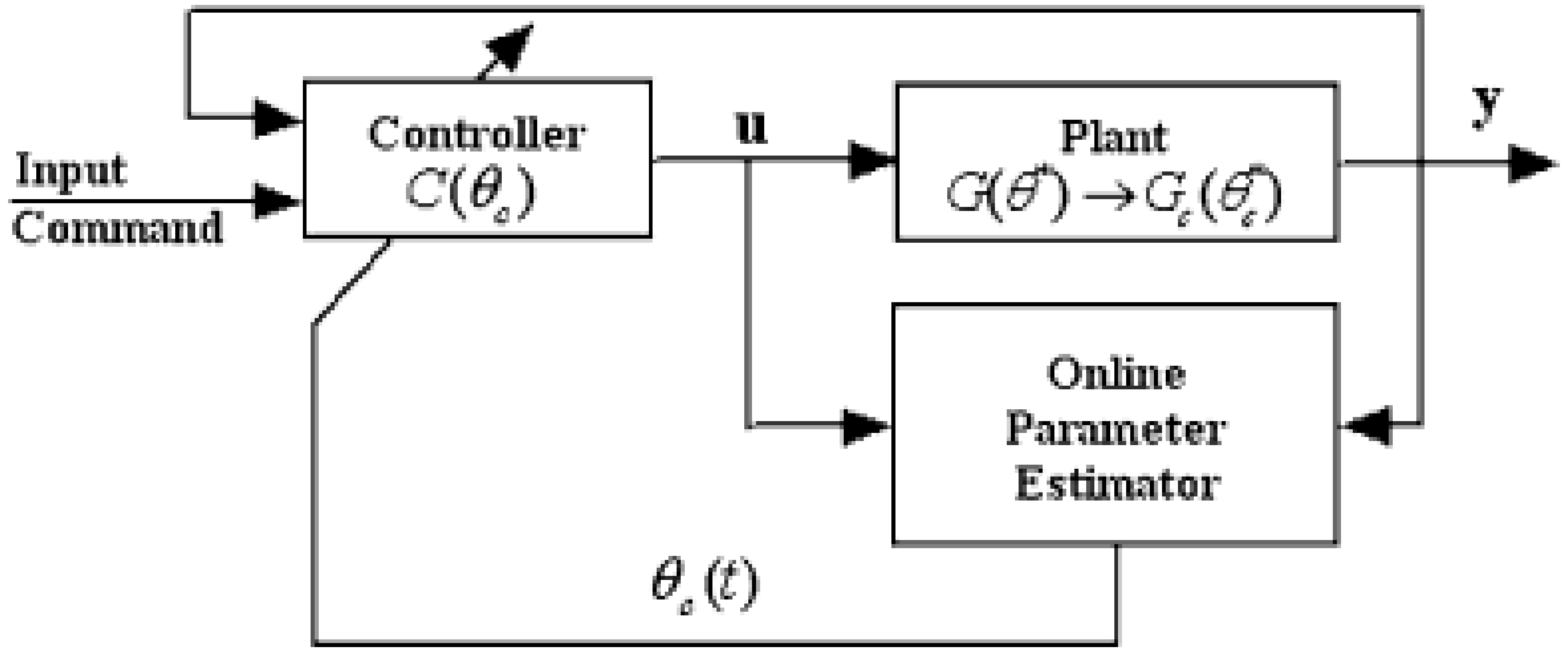
DIRECT AND INDIRECT IMPLEMENTATIONS

INDIRECT or EXPLICIT



Performance specified in terms of the desired plant model

DIRECT or IMPLICIT



Performance specified in terms of realizing the desired behaviour of the closed loop system

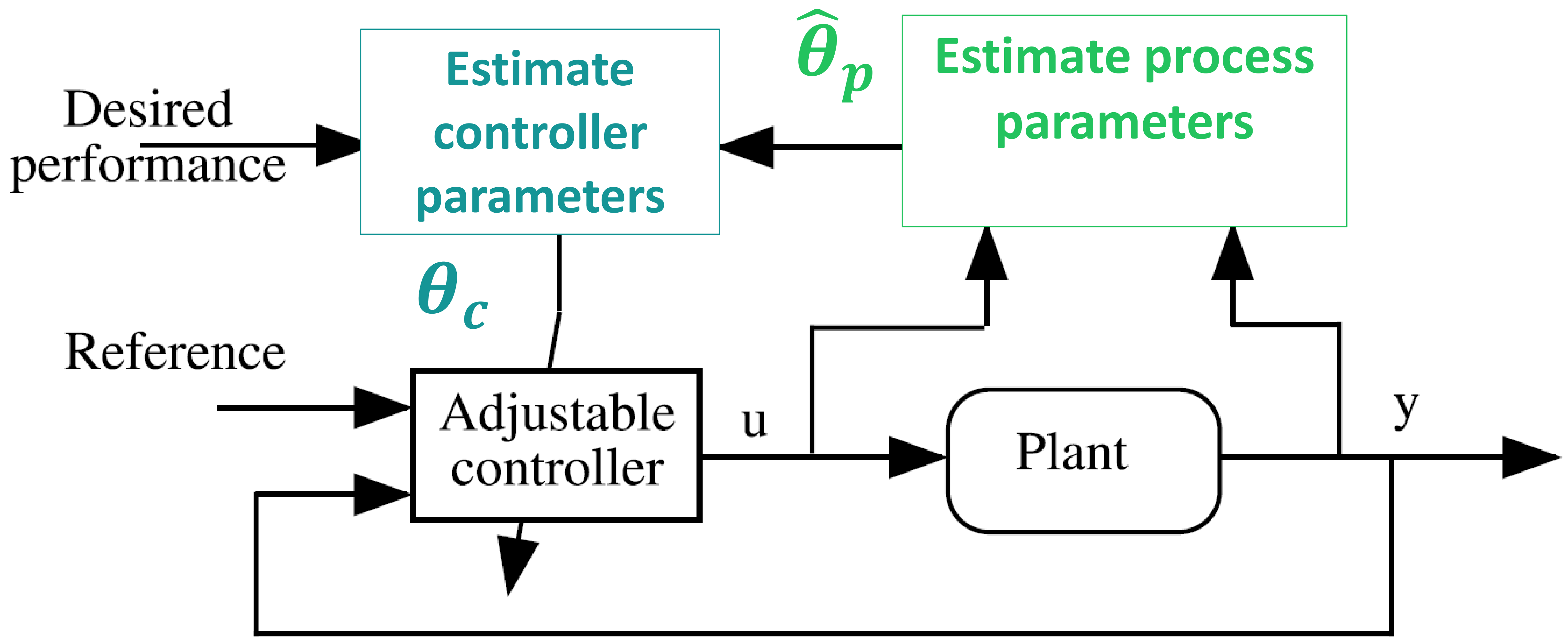
Åström, Karl J. and Wittenmark, Björn. (1995) Adaptive Control. Second Edition
Landau, I.D. , Lozano, R. , M'Saad, M. and Karimi, A. (2011) Adaptive Control. Communications and Control Engineering. Springer
Mahmoud, M, Xia, Y. (2012) Applied Control Sytems Design

SELF-TUNING REGULATORS

Kalman (1958): self tuning controller:
"optimal LQR with explicit identification of parameters"

- Controller parameters converge to the controller that was designed if the process was known.

- Estimates of parameter uncertainties not used in control design.
- **Certainty equivalence principle:** estimated parameters treated as if they were true in designing the controller; additive disturbances.

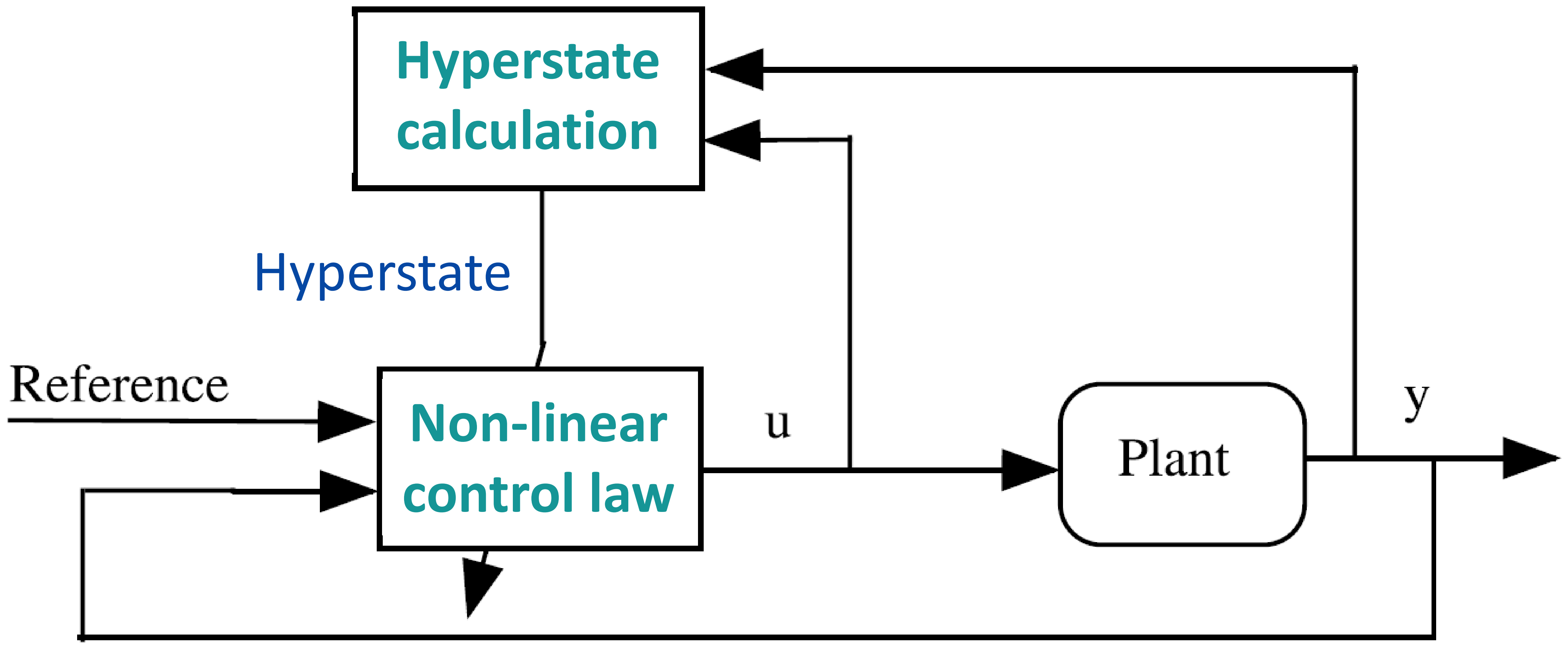


Åström, Karl J. and Wittenmark, Björn. (1995) Adaptive Control. Second Edition
Mahmoud, M, Xia, Y. (2012) Applied Control Sytems Design

STOCHASTIC SELF-TUNING REGULATORS

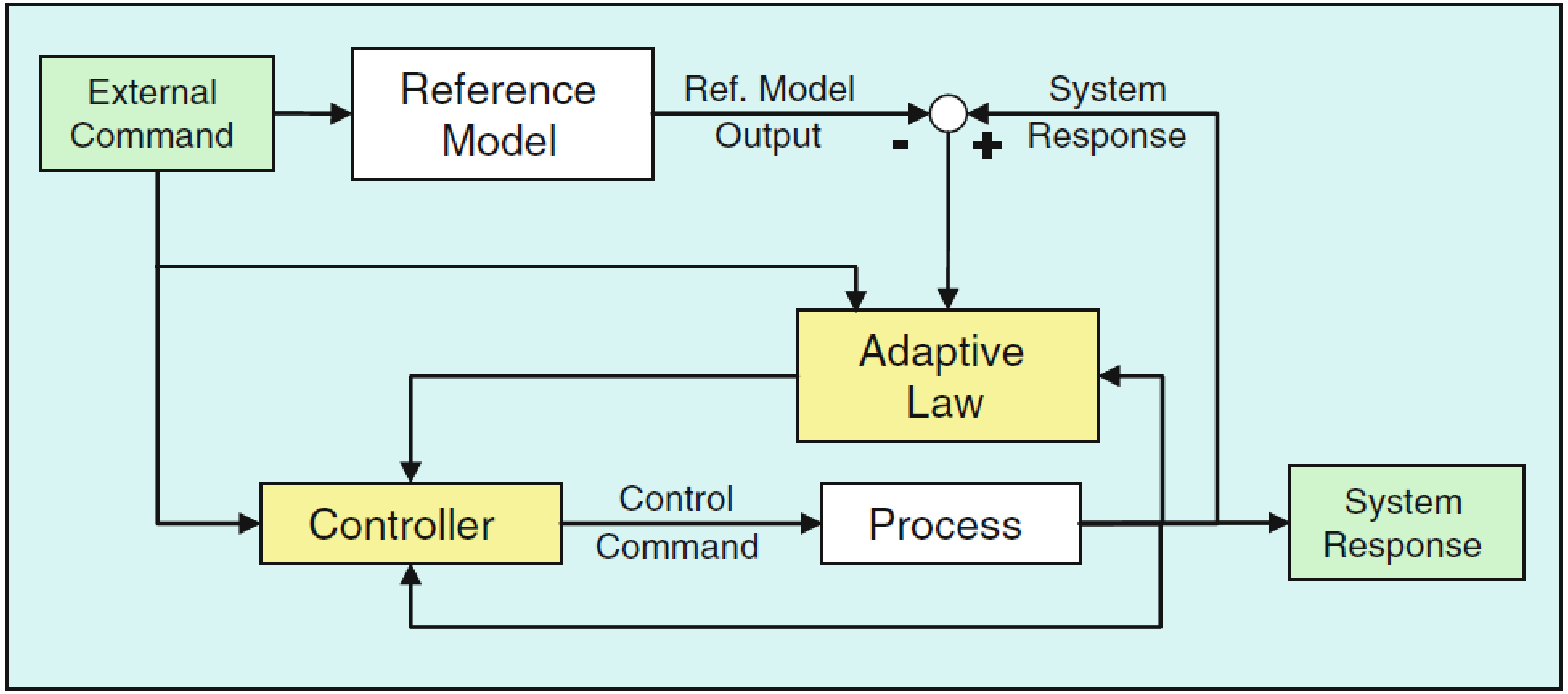
DUAL CONRTROL

- When the input starts decreasing (less excitation) less information is gained about the process and the parameter uncertainties increase.
- Control law as function of **parameter estimates** and the **uncertainties of estimates**.
- The control attempts to drive the output to the desired value but also may introduce perturbations whtn estimates are uncertain → dual control (active learning)



Åström, Karl J. and Wittenmark, Björn. (1995) Adaptive Control. Second Edition
Landau, I.D. , Lozano, R. , M'Saad, M. and Karimi, A. (2011) Adaptive Control. Communications and Control Engineering. Springer
Mahmoud, M, Xia, Y. (2012) Applied Control Sytems Design

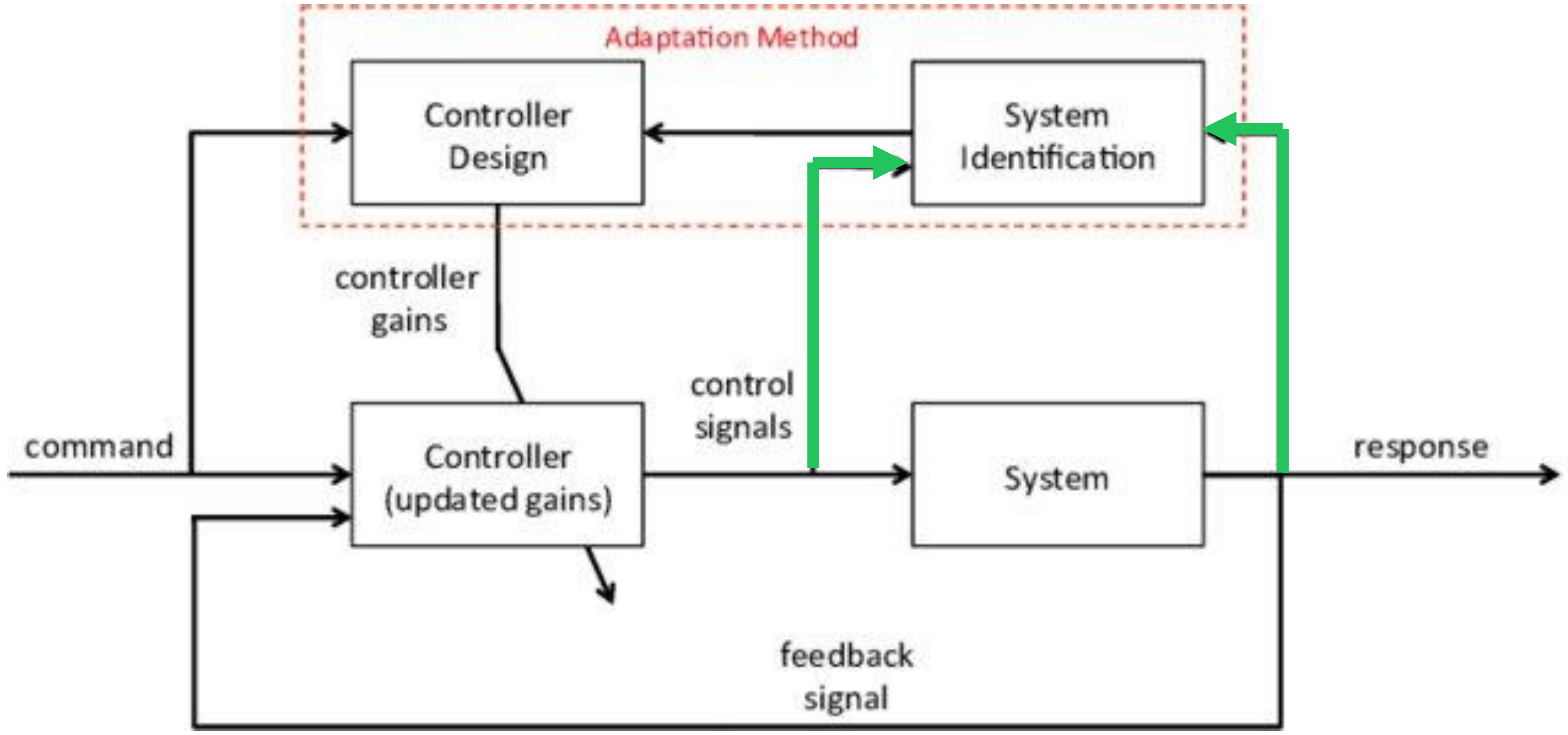
MRAC:
MODEL REFERENCE ADAPTIVE CONTROLLERS

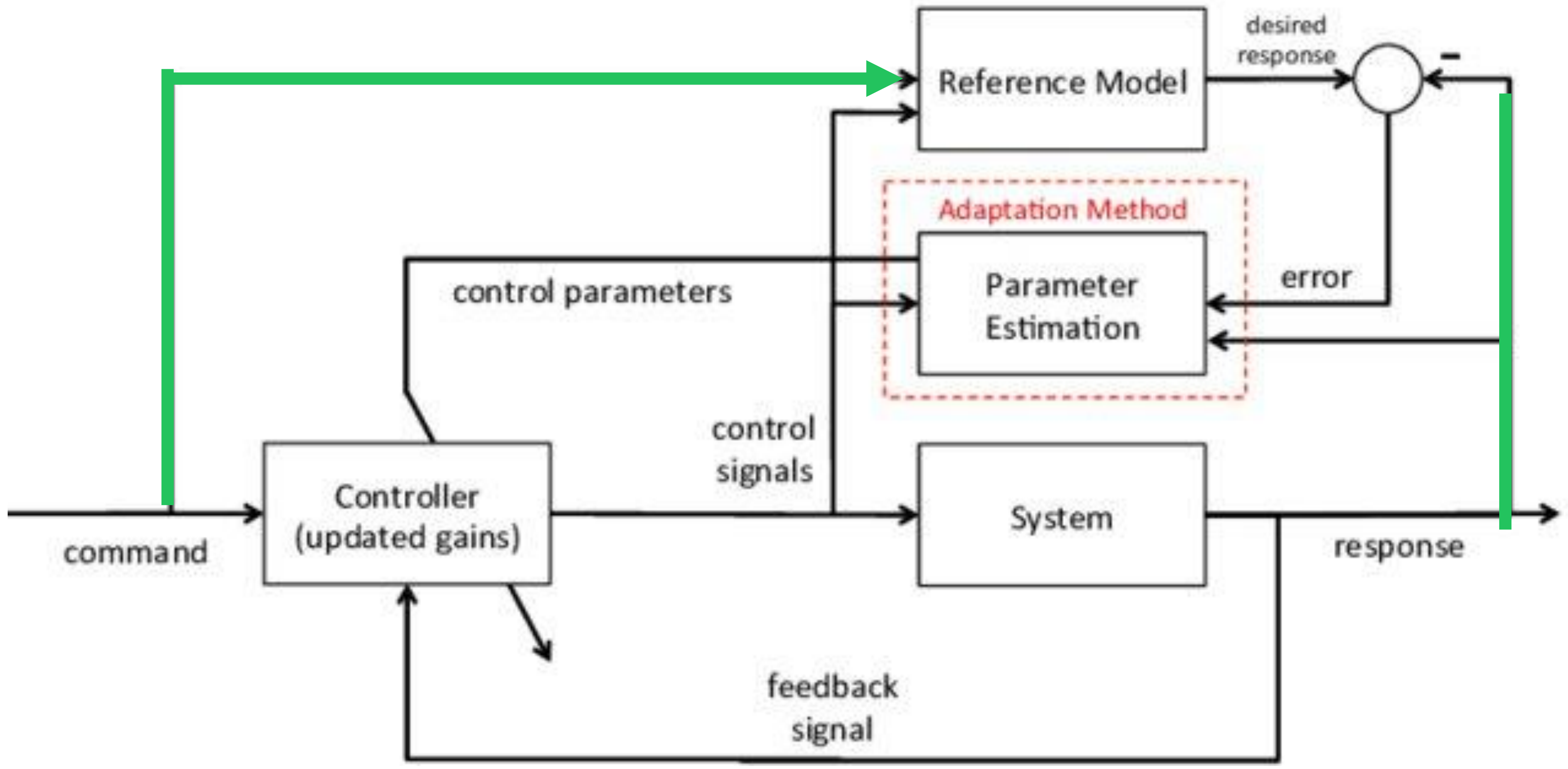


Åström, Karl J. and Wittenmark, Björn. (1995) Adaptive Control. Second Edition
Landau, I.D. , Lozano, R. , M'Saad, M. and Karimi, A. (2011) Adaptive Control. Communications and Control Engineering. Springer
Lavretsky, E. and Wise, K. (2013) Robust and adaptive control with aeropsace appllications

MRAC:
MODEL REFERENCE ADAPTIVE CONTROLLERS

DIRECT AND INDIRECT





INDIRECT | **DIRECT**

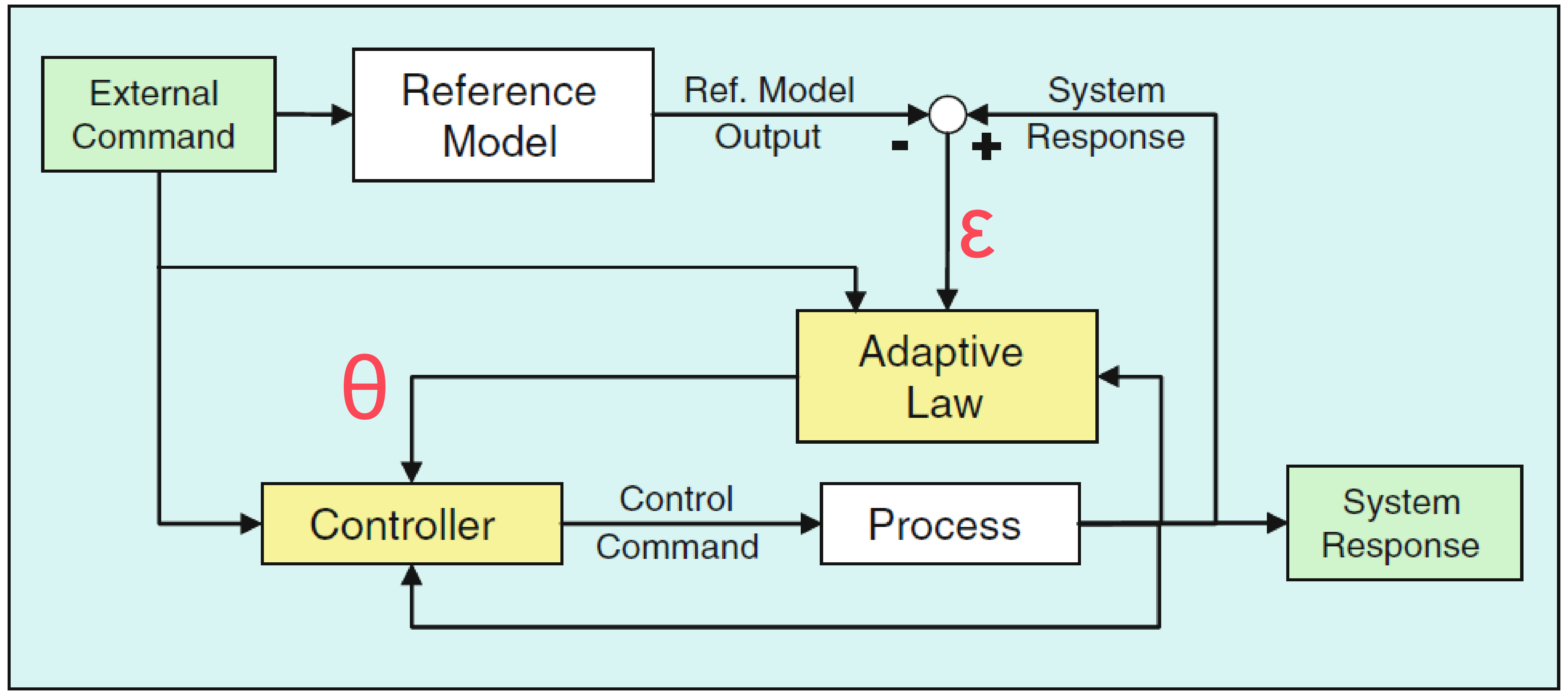
Lavretsky, E. and Wise, K. (2013) Robust and adaptive control with aeropsace appllications
Bhattacharyya, S., Cofer, D., Musliner, D., Mueller, J., & Engstrom, E. (2015). Certification considerations for adaptive systems. In *2015 International Conference on Unmanned Aircraft Systems (ICUAS)* (pp. 270–279). IEEE.

MRAC:
MODEL REFERENCE ADAPTIVE CONTROLLERS

GRADIENT METHOD FOR ADAPTIVE LAW

- Minimize ε^2

$$\frac{d\theta}{dt} = \gamma\varphi\varepsilon \qquad \varphi = -\frac{d\varepsilon}{d\theta}$$



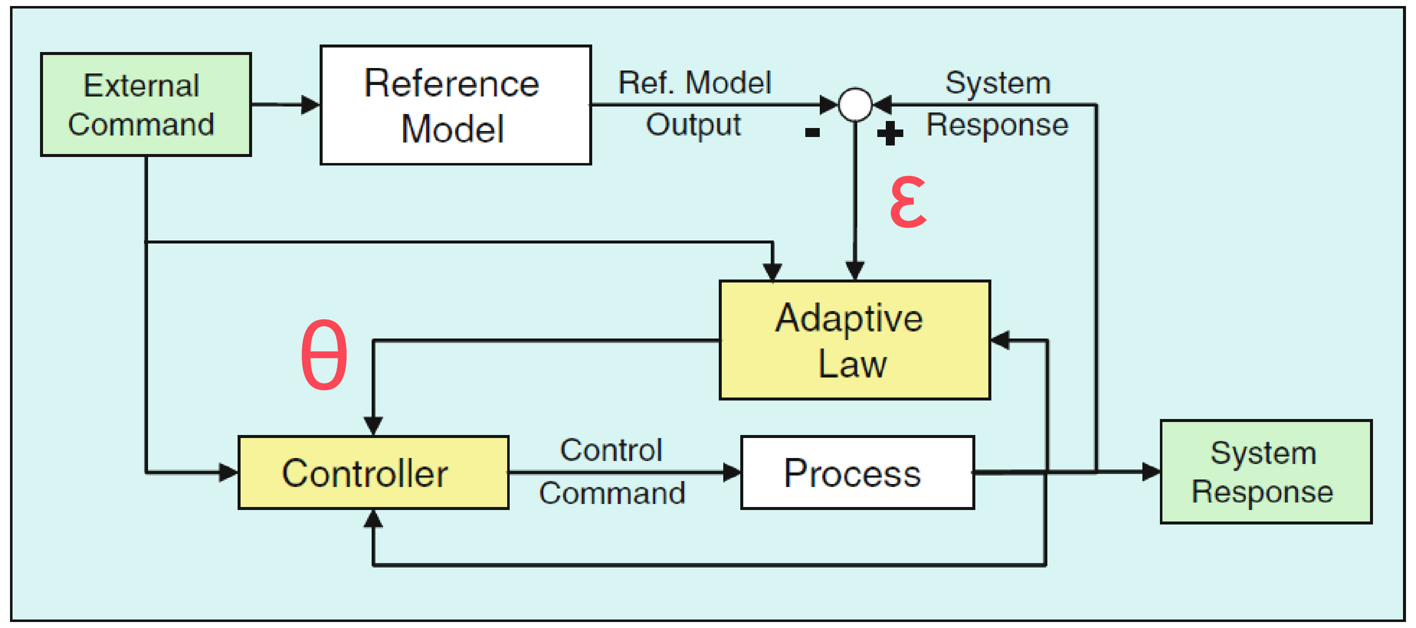
- φ sensitivity derivative → estimations required
  - φ can be a regression vector (filtered)
- ε is the prediction error
- γ is the adaptation gain

Åström, Karl J. and Wittenmark, Björn. (1995) Adaptive Control. Second Edition
Landau, I.D. , Lozano, R. , M'Saad, M. and Karimi, A. (2011) Adaptive Control. Communications and Control Engineering. Springer

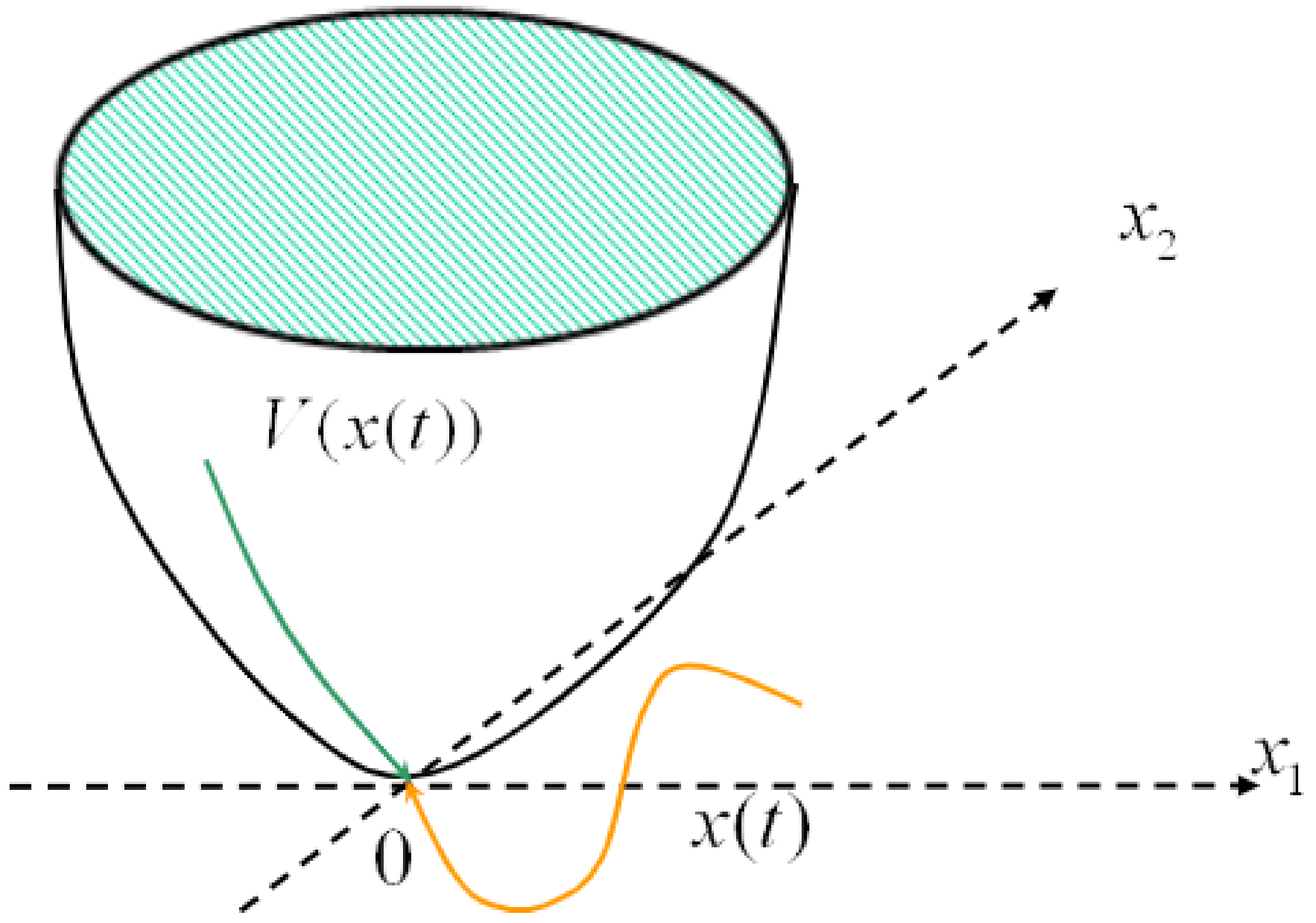
MRAC:
MODEL REFERENCE ADAPTIVE CONTROLLERS

LYAPUNOV-BASED ADAPTIVE LAW

- The design of the adaptive law is formulated as a stability problem.
- *State* is the error ($\varepsilon = y - y_m$) and the parameters (θ) → should be bounded
- Basic steps:
  - Find controller structure
  - Derive error equation
  - Find Lyapunov function
  - Derive a *parameter updating law* such that error will go to zero.



- Error converges to zero.
- Parameters may not converge to their correct values



Åström, Karl J. and Wittenmark, Björn. (1995) Adaptive Control. Second Edition
Landau, I.D. , Lozano, R. , M'Saad, M. and Karimi, A. (2011) Adaptive Control. Communications and Control Engineering. Springer
Mahmoud, M, Xia, Y. (2012) Applied Control Sytems Design

ADAPTIVE CONTROL

STABILITY and CONVERGENCE

- Adaptive law → multiplicative nonlinearity → non-linear closed-loop plant, often time-varying
- Proofs of global closed-loop stability and asymptotic convergence of the tracking error to zero
  - Not when approximate sensitivity functions are used

Åström, Karl J. and Wittenmark, Björn. Adaptive Control. Second Edition (1995)

Ioannou, Petros, A. and Sun, Jing. Robust Adaptive Control. (2012)

Landau, I.D. , Lozano, R. , M'Saad, M. and Karimi, A. (2011) Adaptive Control. Communications and Control Engineering. Springer

ADAPTIVE CONTROL

STABILITY and CONVERGENCE

- Adaptive law → multiplicative nonlinearity → non-linear closed-loop plant, often time-varying
- Proofs of global closed-loop stability and asymptotic convergence of the tracking error to zero
  - Not when approximate sensitivity functions are used

Direct method

- The stable error dynamics and adaptive laws are derived using the structure of the control signal

Indirect method

- The stable error dynamics and adaptive laws are derived independent of the control signal

Åström, Karl J. and Wittenmark, Björn. Adaptive Control. Second Edition (1995)

Ioannou, Petros, A. and Sun, Jing. Robust Adaptive Control. (2012)

Landau, I.D. , Lozano, R. , M'Saad, M. and Karimi, A. (2011) Adaptive Control. Communications and Control Engineering. Springer

ADAPTIVE CONTROL

STABILITY

- Adaptive control theorems:
  - If A, B and C hold, then all the signals in the loop are bounded and convergence occurs.
- Unknowness of the plant and a performance index that should be minimized.
  - But if plant is unknown, can performance index be minimized?

Åström, Karl J. and Wittenmark, Björn. (1995) Adaptive Control. Second Edition
Landau, I.D. , Lozano, R. , M'Saad, M. and Karimi, A. (2011) Adaptive Control. Communications and Control Engineering. Springer
Mahmoud, M, Xia, Y. (2012) Applied Control Sytems Design

ADAPTIVE CONTROL

STABILITY

- Adaptive control theorems:
  - If A, B and C hold, then all the signals in the loop are bounded and convergence occurs.
- Unknowness of the plant and a performance index that should be minimized.
  - But if plant is unknown, can performance index be minimized?
- Time-scale of identification step needs to be faster than plant variation timescale.

Åström, Karl J. and Wittenmark, Björn. (1995) Adaptive Control. Second Edition
Landau, I.D. , Lozano, R. , M'Saad, M. and Karimi, A. (2011) Adaptive Control. Communications and Control Engineering. Springer
Mahmoud, M, Xia, Y. (2012) Applied Control Sytems Design

ADAPTIVE CONTROL

STABILITY

- Adaptive control theorems:
  - If A, B and C hold, then all the signals in the loop are bounded and convergence occurs.
- Unknowness of the plant and a performance index that should be minimized.
  - But if plant is unknown, can performance index be minimized?
- Time-scale of identification step needs to be faster than plant variation timescale.
- Interaction between two processes can generate instability

X-15-3 flight accident.

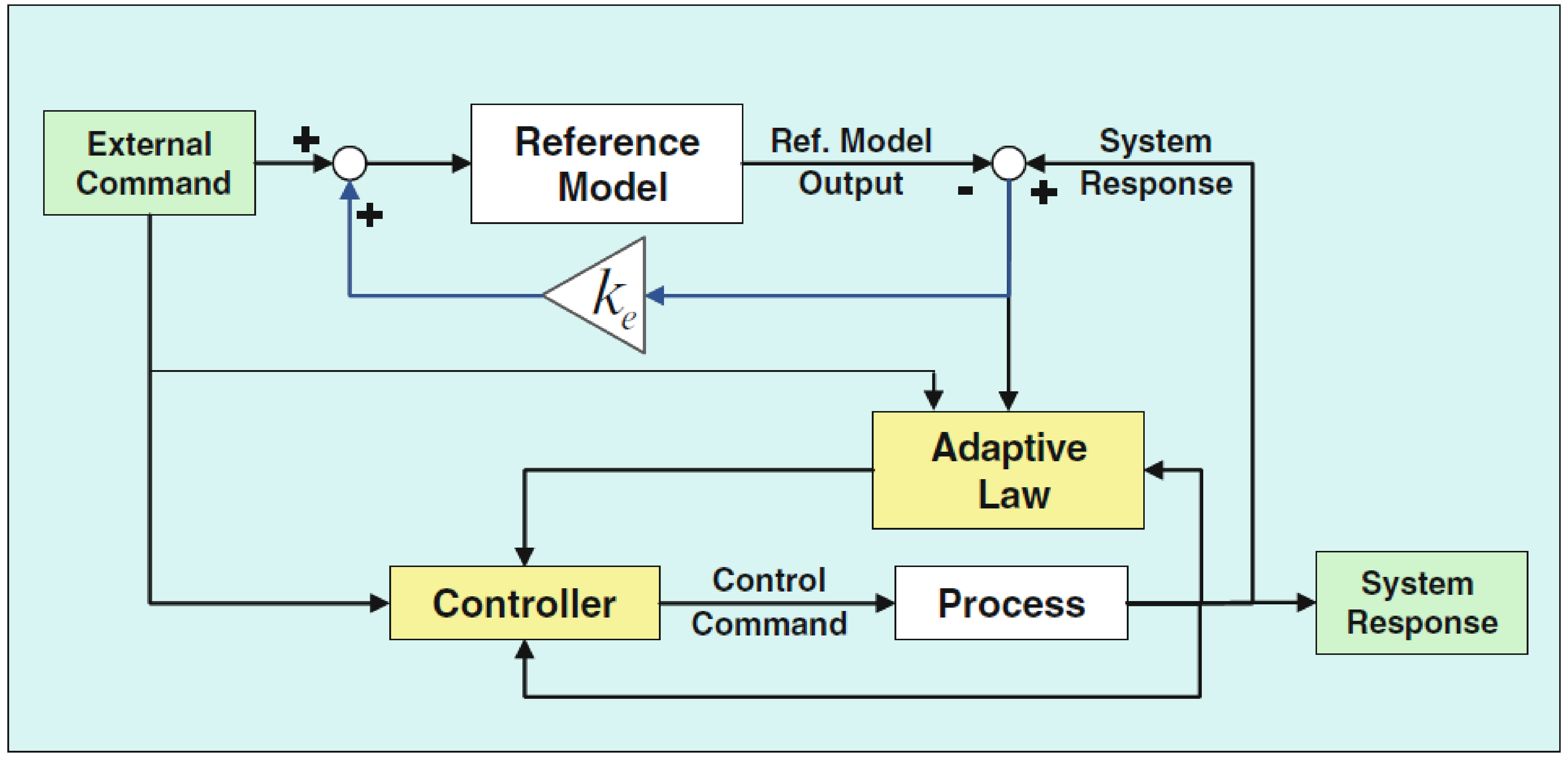
Limit cycle (1967)

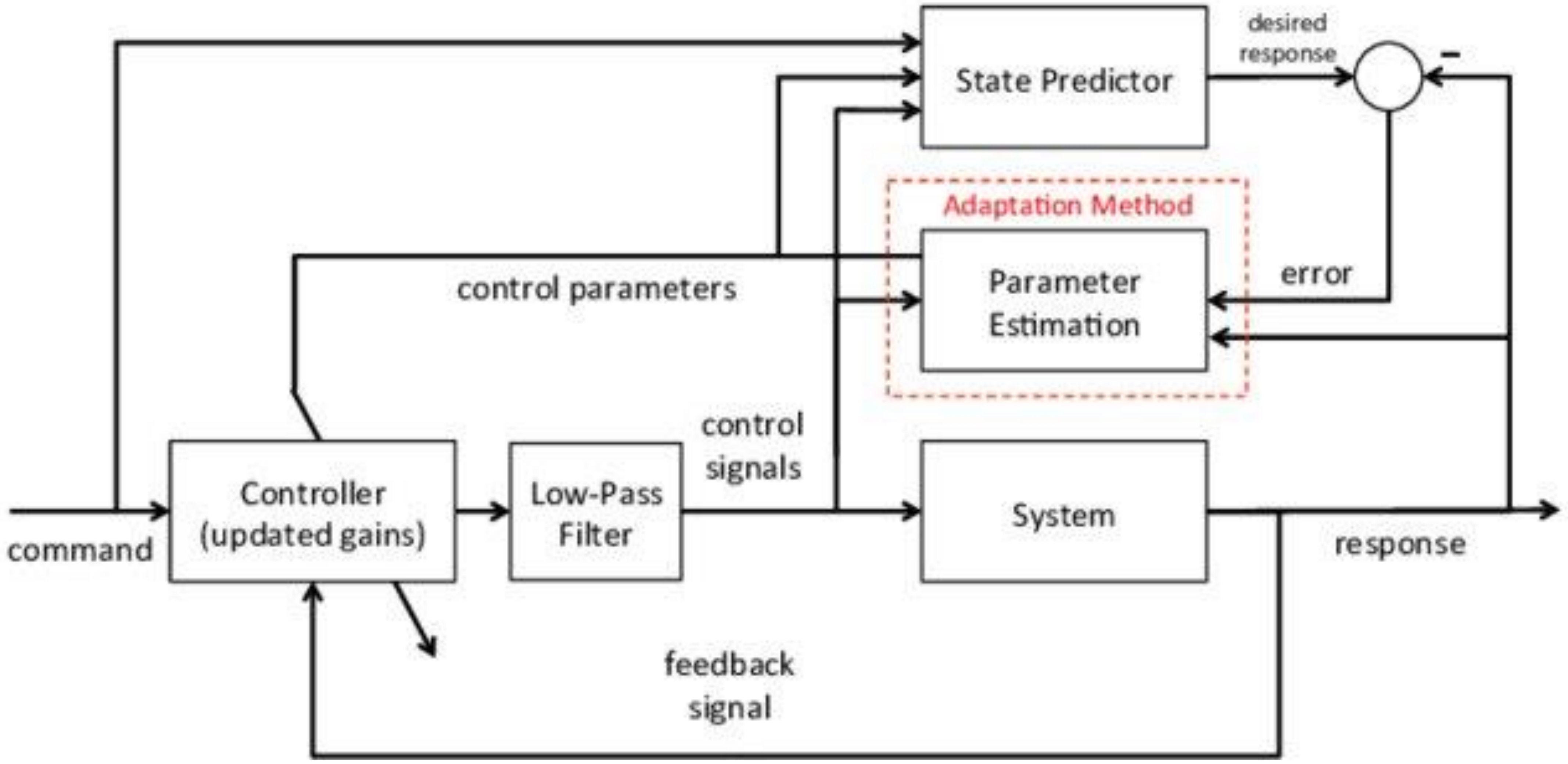
Åström, Karl J. and Wittenmark, Björn. (1995) Adaptive Control. Second Edition
Landau, I.D. , Lozano, R. , M'Saad, M. and Karimi, A. (2011) Adaptive Control. Communications and Control Engineering. Springer
Mahmoud, M, Xia, Y. (2012) Applied Control Sytems Design

ADAPTIVE CONTROL

STABILITY and CONVERGENCE

- Robust adaptive controller:
    - Guarantees signal boundedness in the presence of «reasonable» classes of unmodeled dynamics and bounded disturbances as well as performance error bounds within the modeling error.



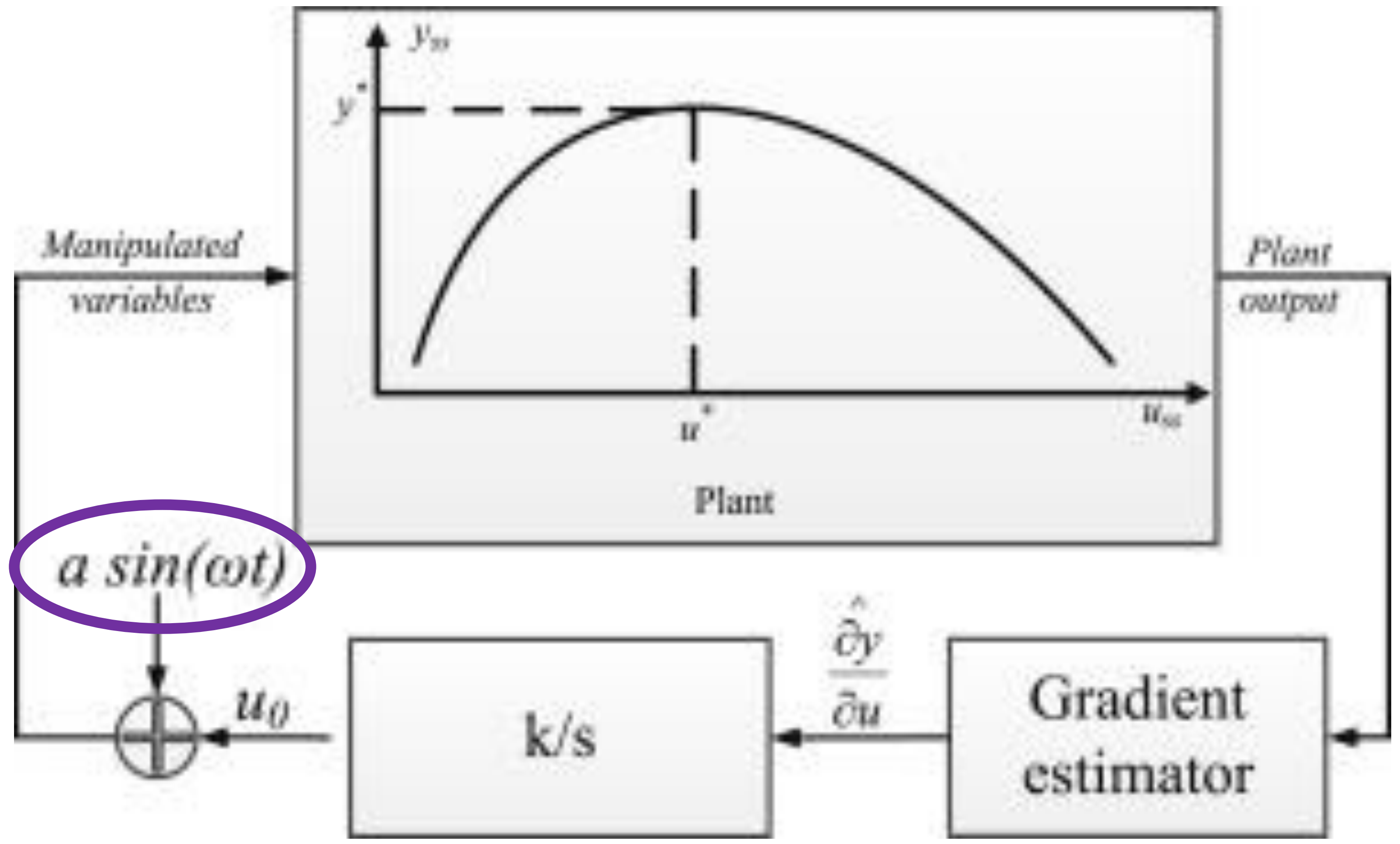


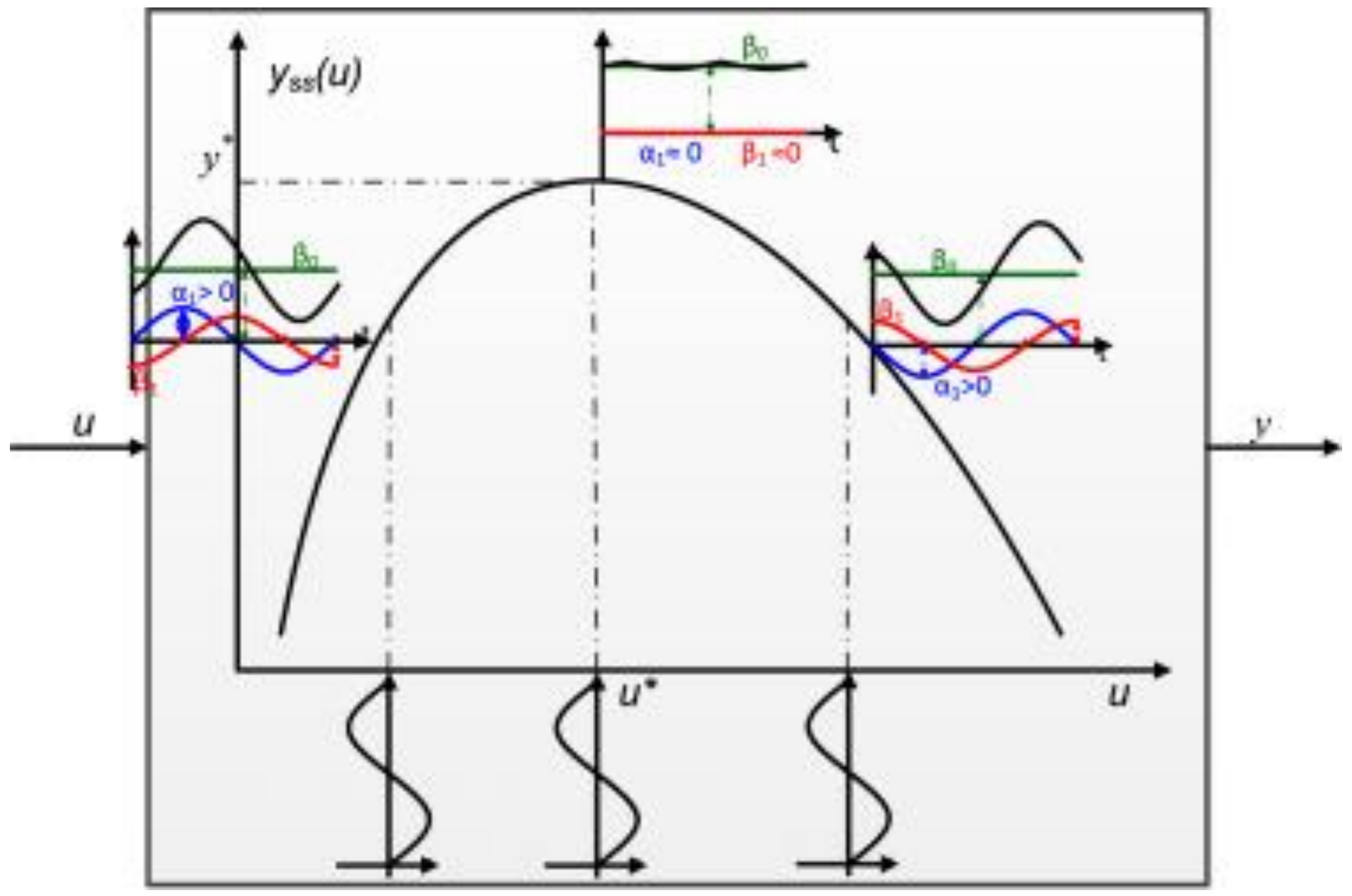
Bhattacharyya, S., Cofer, D., Musliner, D., Mueller, J., & Engstrom, E. (2015). Certification considerations for adaptive systems. In *2015 International Conference on Unmanned Aircraft Systems (ICUAS)* (pp. 270–279). IEEE.
Ioannou, Petros, A. and Sun, Jing. Robust Adaptive Control. (2012)
Lavretsky, E. and Wise, K. (2013) Robust and adaptive control with aeropsace appllications

EXTREMUM-SEEKING CONTROL

PRINCIPLE

- Single objective on-line (local) optimization.
- Data driven adaptive control (model-free)
- Setpoint selected to achieve a maximum of an uncertain *reference-to-output equilibrium map* opposed to known setpoints or reference trajectories
- Proof of stability exists



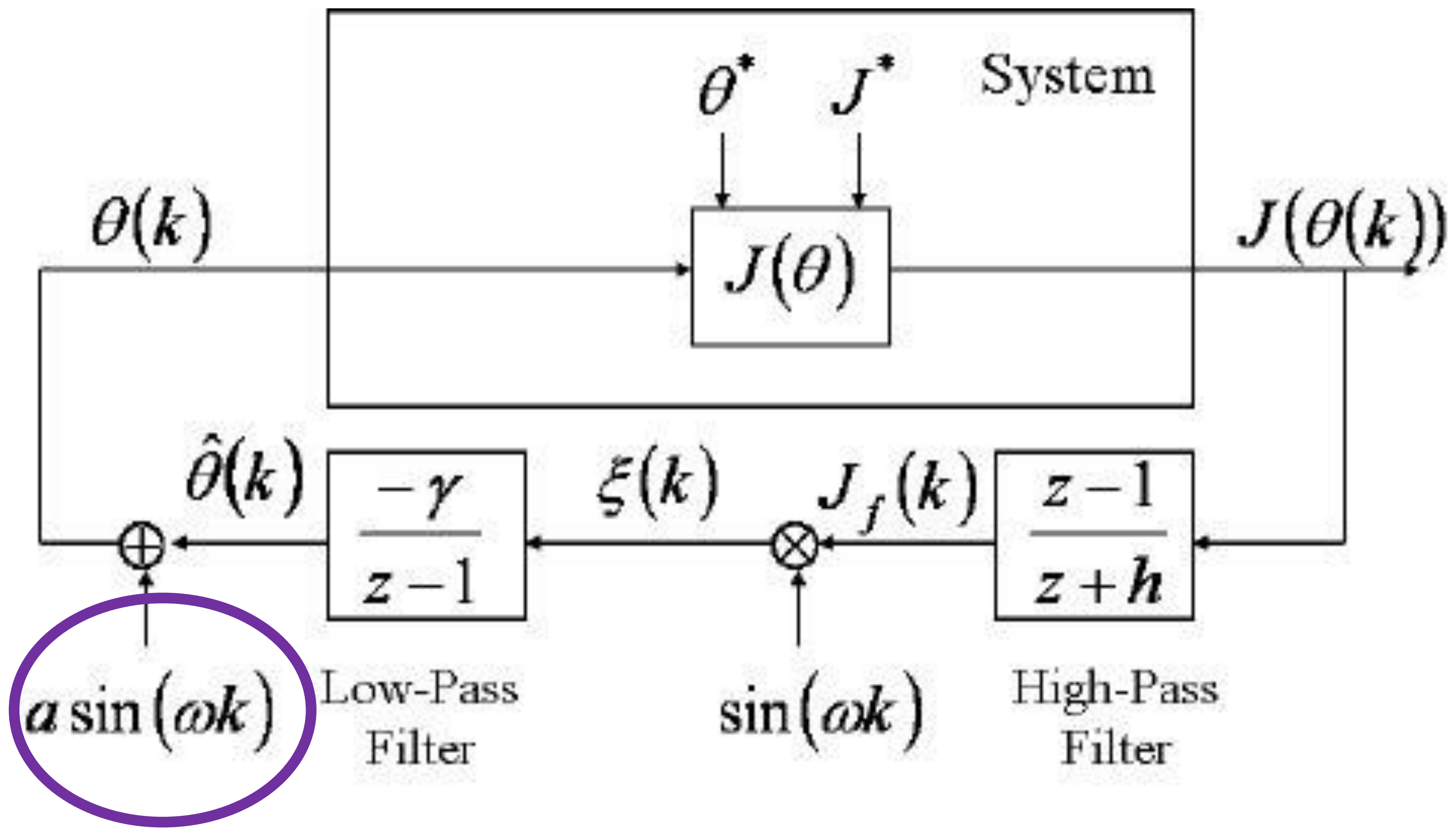


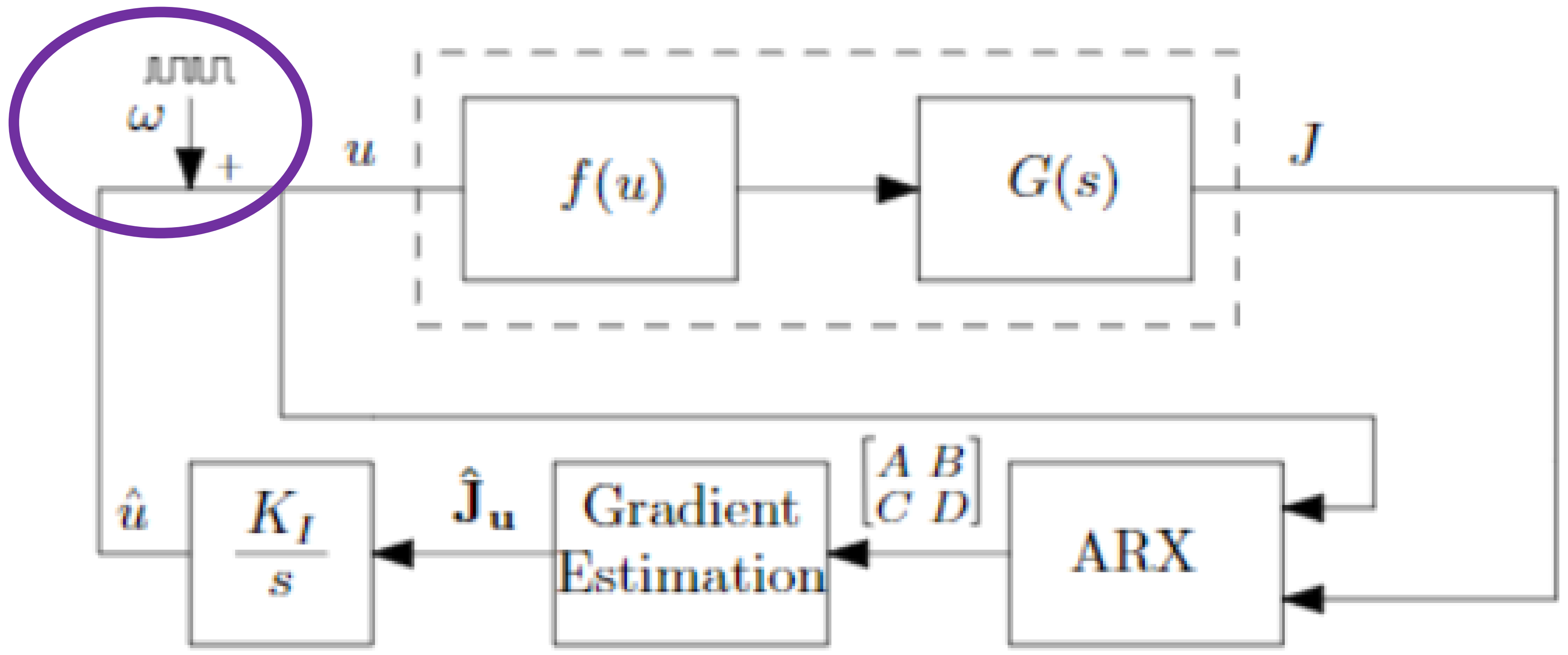
Krstić, M., & Wang, H.-H. (2000). Stability of extremum seeking feedback for general nonlinear dynamic systems. *Automatica*, *36*(4), 595–601.
Reghenzani, F., Formentin, S., Massari, G., & Fornaciari, W. (2018). A constrained extremum-seeking control for CPU thermal management. In *Proceedings of the 15th ACM International Conference on Computing Frontiers - CF '18* (pp. 320–325). New York, New York, USA: ACM Press.
Atta, K. T., Johansson, A., & Gustafsson, T. (2015). Extremum seeking control based on phasor estimation. *Systems & Control Letters*, *85*, 37–45.

EXTREMUM-SEEKING CONTROL

DIFFERRENT IMPLEMENTATIONS

- Single objective (local) on-line optimization.
- Data driven adaptive control (model-free)
- Setpoint selected to achieve a maximum of an uncertain *reference-to-output equilibrium map* opposed to known setpoints or reference trajectories
- Proof of stability exists

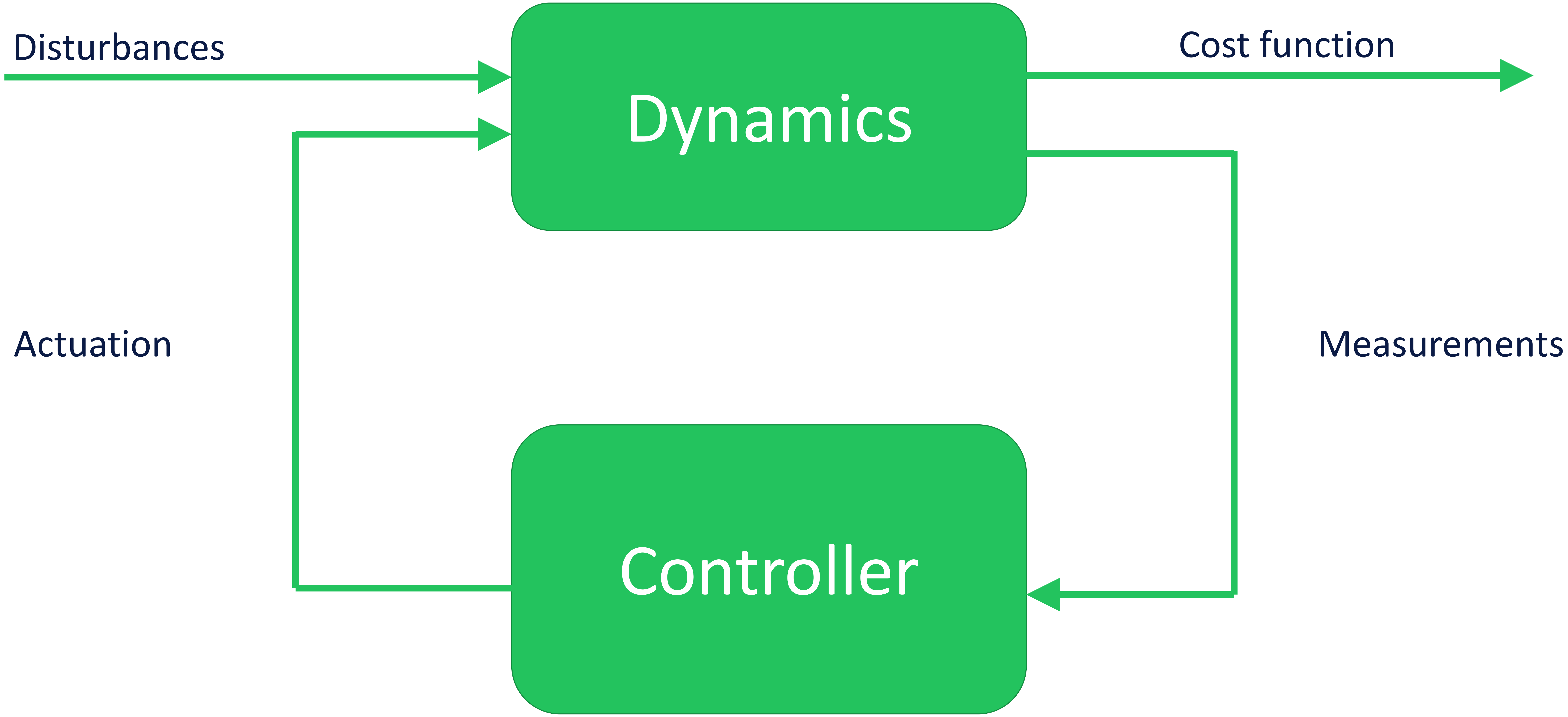




Krstić, M., & Wang, H.-H. (2000). Stability of extremum seeking feedback for general nonlinear dynamic systems. *Automatica*, *36*(4), 595–601.
Krishnamoorthy, D., Ryu, J., & Skogestad, S. (2019). A Dynamic Extremum Seeking Scheme Applied to Gas Lift Optimization. *IFAC-PapersOnLine*, *52*(1), 802–807.
Ou, Y., Xu, C., Schuster, E., Luce, T. C., Ferron, J. R., Walker, M. L., & Humphreys, D. A. (2008). Design and simulation of extremum-seeking open-loop optimal control of current profile in the DIII-D tokamak. *Plasma Physics and Controlled Fusion*, *50*(11), 115001.

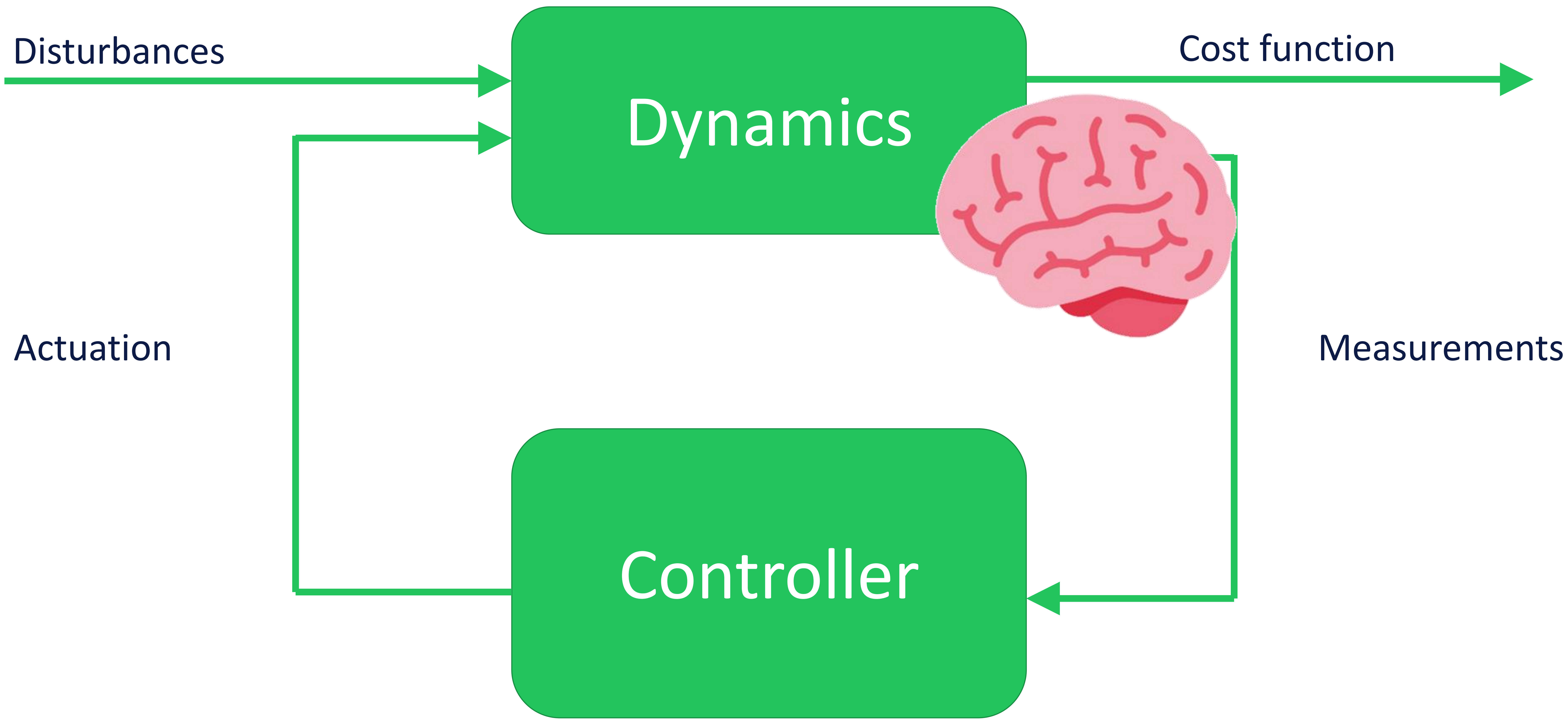
SELF-LEARNING CONTROL

MAIN IDEA



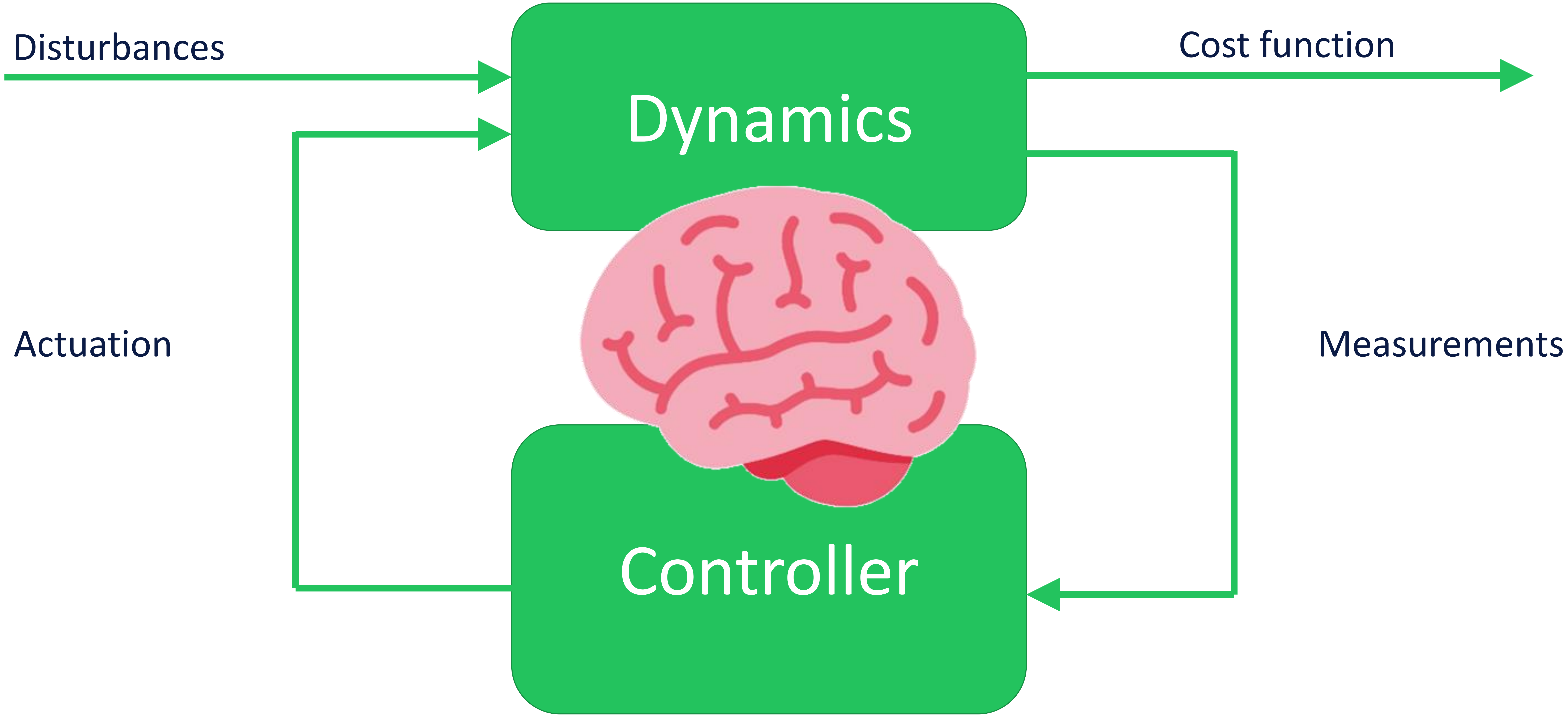
SELF-LEARNING CONTROL

MAIN IDEA



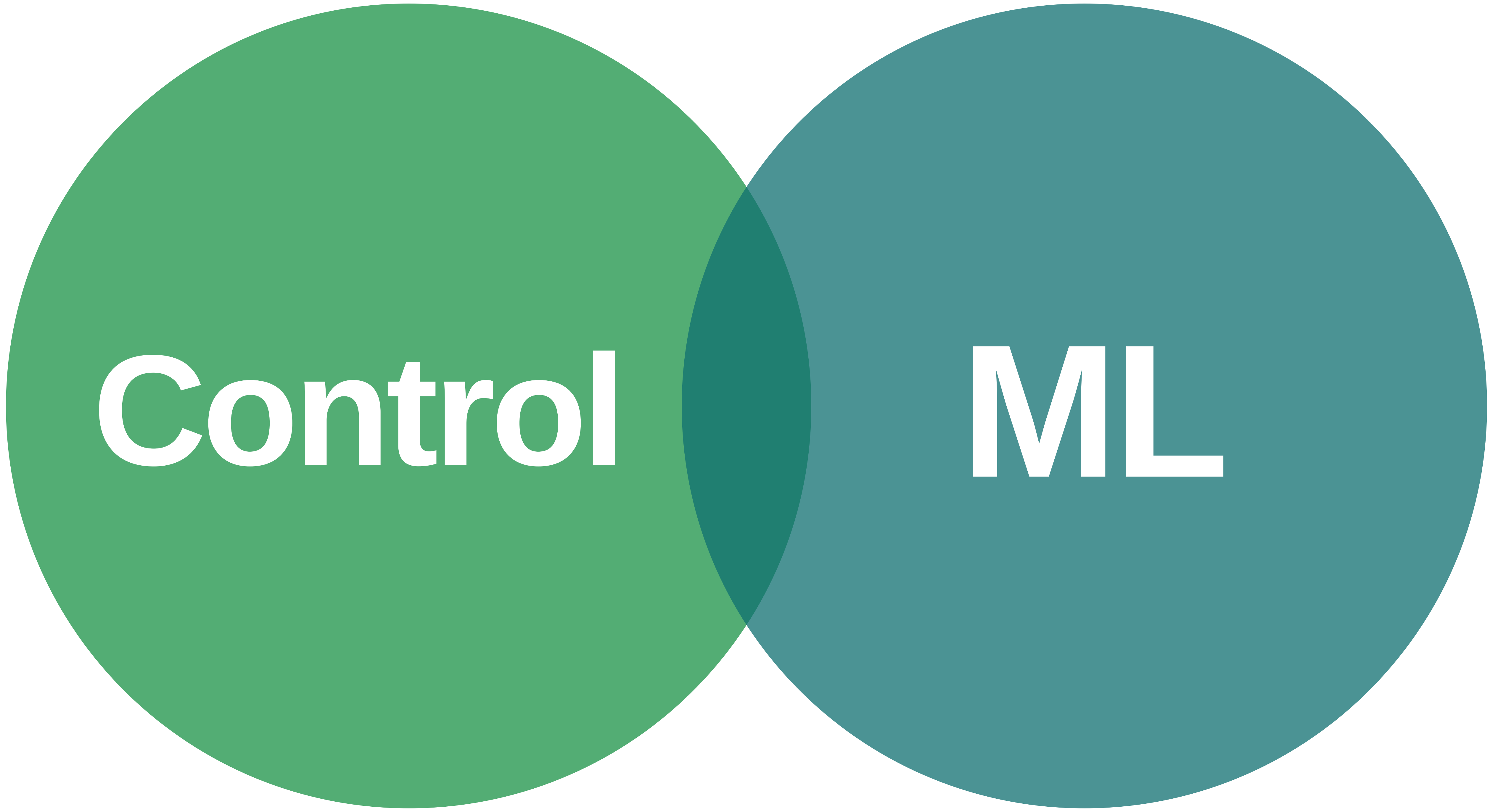
SELF-LEARNING CONTROL

MAIN IDEA



MACHINE LEARNING FOR CONTROL

MAIN IDEA

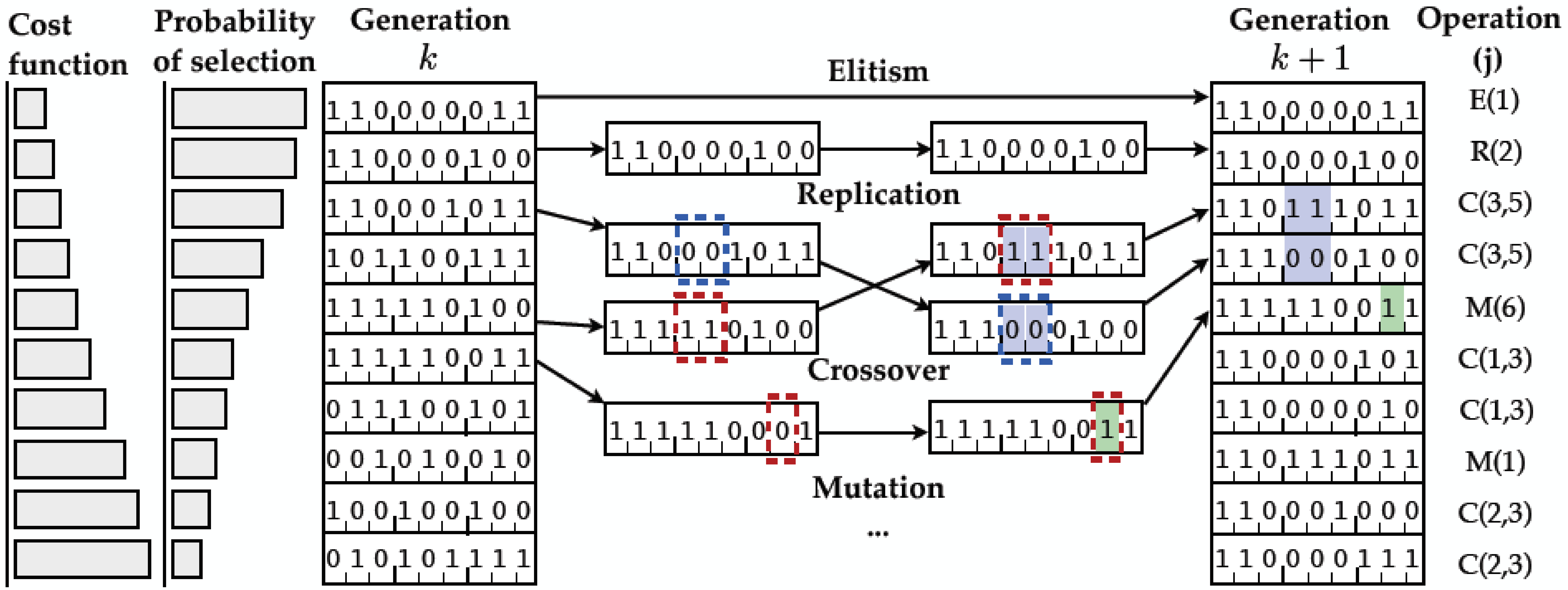


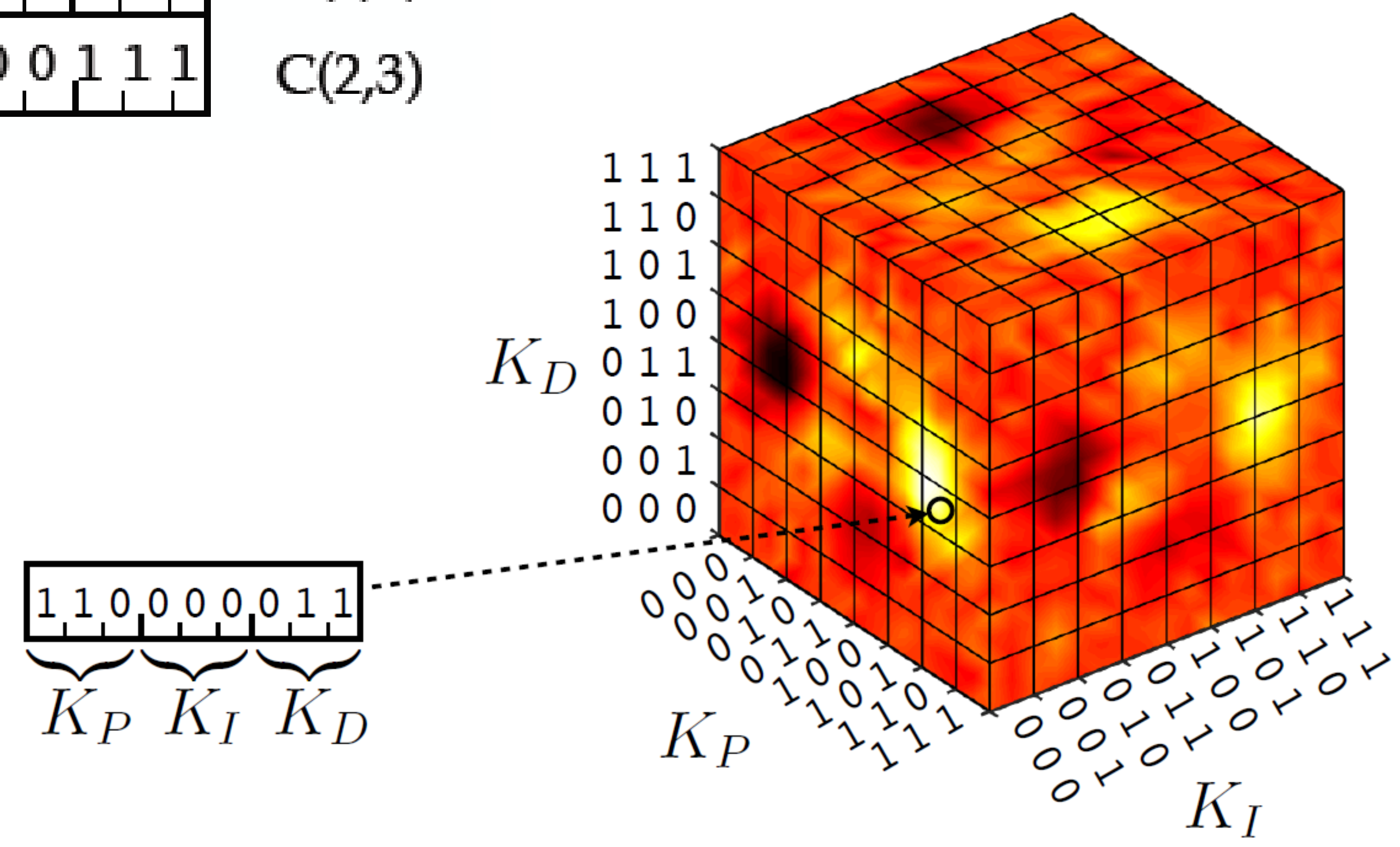
Optimization constrained by dynamics

Optimization based on data → without having *a priori* models of the dynamics

GENETIC ALGORITHMS IN CONTROL

MAIN IDEA





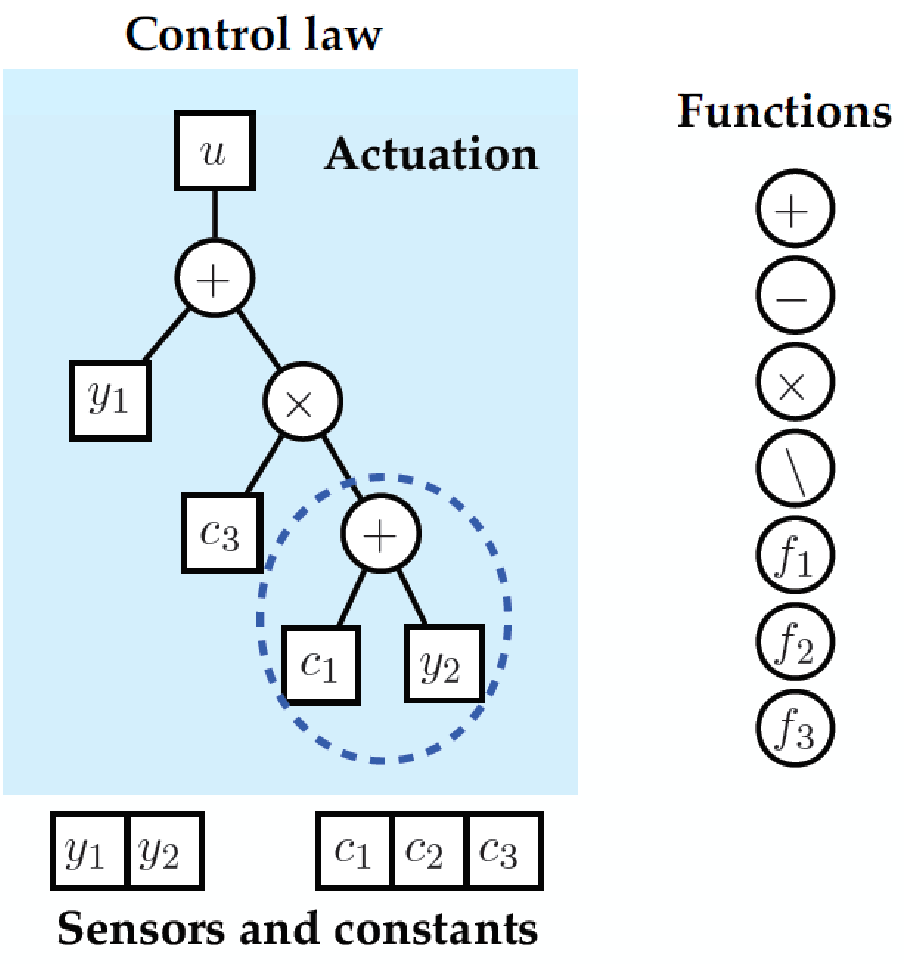
- Parameter estimation/ Model identification

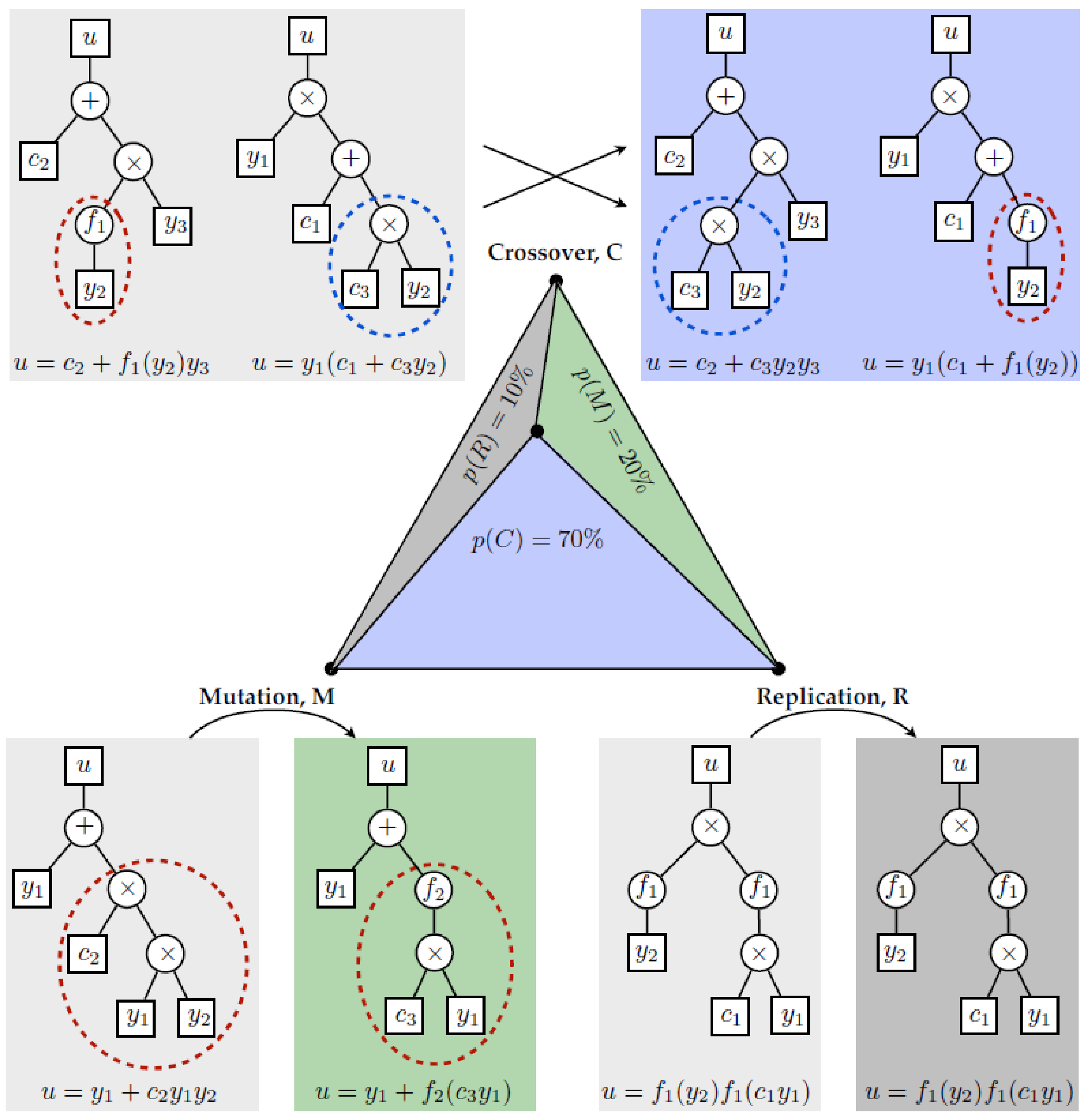
Brunton, Steven and Kutz, Nathan. (2017) *Data Driven Science & Engineering. Machine Learning, Dynamical Systems and Control*

GENETIC PROGRAMMING

MAIN IDEA

- Simultaneously learns structure and parameters of the controller.
- Similar operations as genetic algorithms
- Functions can also include transfer functions (e.g. integration of error).
- Control law defined by tree
- Requires a large number of experiments
  - The effect of the changed control law and parameters should be measured fast





Duriez, Thomas, Brunton, Steven, Noack, Bernd R. (2017) *Machine Learning Control – Taming Nonlinear Dynamics and Turbulence* Springer
Brunton, Steven and Kutz, Nathan. (2017) *Data Driven Science & Engineering. Machine Learning, Dynamical Systems and Control*

REINFORCEMENT LEARNING

MAIN IDEA

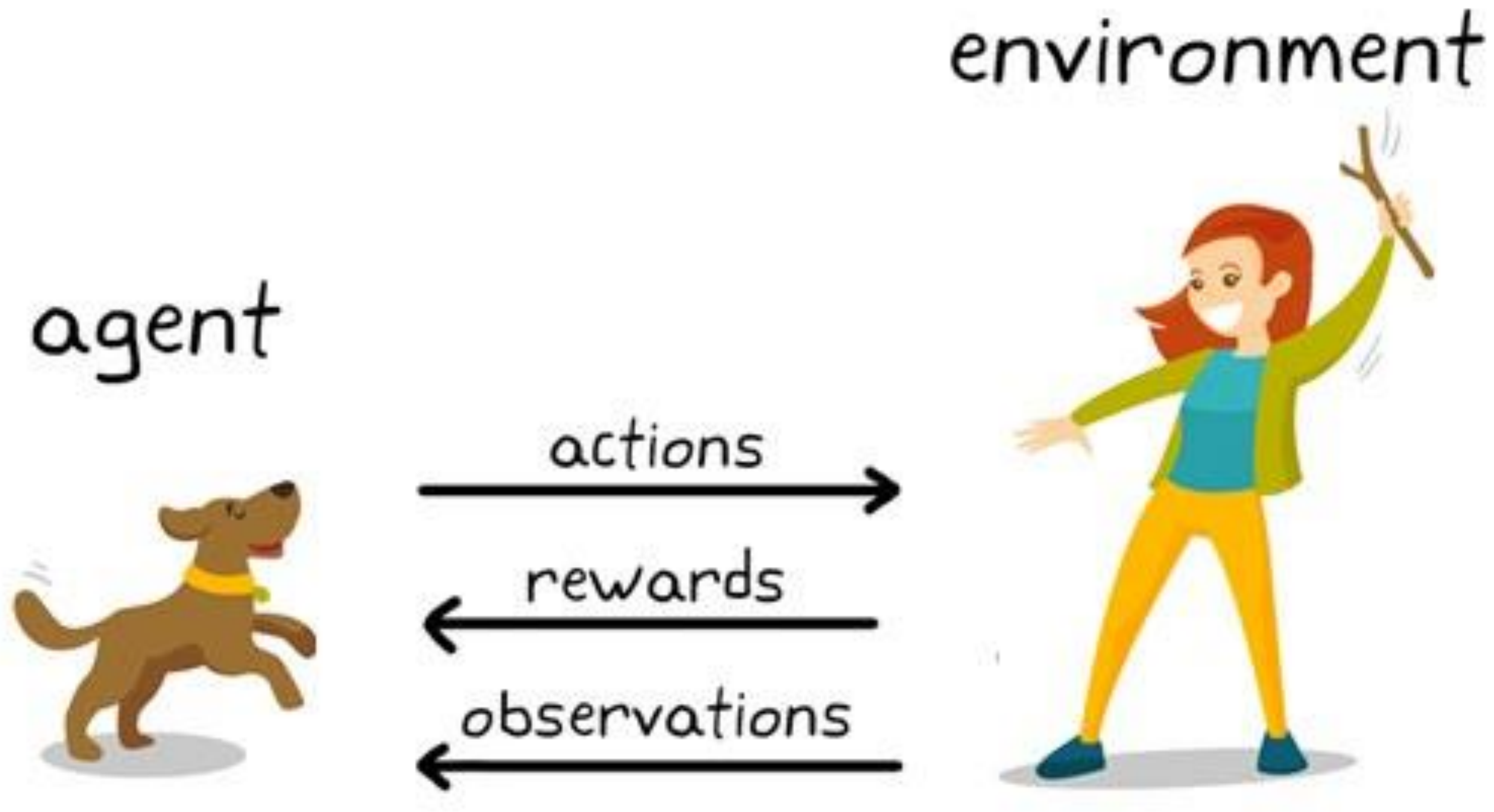
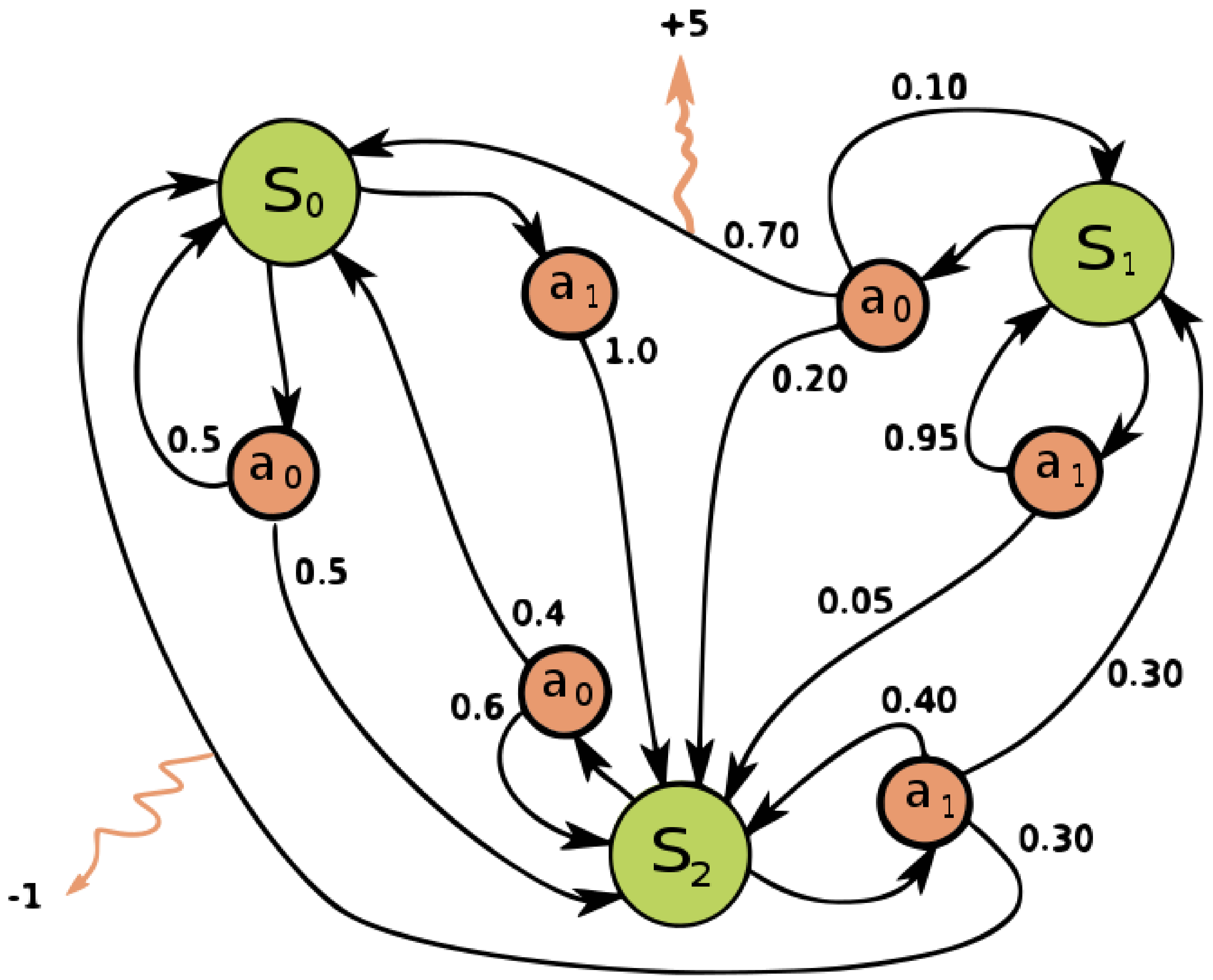


Image taken from www.kdnuggets.com/2019/10/mathworks-reinforcement-learning.html

REINFORCEMENT LEARNING

MAIN IDEA



Markov Decision Processes

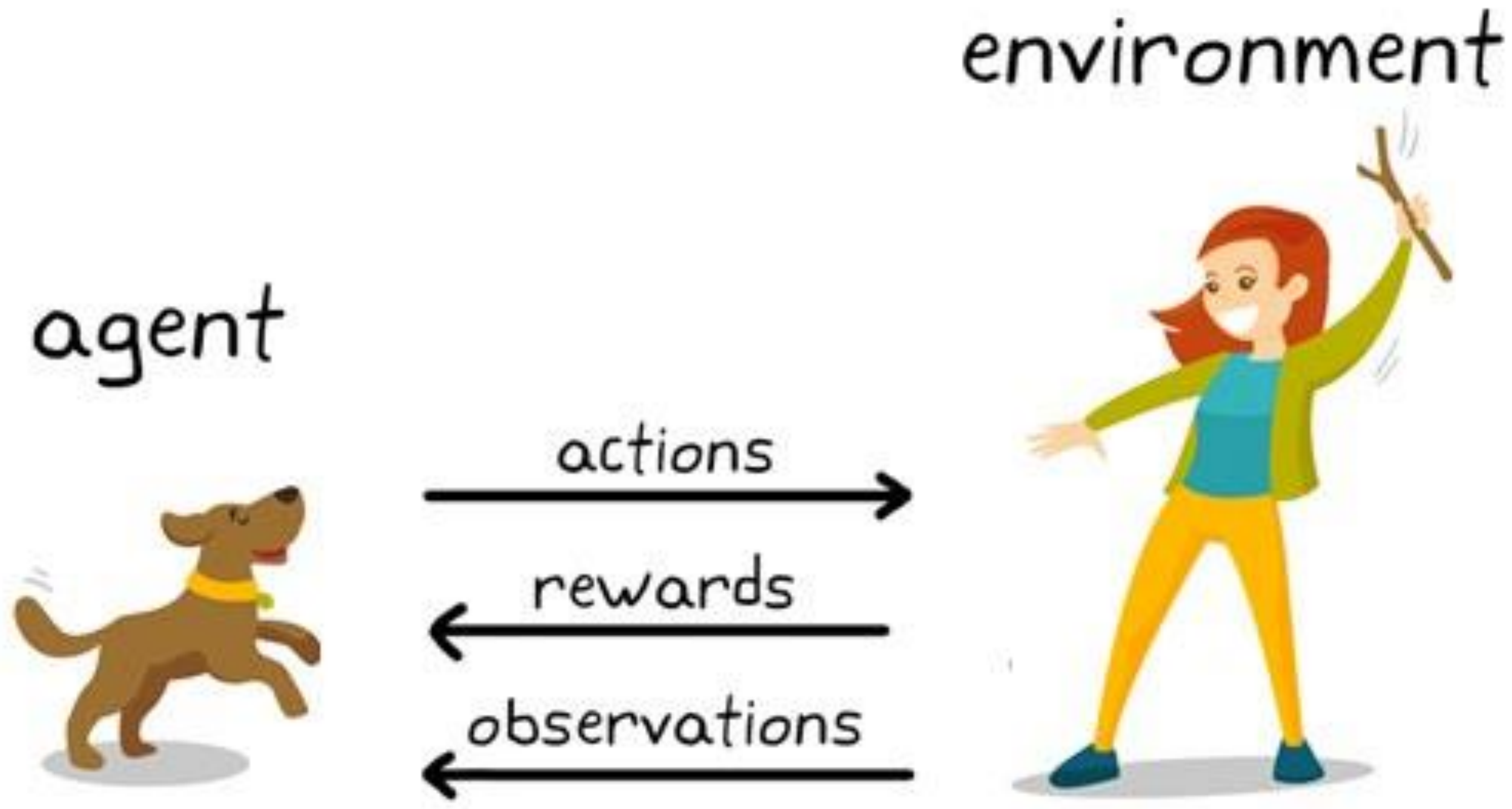
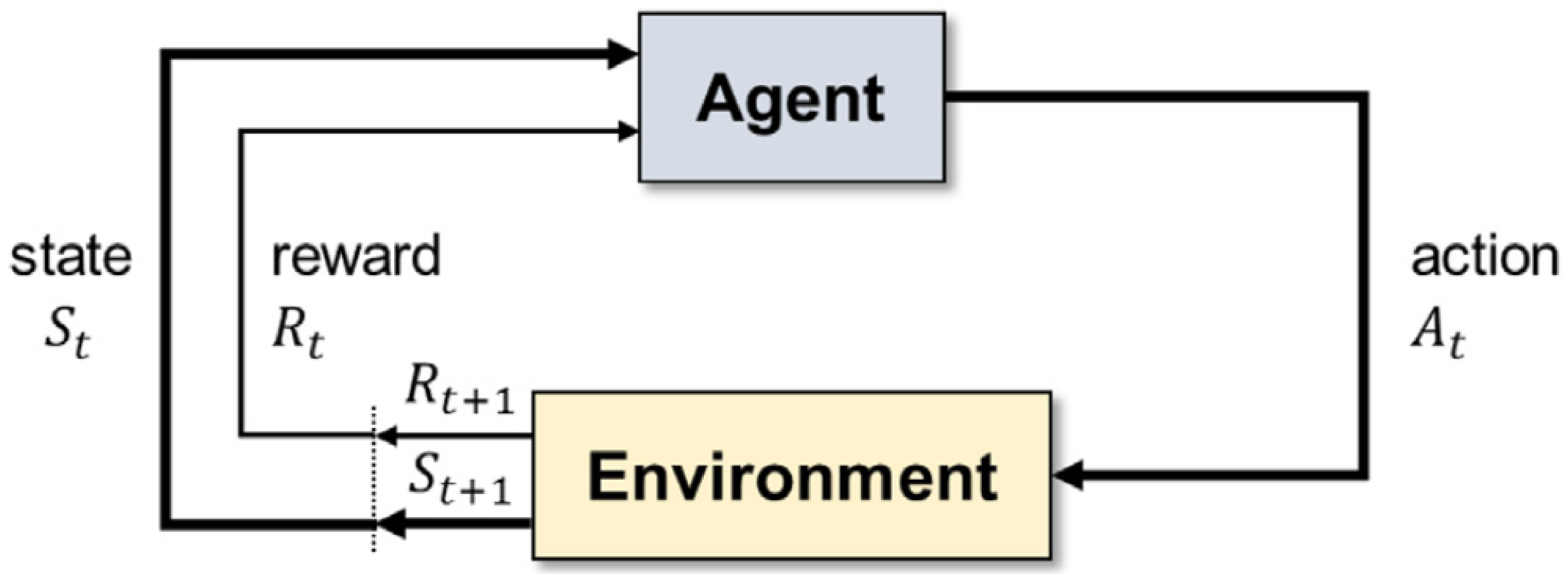
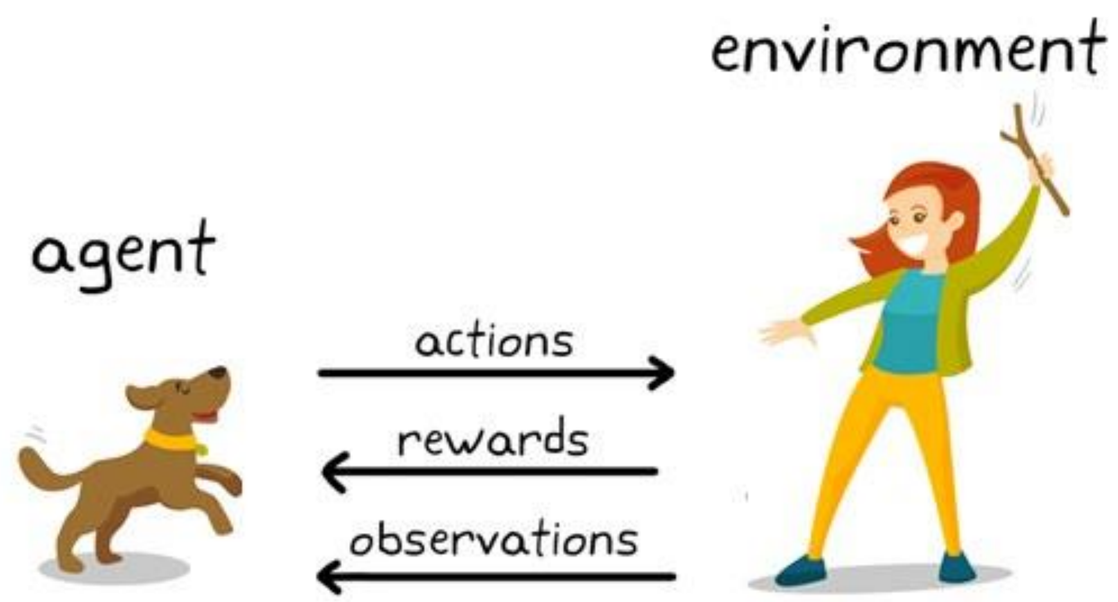


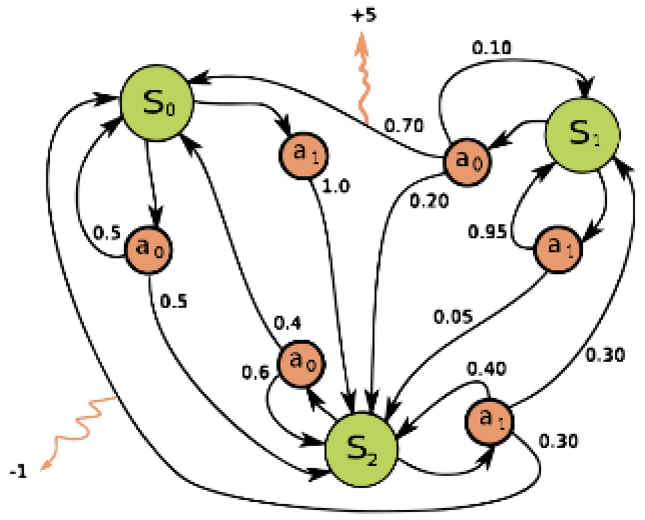
Image taken from www.kdnuggets.com/2019/10/mathworks-reinforcement-learning.html

REINFORCEMENT LEARNING

BASIC SETTING







Time *t-1*

Time *t*

The Agent takes an action A_t that affects the environment

Time *t+1*

Environment transitions from state S_t to S_{t+1}

An immediate reward R_{t+1} is generated

State information

Reward information

The Agent takes an action A_{t+1} that affects the environment

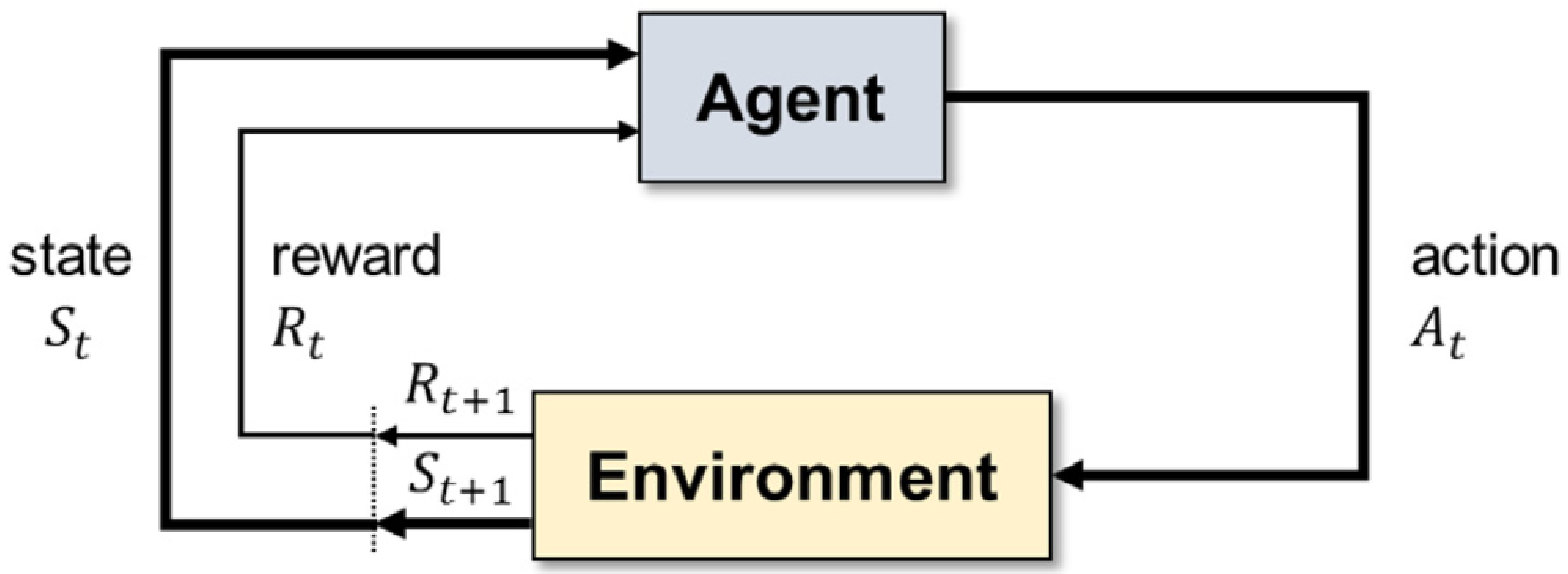
Time *t+2*

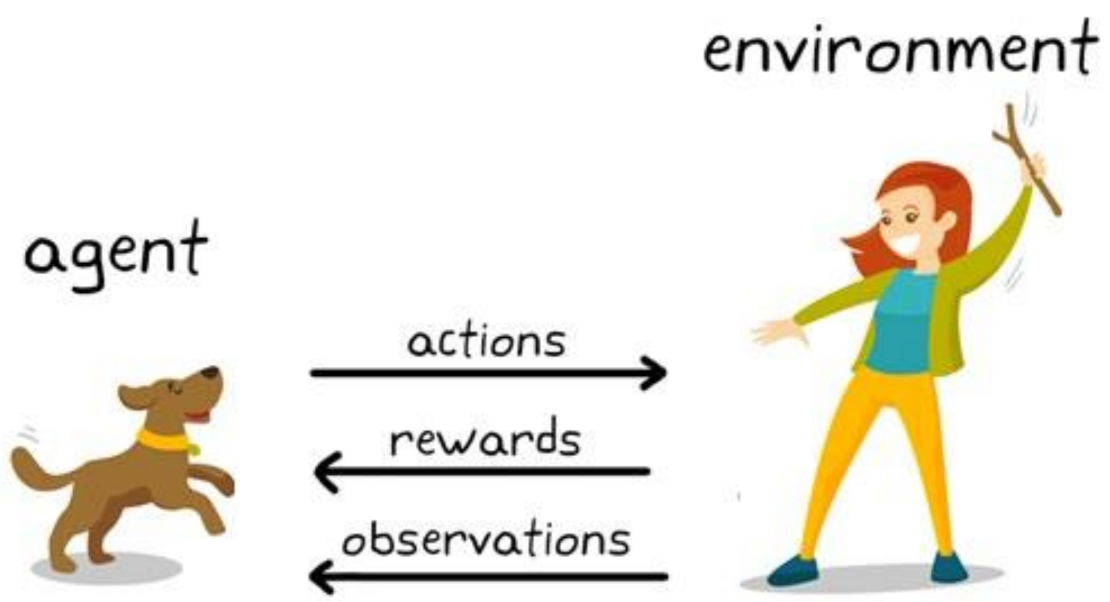
Shin, J., Badgwell, T. A., Liu, K.-H., & Lee, J. H. (2019). Reinforcement Learning – Overview of recent progress and implications for process control. *Computers & Chemical Engineering*, *127*, 282–294.

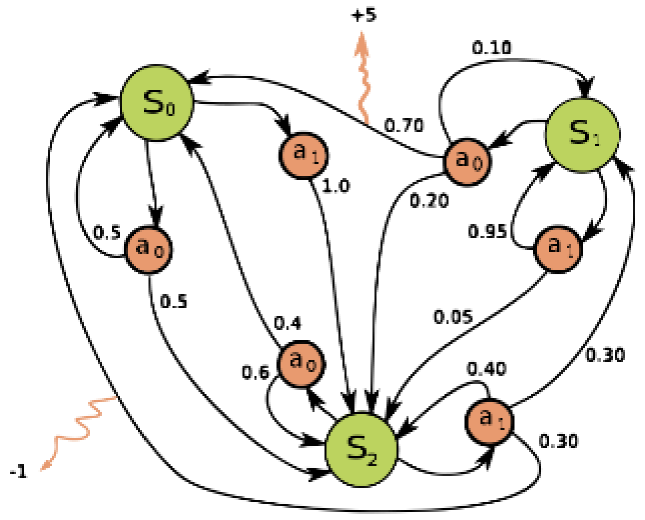
REINFORCEMENT LEARNING

BASIC SETTING

POLICY π
Mapping from states to actions







Time *t-1*

Time *t*

The Agent takes an action A_t that affects the environment

Time *t+1*

Environment transitions from state S_t to S_{t+1}

An immediate reward R_{t+1} is generated

State information

Reward information

The Agent takes an action A_{t+1} that affects the environment

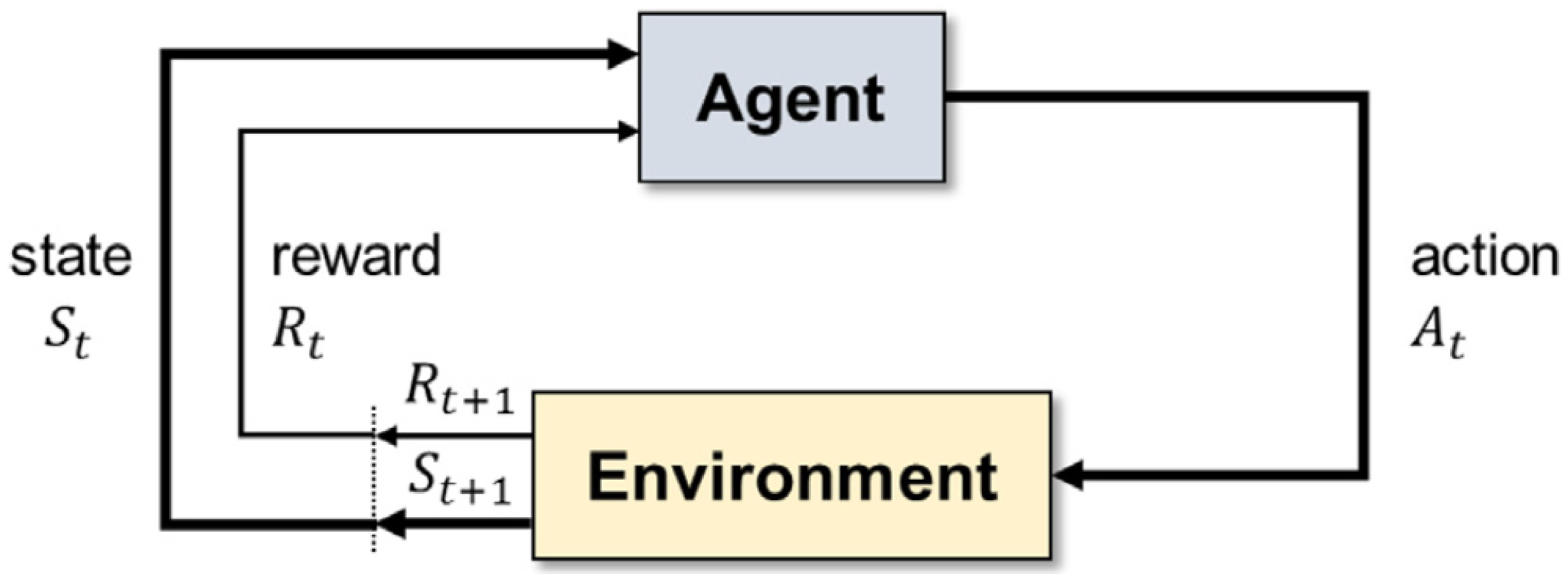
Time *t+2*

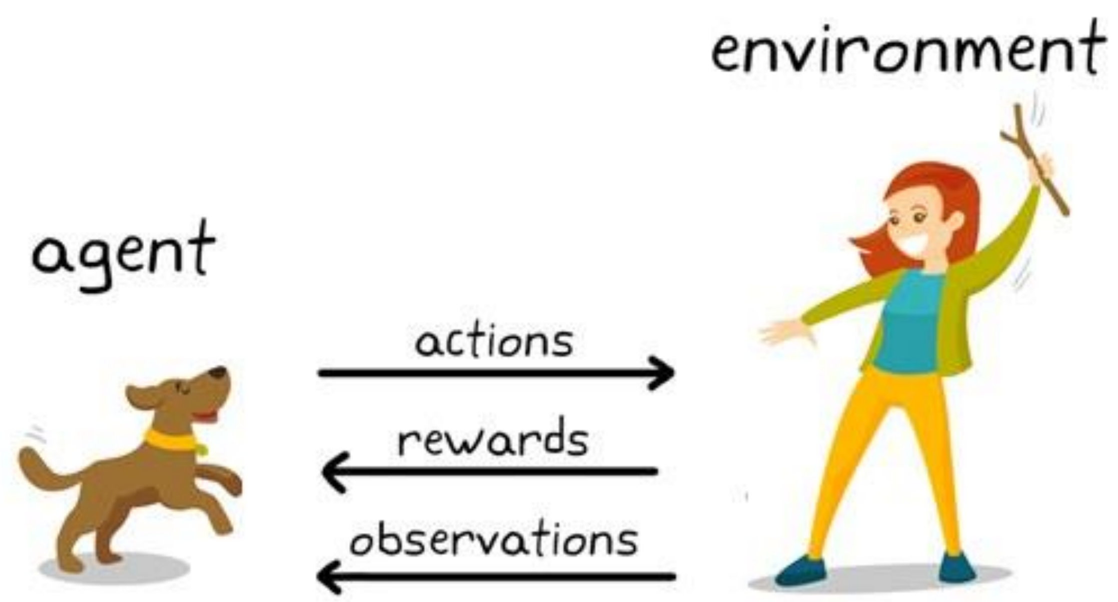
Shin, J., Badgwell, T. A., Liu, K.-H., & Lee, J. H. (2019). Reinforcement Learning – Overview of recent progress and implications for process control. *Computers & Chemical Engineering*, *127*, 282–294.

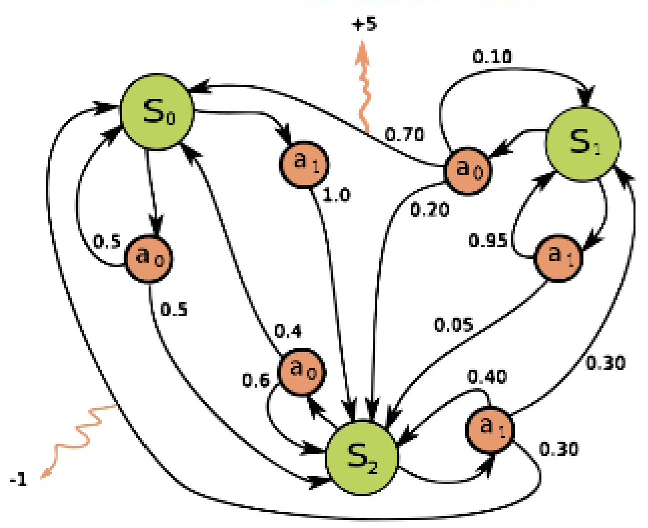
REINFORCEMENT LEARNING

BASIC SETTING

POLICY π
Mapping from states to actions







GOAL:
To learn a policy that maximizes the value function

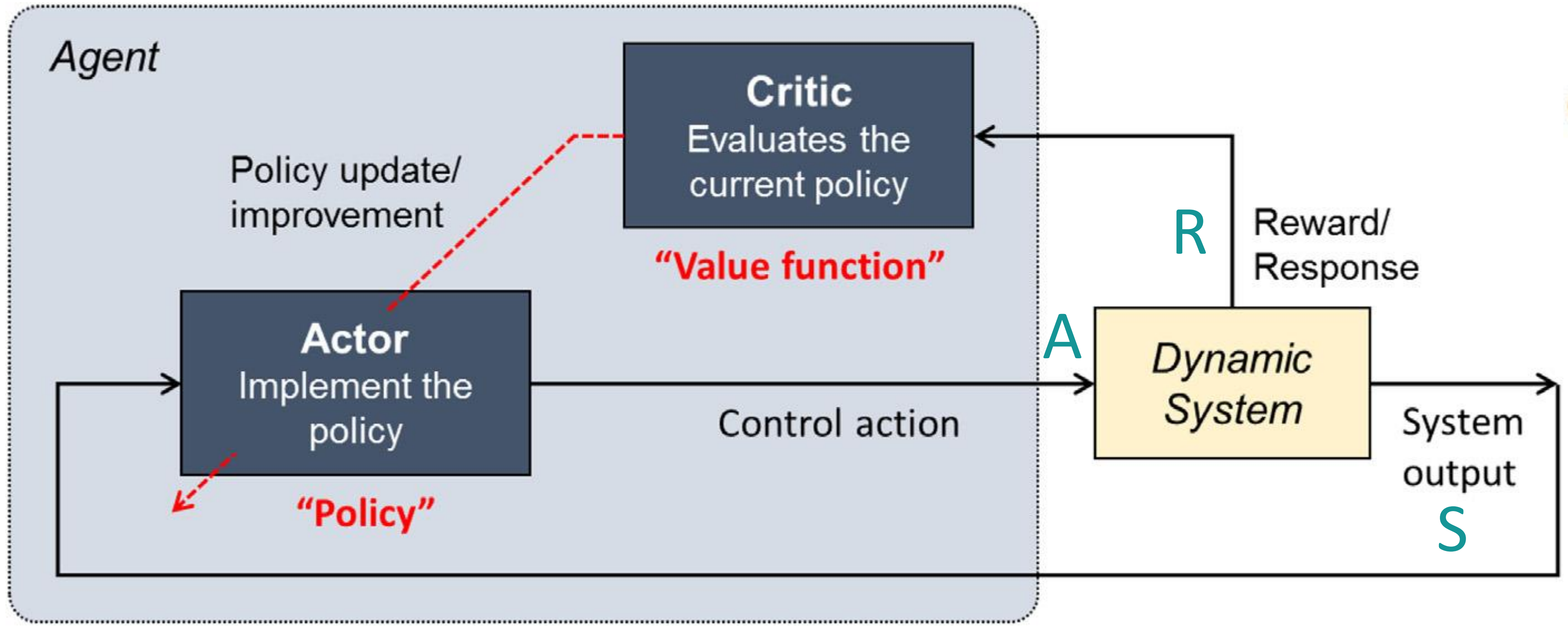
VALUE FUNCTION:
Long term sum of (expected) future rewards

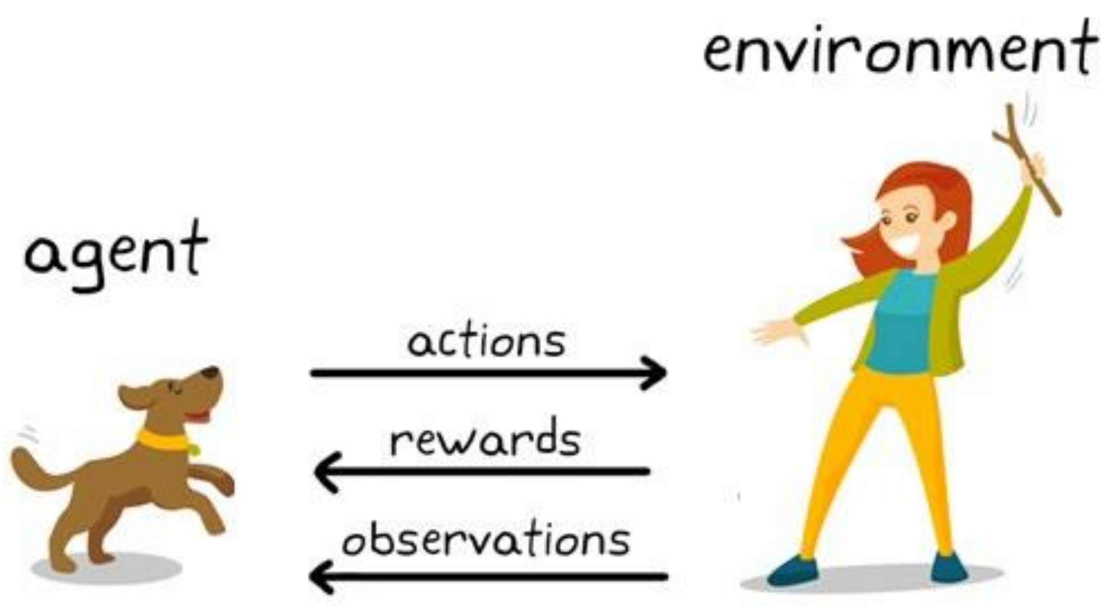
Shin, J., Badgwell, T. A., Liu, K.-H., & Lee, J. H. (2019). Reinforcement Learning – Overview of recent progress and implications for process control. *Computers & Chemical Engineering*, *127*, 282–294.

REINFORCEMENT LEARNING

BASIC SETTING

POLICY π
Mapping from states to actions





GOAL:
To learn a policy that maximizes the value function

VALUE FUNCTION:
Long term sum of (expected) future rewards

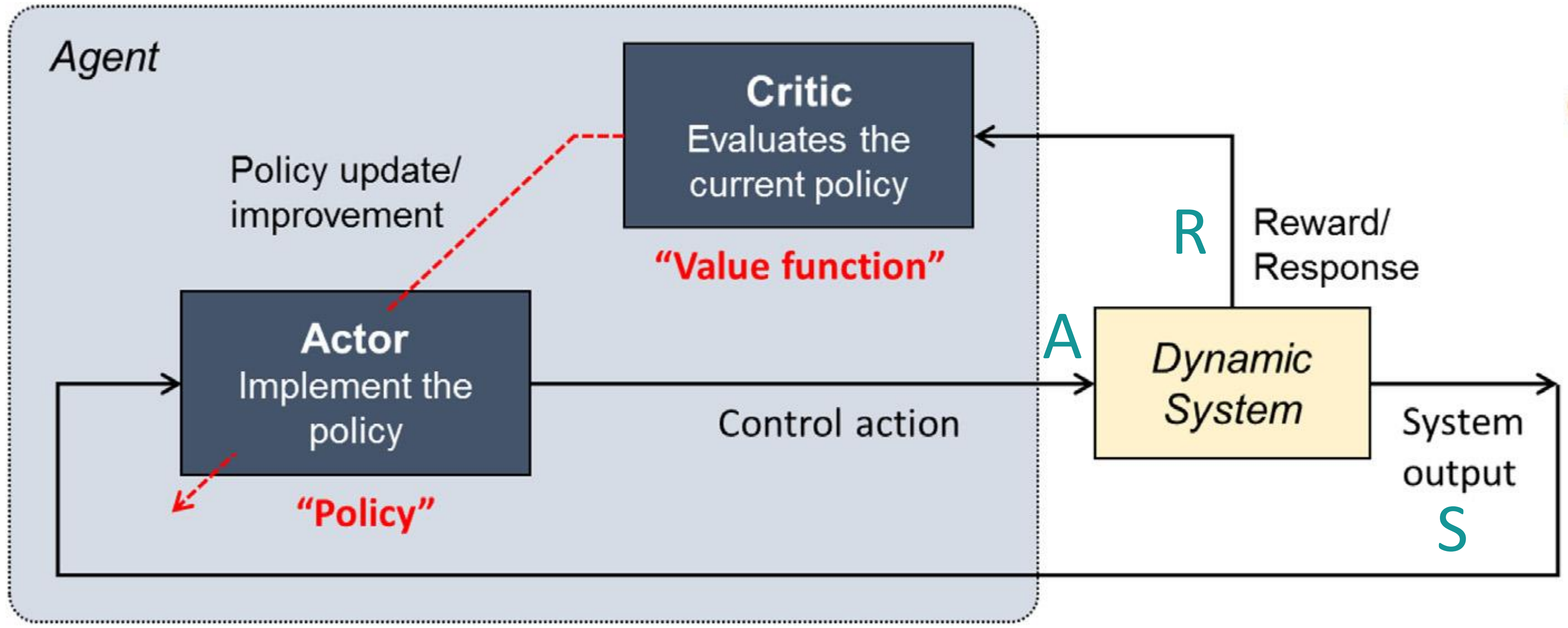
$$v_\pi(s) = E\left\{R_{t+1} + \gamma R_{t+2} + \gamma^2 R_{t+3} \cdots |S_t = s\right\}$$

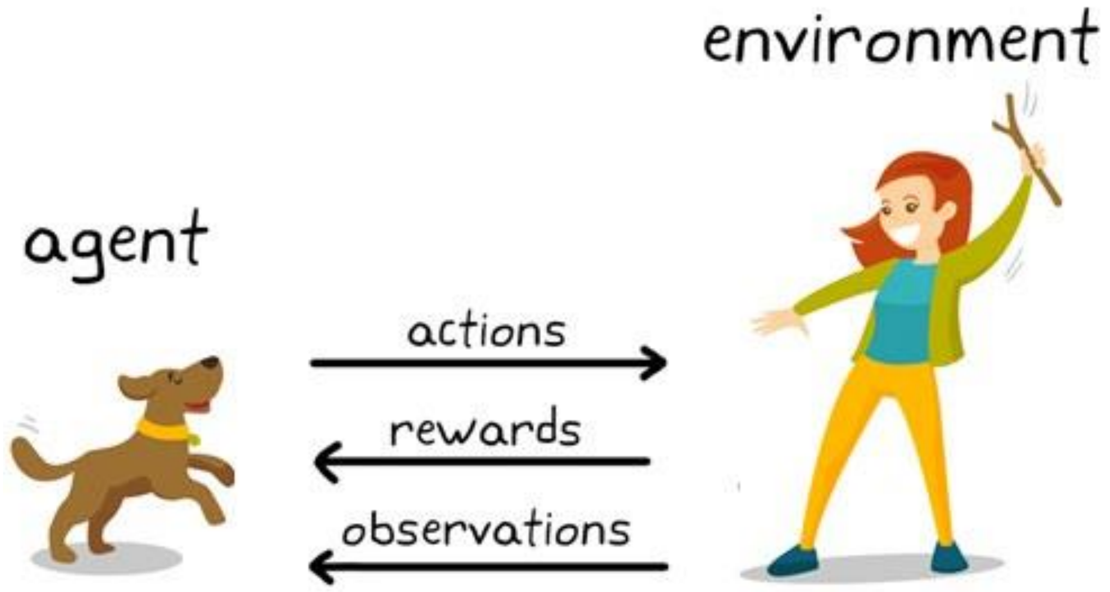
Shin, J., Badgwell, T. A., Liu, K.-H., & Lee, J. H. (2019). Reinforcement Learning – Overview of recent progress and implications for process control. *Computers & Chemical Engineering*, *127*, 282–294.

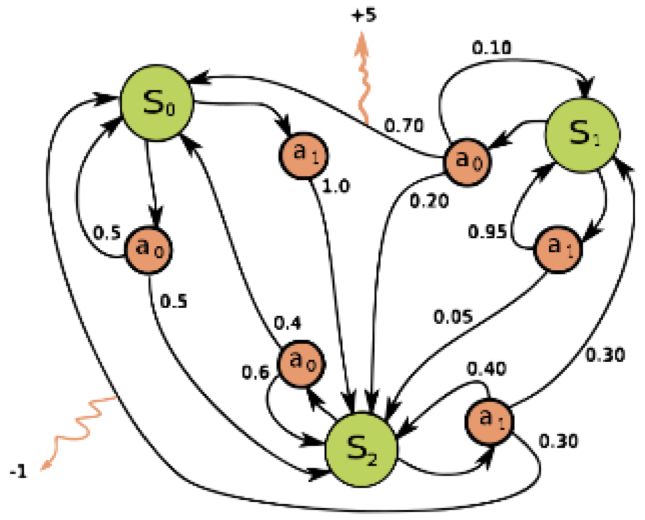
REINFORCEMENT LEARNING

BASIC SETTING

POLICY π
Mapping from states to actions







GOAL:
To learn a policy that maximizes the value function

VALUE FUNCTION:
Long term sum of (expected) future rewards

$$v_\pi(s) = E\left\{R_{t+1} + \gamma R_{t+2} + \gamma^2 R_{t+3} \cdots | S_t = s\right\}$$

$$v_*(s) = \max_a \sum_{s',r} p(s',r|s,a)\left[r + \gamma v_*(s')\right]$$

Bellman's optimality equation

Shin, J., Badgwell, T. A., Liu, K.-H., & Lee, J. H. (2019). Reinforcement Learning – Overview of recent progress and implications for process control. *Computers & Chemical Engineering*, *127*, 282–294.

REINFORCEMENT LEARNING

LIMITATIONS OF BASIC SETTING

- Model is unknown
- State dimension is large

$$v_*(s) = \max_a \sum_{s',r} p(s', r|s, a)\left[r + \gamma v_*(s')\right]$$

SOLUTION APPROACH

- Model-based
- Value-based (model-free)
- Policy-gradient (model-free)
- Actor-critic (model-free)

Convergence is achieved.

Shin, J., Badgwell, T. A., Liu, K.-H., & Lee, J. H. (2019). Reinforcement Learning – Overview of recent progress and implications for process control. *Computers & Chemical Engineering*, *127*, 282–294.

REINFORCEMENT LEARNING

CHALLENGES

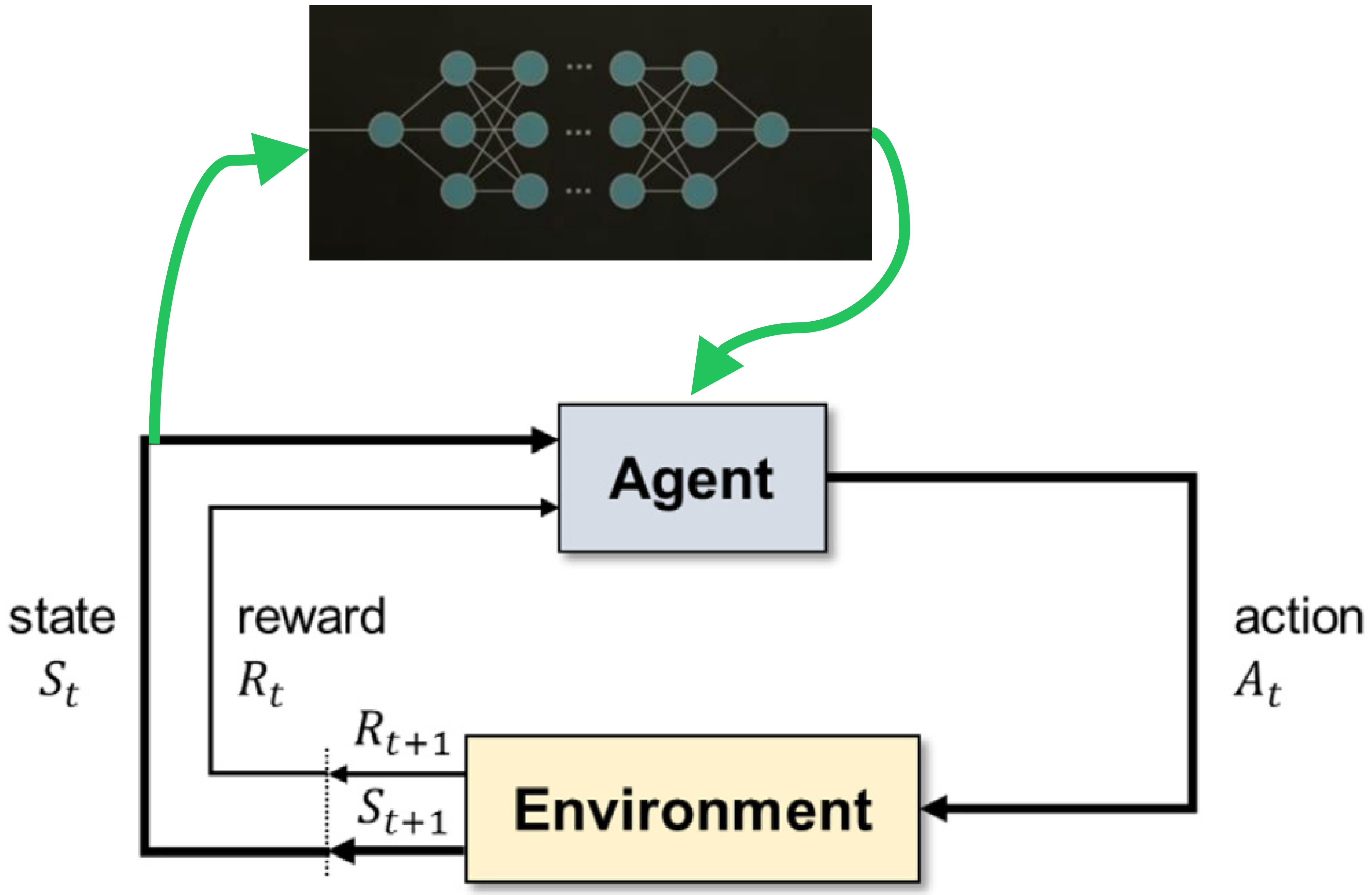
- **Stability:**
  - learning requires data, might bring process to unstable regions
  - choice of meta-parameters to get reliable convergence
- **Sample efficiency:**
  - iterations required to achieve convergence.
  - if policy brings the process to a poor space, it might not recover.
- **Causality**
- **Assignation of rewards**
- **Exploitation vs exploration:** online performance vs information aquisition
- **Types of state variables:** physical? interpretation
- **Value function approximation:** for parameter estimation
- **Episodic vs infinite horizon:** choice of algorithm
- **Continuous vs discrete:** choice of algorithm
- **Stochastic vs deterministic:** policies, environments

Shin, J., Badgwell, T. A., Liu, K.-H., & Lee, J. H. (2019). Reinforcement Learning – Overview of recent progress and implications for process control. *Computers & Chemical Engineering*, *127*, 282–294.

DEEP REINFORCEMENT LEARNING

MAIN CONCEPT

- The use of Deep Neural Networks (DNNs) to approximate the value (e.g. probabilities) and policy functions

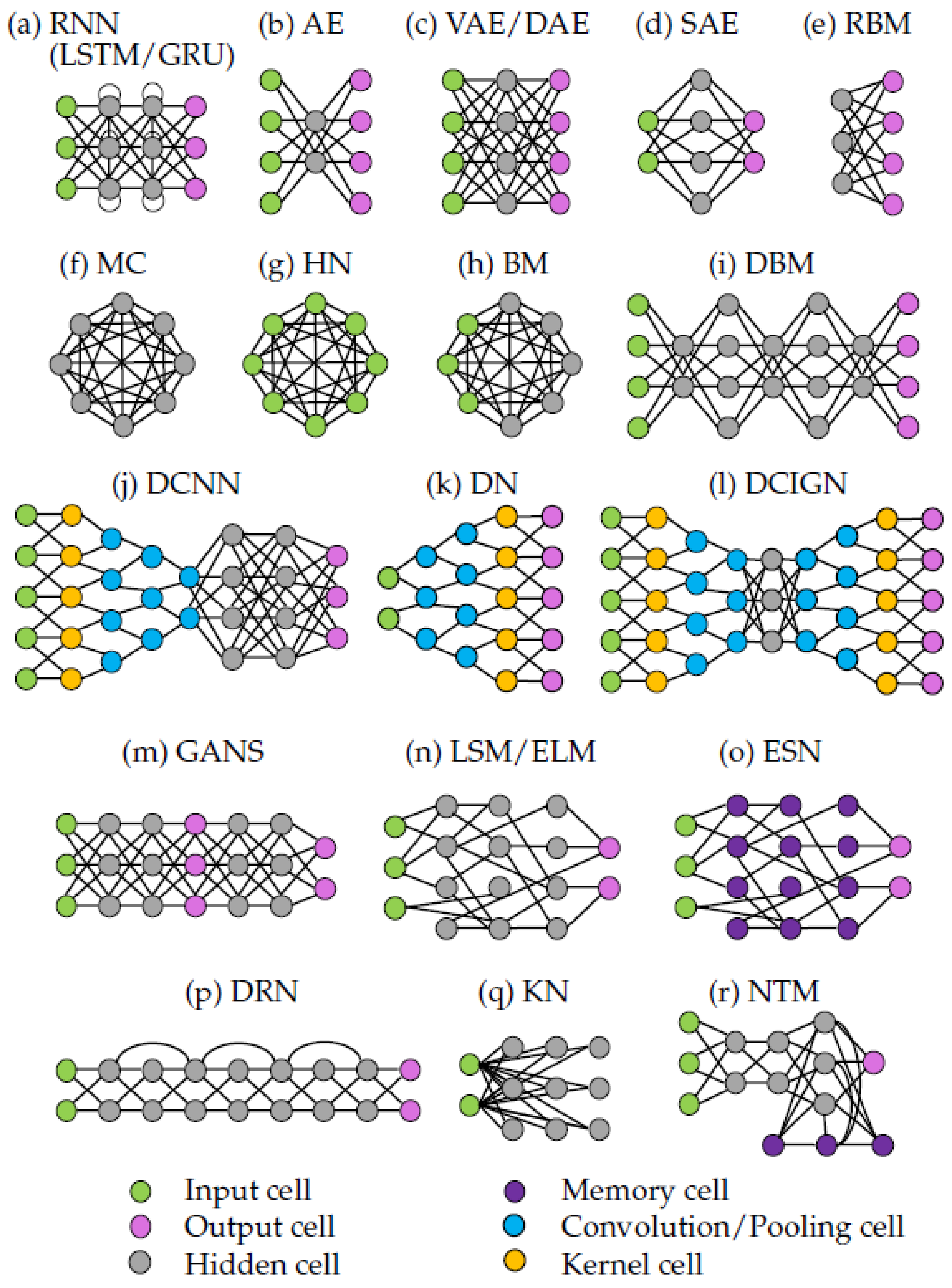


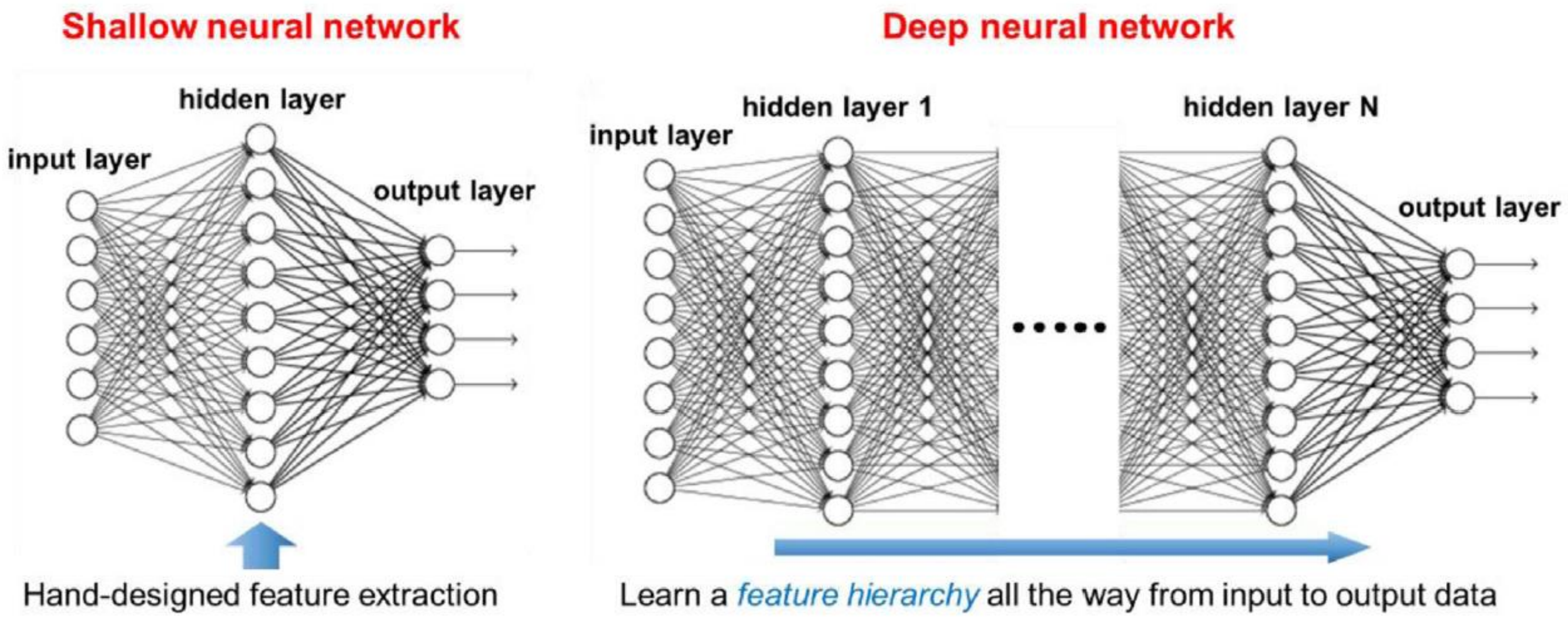
Brunton, Steven and Kutz, Nathan. (2017) *Data Driven Science & Engineering. Machine Learning, Dynamical Systems and Control*

Shin, J., Badgwell, T. A., Liu, K.-H., & Lee, J. H. (2019). Reinforcement Learning – Overview of recent progress and implications for process control. *Computers & Chemical Engineering*, *127*, 282–294.

DEEP REINFORCEMENT LEARNING

NEURAL NETWORKS



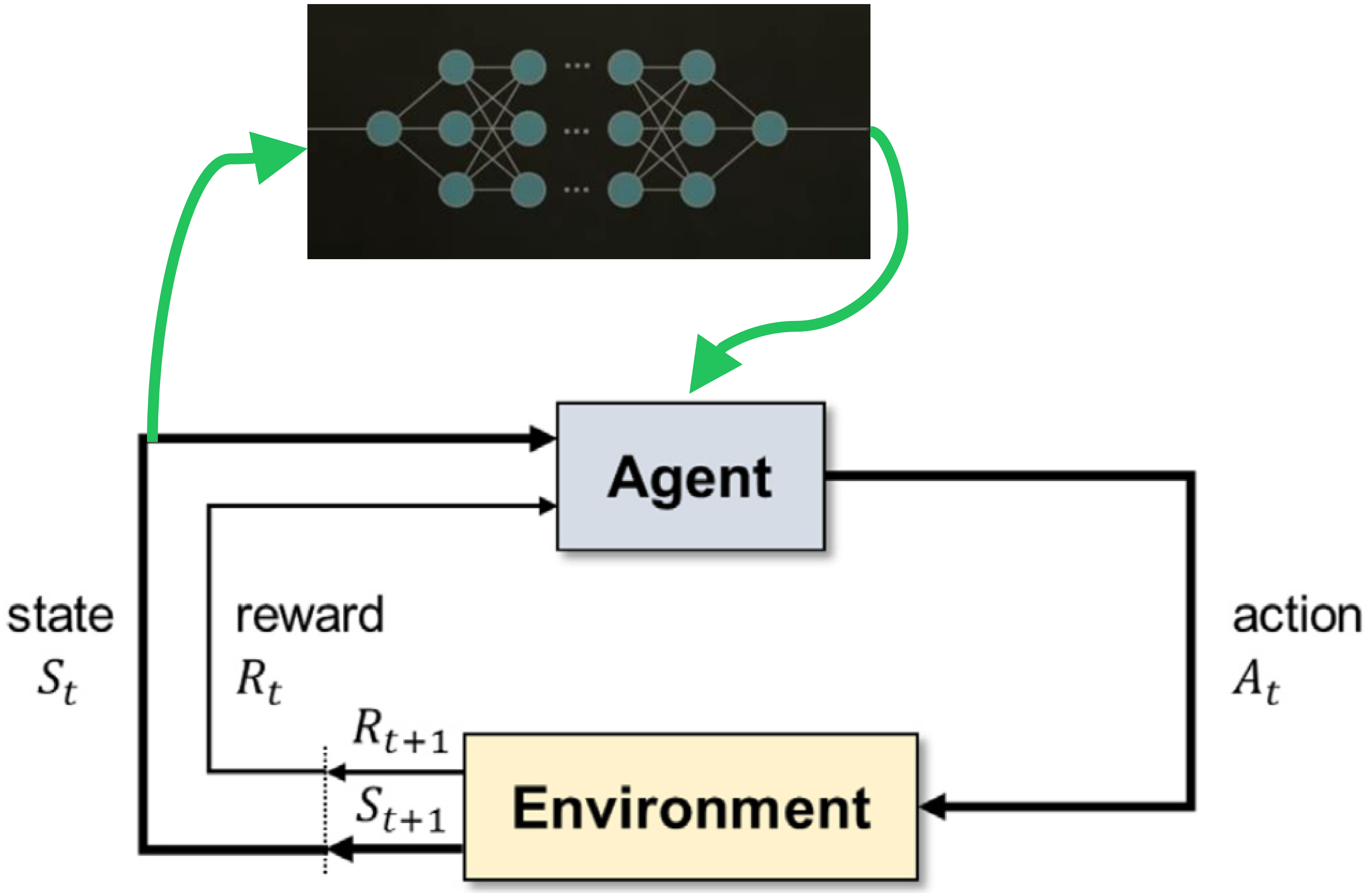


Brunton, Steven and Kutz, Nathan. (2017) *Data Driven Science & Engineering. Machine Learning, Dynamical Systems and Control*
Shin, J., Badgwell, T. A., Liu, K.-H., & Lee, J. H. (2019). Reinforcement Learning – Overview of recent progress and implications for process control. *Computers & Chemical Engineering*, *127*, 282–294.

DEEP REINFORCEMENT LEARNING

CHALLENGES FOR USE OF NEURAL NETWORKS IN RL

- Overfitting:
  - There may be too many available degrees of freedom
  - Need to crossvalidate data.
- NN rely on exploration
- NN are in general not generalizable
- States and behavior may not be interpretable
- Estimates may be noisy
- How to incorporate physical knowledge?
  - Do not disregard what we (partially) know about the system.



Shin, J., Badgwell, T. A., Liu, K.-H., & Lee, J. H. (2019). Reinforcement Learning – Overview of recent progress and implications for process control. *Computers & Chemical Engineering*, *127*, 282–294.

SELF-LEARNING CONTROLLERS IN CONTRAST TO ADAPTIVE CONTROLLERS

With focus on Model-Free Reinforcement Learning and Model Reference Adaptive Control

	SELF-LEARNING CONTROLLERS (RL)	ADAPTIVE CONTROLLERS (MRAC)
Underlying system assumption	Markov decision process	Fixed structure of the process (transfer function, state-space)
Goal(s)	Win reward	Adapt parameters of controller, minimize error
Modeled component	Value function or policy	Process and/or controller
Model learning paradigm	Model learned from trial and error (simulation or real process)	Given structure, calculation of parameters given system response
Exploration/exploitation	Simultaneous	Exploration to get model, exploitaton thereafter
Feedback	Value function or policy	Error, y_{model}-y_{system}
Stability	Closed-loop stability not considered	Stability analysis; proofs
Failure tolerance	Failure is necessary for learning	Failure is not tolerated

Shin, J., Badgwell, T. A., Liu, K.-H., & Lee, J. H. (2019). Reinforcement Learning – Overview of recent progress and implications for process control. *Computers & Chemical Engineering*, *127*, 282–294.

FINAL COMMENT

WHAT DRIVES IMPLEMENTATION?

REQUIRED EFFORT

- Implementation
- Use
- Maintenance

CONFIDENCE IN THE CONTROLLER

- Does it fulfill the control objectives?

COST

SELF-LEARNING CONTROLLERS IN CONTRAST TO ADAPTIVE CONTROLLERS

TRIAL LECTURE

Thank you for your attention!