EE550
Computational Biology

Week 14 Course Notes

Instructor: Bilge Karaçalı, PhD

Topics

- Regulation of gene transcription
  - Regulation in a biomolecular network
  - Primary regulation mechanisms
    - Autoregulation
    - Feed-forward loops

Regulation in a Biomolecular Network

- Selection promotes
  - efficiency
    - Cells operate in an environment with limited resources
    - The resources must be spent on supplying the mechanisms that are of higher priority than others
  - adaptability
    - The extracellular environment and the conditions it imposes on the cells change in time
    - The cells must be able to respond to these changes by adjusting their priorities
  - rapid response
    - The quicker the cells adapt to the changing conditions the better for maintaining efficiency
  - robustness
    - At the same time, the cellular operations must also be shielded from random fluctuations in the environmental conditions

Regulation in a Biomolecular Network

- Tightly controlled regulation of gene transcription is a result of natural selection
  - Genetic variability produces diverse organisms with slightly different regulatory skills
  - The organisms possessing the regulatory skills that endow them with a higher fitness undergo positive selection
- Several primary regulatory mechanisms for gene transcription are present “conspicuously” across different species
  - Autoregulation
  - Feed-forward loop

Autoregulation

- Regulation of a gene Y by another gene X is indicated by an edge in the network graph between the nodes X and Y
  - If the regulation is activation, the edge is an arrow

    $$X \rightarrow Y$$

  - Conversely if X represses Y, the edge ends with a line stop

    $$X \dashv Y$$

- In autogenous regulation, a gene's product acts as its own transcription factor
  - Such cases are indicated by a self-edge
  - The edge can be activation or repression as any other edge in the regulatory network

Production Rates of Autogenously Regulated Genes

- Positive autoregulation
  - This situation refers to the case where the gene's own protein product acts as a transcription factor activating its expression
  - The input function governing a positively autoregulated gene X is given by the usual Hill function

$$\text{rate of production of } X = f([X^*]) = \frac{\beta [X^*]^n}{\kappa^n + [X^*]^n}$$

- Negative autoregulation
  - The gene's product represses its expression
  - The input function is given by

$$\text{rate of production of } X = f([X^*]) = \frac{\beta}{1 + \left(\frac{[X^*]}{\kappa}\right)^n}$$

Production Rates of Autogenously Regulated Genes

- Note that these functions do not characterize a static system
  - By definition, a positive production rate increases $[X]$
  - Since we assume that the signal S_X is always present, all $[X]$ is readily transformed into the active state $[X^*]$
  - Thus, the concentration does not remain on the initial value of $[X^*]$
  - Same thing happens with a negative production rate
- Instead, they represent instantaneous production rates
  - That vary with time
- Consequently, the system becomes a dynamic one

Transients of Autoregulation

- Dynamic response in negative autoregulation – kinetic modelling
  - The equation governing the temporal variation of a gene product is

$$\frac{d}{dt}([X])(t) = f\left(([X^*])(t)\right) - \alpha([X])(t)$$

where the production rate follows the relationship

$$f\left(([X^*])(t)\right) = \frac{\beta}{1 + \left(\frac{([X^*])(t)}{\kappa}\right)^n}$$

  - Assuming S_X is always present allows $[X^*] = [X]$ and produces

$$\frac{d}{dt}([X])(t) = \frac{\beta}{1 + \left(\frac{S_X(t)([X])(t)}{\kappa}\right)^n} - \alpha([X])(t)$$

Transients of Autoregulation

- Dynamic response in negative autoregulation – approximate analysis
  - However, before solving the dynamic equation above, it is possible to predict how the system will respond using the logic approximation to the Hill function in repression
  - The logic approximation for repression provides

$$\frac{d}{dt}([X])(t) \simeq \beta \mathbf{1}(S_X(t)([X])(t) < \kappa) - \alpha([X])(t)$$

- When $([X])(t) < \kappa$, $[X]$ is simply regulated with

$$\frac{d}{dt}([X])(t) = \beta - \alpha([X])(t)$$

resulting in an exponential rise towards the β/α with $T_{1/2} = \log(2)/\alpha$

- When $([X])(t) > \kappa$, however, the production ceases and exponential decay starts

➔ stability around $[X] = \kappa$

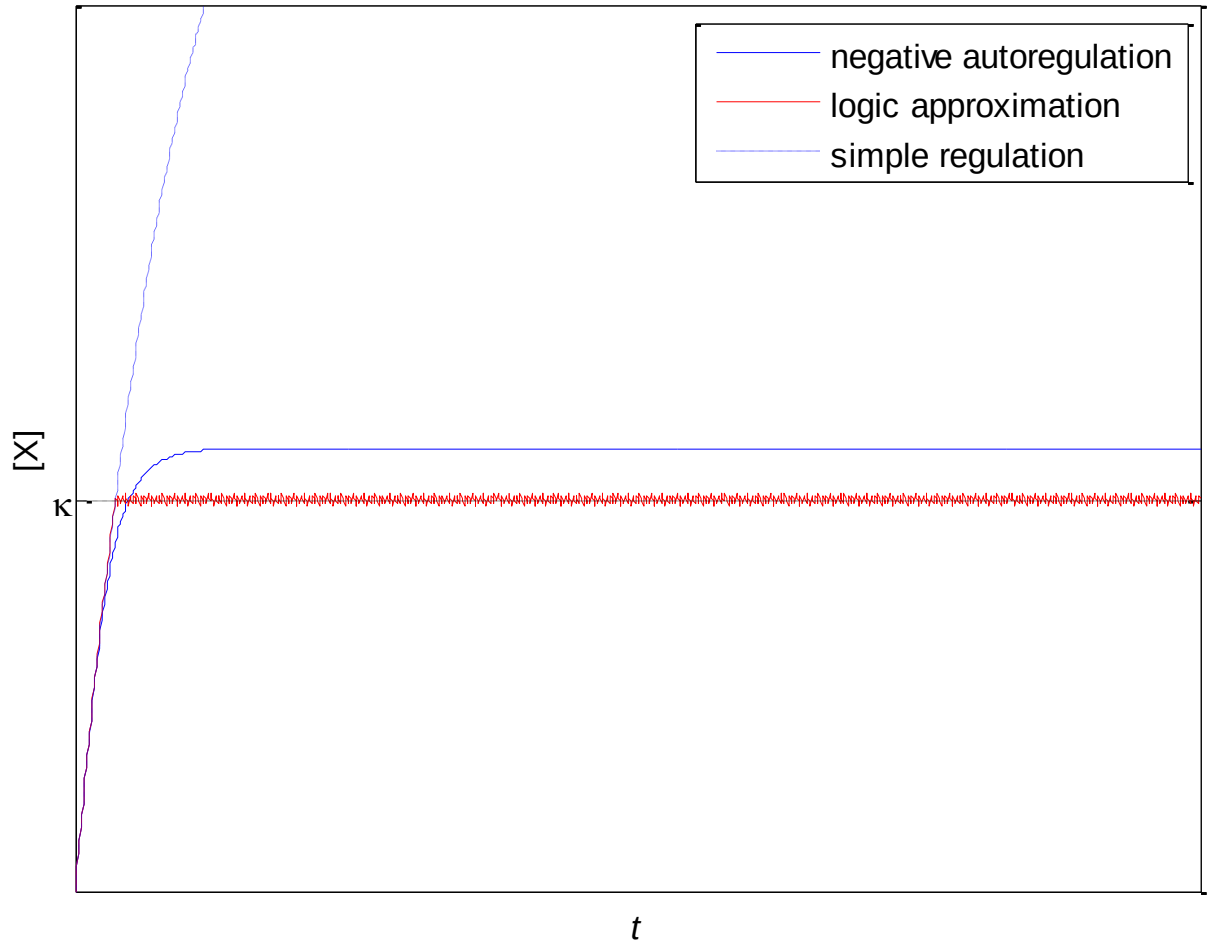
Transients of Autoregulation

- Dynamic response in negative autoregulation – numerical analysis
  - The ordinary differential equation is to be solved numerically using Euler's method that provides

$$([X])(t+\Delta t) \simeq ([X])(t) + \Delta t \frac{d}{dt}([X])(t)$$

  - Thus, starting at $t = 0$ with $([X])(t) = [X]_0$ and for $\Delta t \ll 1$
    - Calculate $\frac{d}{dt}([X])(t)$ using the formula in the differential equation
    - Set $([X])(t+\Delta t) = ([X])(t) + \Delta t \frac{d}{dt}([X])(t)$
    - Let $t \leftarrow t + \Delta t$
    - Repeat until convergence

Dynamic Response in Negative Autoregulation



Dynamic Response in Negative Autoregulation

- Negative autoregulation alters the response time of gene activation
  - The time to half steady state (around $\kappa < \beta/\alpha$) is given by

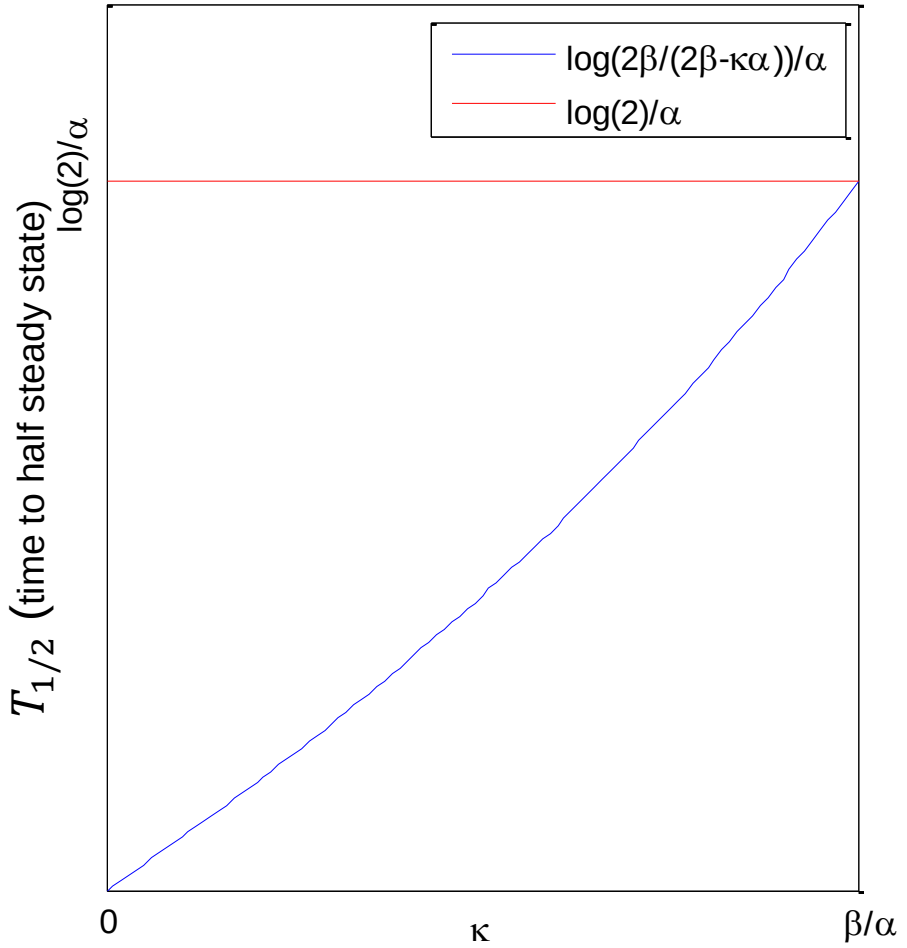
$$\frac{\kappa}{2} = \frac{\beta}{\alpha}\left(1 - e^{-\alpha T_{1/2}}\right)$$

$$\Rightarrow T_{1/2} = \log\left(\frac{2\beta}{2\beta - \kappa\alpha}\right)/\alpha$$

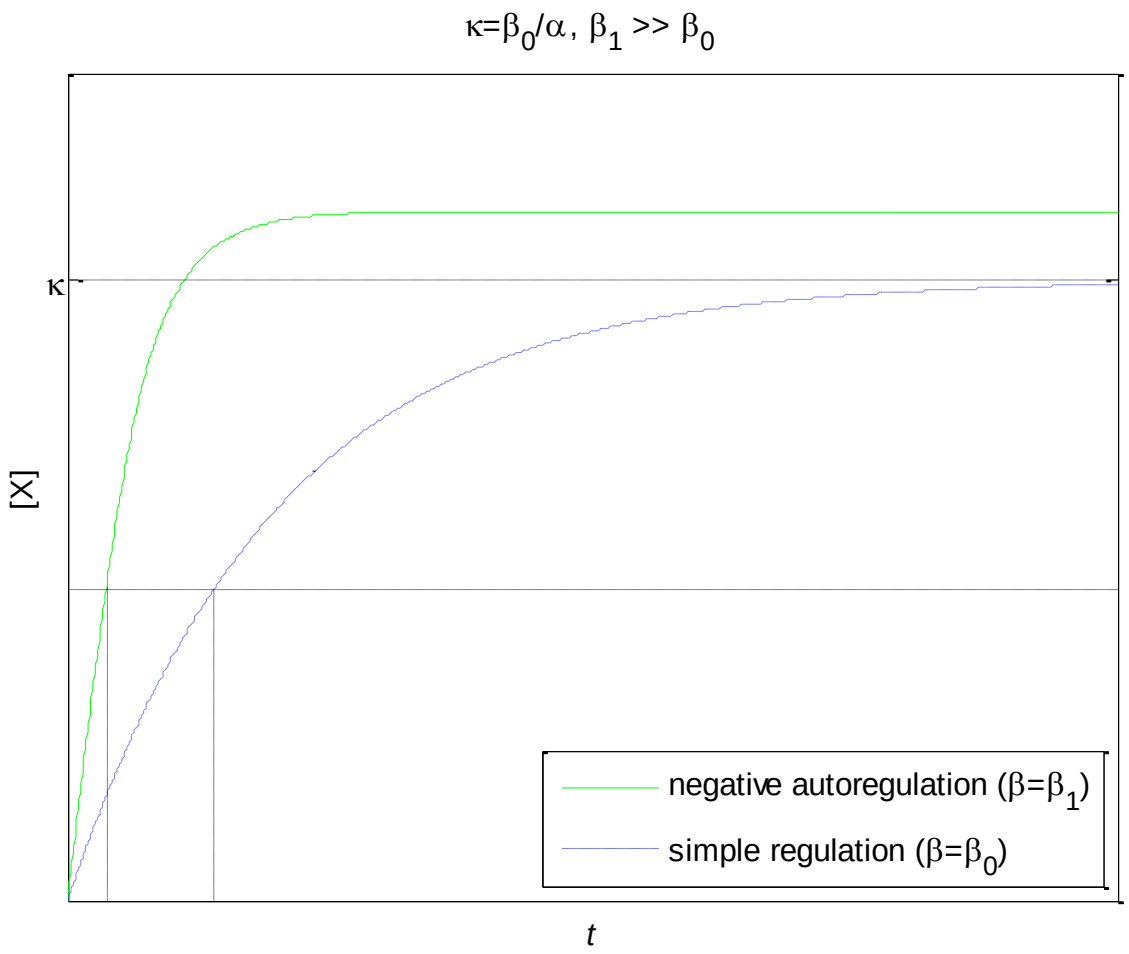
  - Compare that to $\log(2)/\alpha$ in a simple regulation alternative with

$$\beta' = \kappa\alpha$$

  that achieves the same steady state level κ



Dynamic Response in Negative Autoregulation



Dynamic Response in Negative Autoregulation

- In addition to a faster rise, negative autoregulation provides robustness in gene expression against random fluctuations in the production rate β
  - Twin bacterial cells show variations in their respective production rates
    - Differences in capacity leads to variations from a few percents to tens
  - The production rate also varies in time due to random effects
  - The steady state level in **simple regulation** is directly affected by the **production rate fluctuations**
    - Note that the steady state level is given by β/α
  - The **threshold** κ on the other hand is a biochemical property of the input function, and is much more **stable across individuals and in time**

  ➔ The steady state expression level in negative autoregulation is stable even though the production rate may fluctuate

Dynamic Response in Positive Autoregulation

- In positive autoregulation, a gene product improves the expression rate of its own gene
  - Kinetic modelling: Using the Hill function and positive autoregulation transient equation provides

$$\frac{d}{dt}([X])(t) = \frac{\beta([X])^n(t)}{\kappa^n + ([X])^n(t)} - \alpha([X])(t)$$

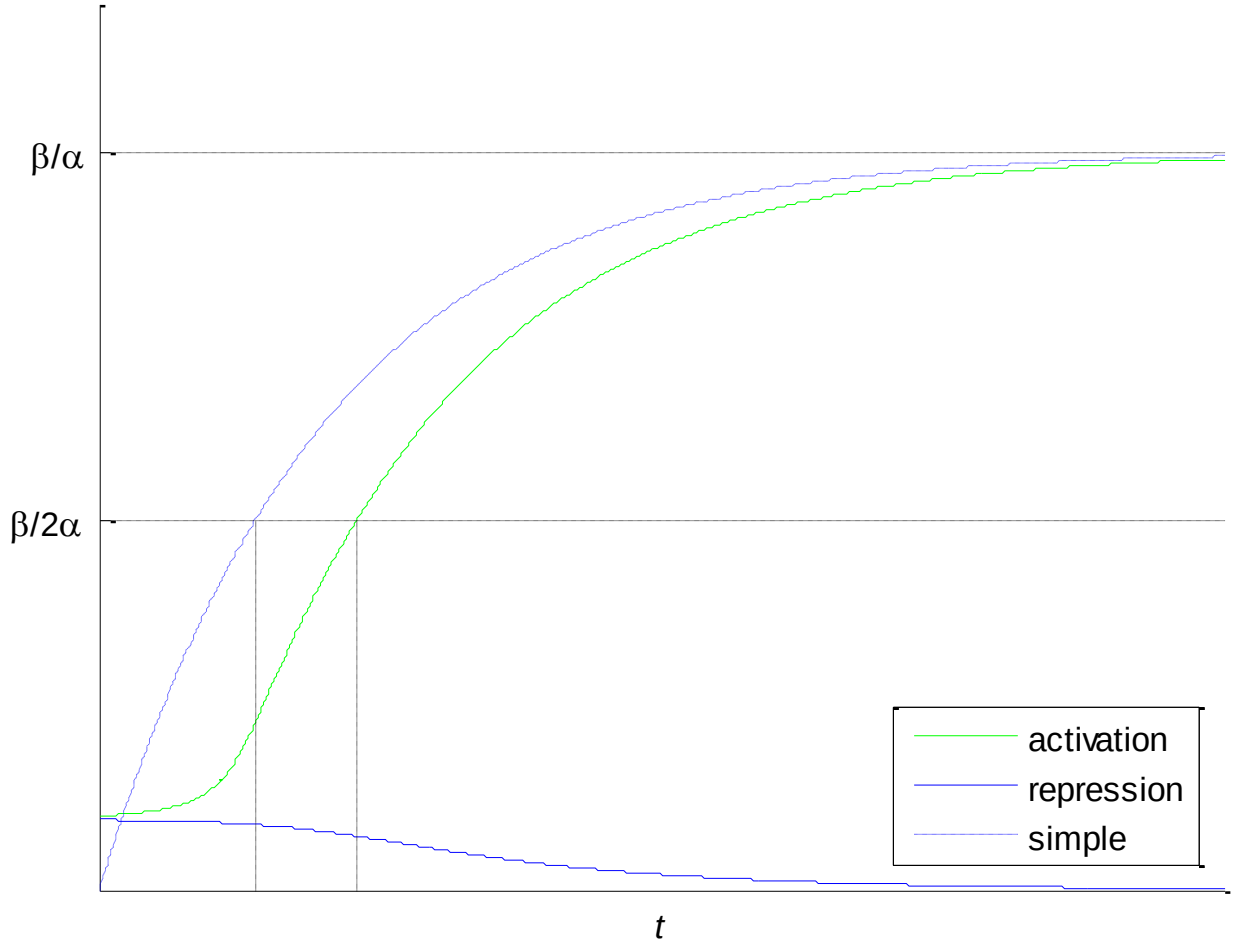
  - The logic function approximation leads to

$$\frac{d}{dt}([X])(t) = \beta \mathbf{1}\big(([X])(t) > \kappa\big) - \alpha([X])(t)$$

  - This suggests that
    - If $[X]$ is low, it stays low
    - If $[X]$ is high (at the steady state level), it stays high

    ➔ Bi-stability in gene expression

Dynamic Response in Positive Autoregulation

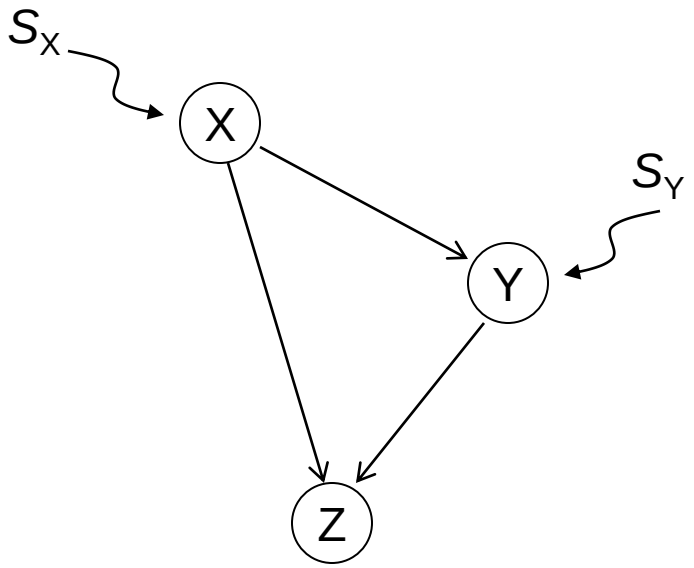


Dynamic Response in Positive Autoregulation

- Bi-stability represents permanent decision making
  - Once a gene is activated by **some other regulatory means**, it remains active
  - Such decisions are frequently made in the early stages of development
    - In cellular differentiation, identical stem cells are set to grow into different tissues and organs
  - The state of positively autoregulated genes thus represents a bar-code for the cell's identity
    - This set would naturally include the genes that are governed by positive autoregulation cascades
- Delay represents timing priorities
  - The genes that produce proteins required at a specific stage of a process are delayed to wait for the completion of the preceding stages

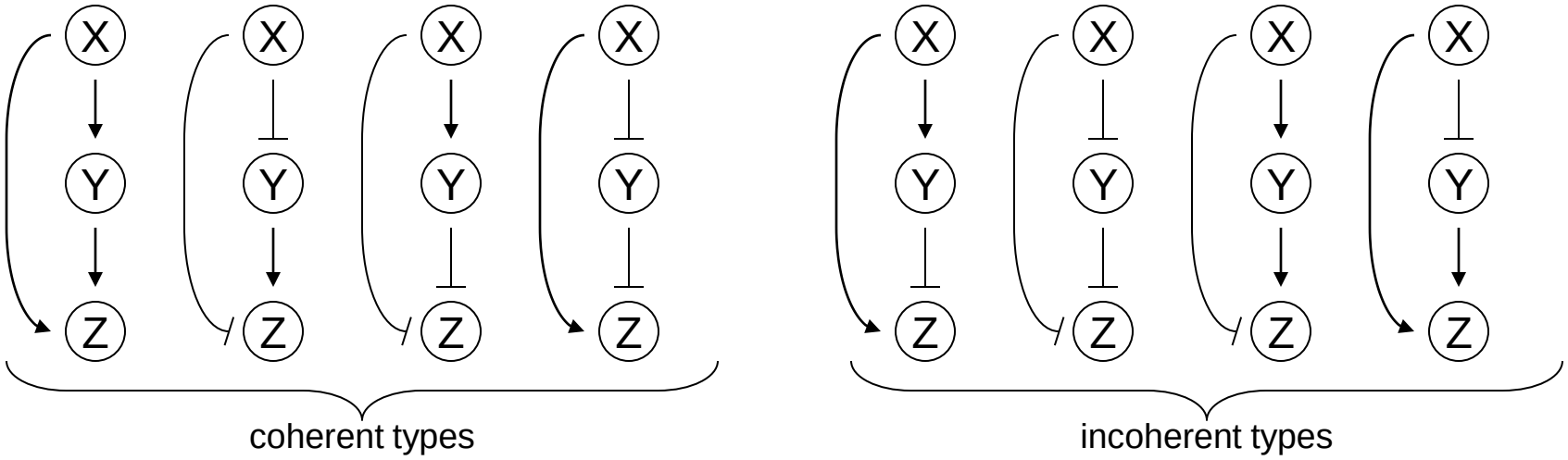
The Feed-Forward Loop

- Another common regulation mechanism in gene transcription networks is the feed-forward loop
  - Consists of three nodes
    - First node regulates the other two
    - The second is regulated by the first and regulates the third
    - The third is regulated jointly by the first two
  - The regulatory mechanism consists of the effects of the signals to the first two nodes onto the expression of the third



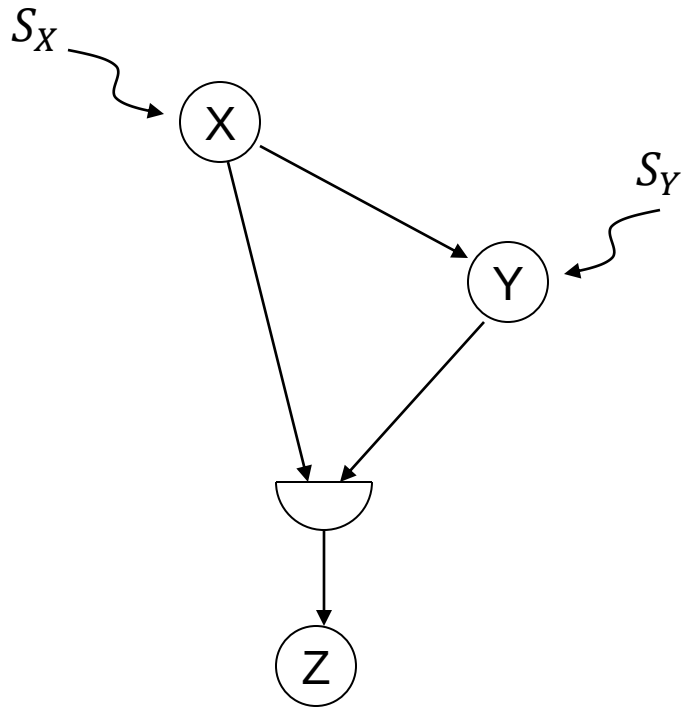
The Feed-Forward Loop

- Depending on the functionality on the edges, the regulatory function of the feed-forward loop changes
  - Coherent type: The regulatory effects of both paths are the same
  - Incoherent type: The regulatory effects conflict with each other
- An additional control mechanism is in the integration of the regulatory inputs from both paths at the third node
  - AND or OR (SUM is not particularly interesting; it merely provides a linear combination of both paths)



Coherent Type-1 Feed Forward Loop with AND Integration

- Characteristics of the regulatory mechanism:
  - All regulatory edges are activations
    - X → Y with κ_{XY}
    - X → Z with κ_{XZ}
    - Y → Z with κ_{YZ}
  - Two alternate paths with the same regulatory function on gene Z
  - Activation signals from both paths are required to express Z
    - AND integration



Coherent Type-1 Feed Forward Loop with AND Integration

- Kinetic model
  - Premises:
    - S_Y is present, S_X becomes present at time $t = 0$
    - $[X]$ is constant at steady state, $[Y]$ and $[Z]$ are initially zero

$$([X])(0^-) = [X]_{st},\ ([Y])(0^-) = ([Z])(0^-) = 0$$

  - X → Y:
    - Simply regulated
    - The expression of Y begins at time $t = 0$ when X is activated into X*
    - The dynamics are governed by

$$\frac{d}{dt}([Y])(t) = \beta_Y \cdot S_X(t) - \alpha_Y([Y])(t)$$

  - X AND Y → Z:
    - Both are simply regulated as well
    - Since $[X] > \kappa_{XZ}$ already, the expression of Z begins after $[Y]$ crosses the threshold κ_{YZ}

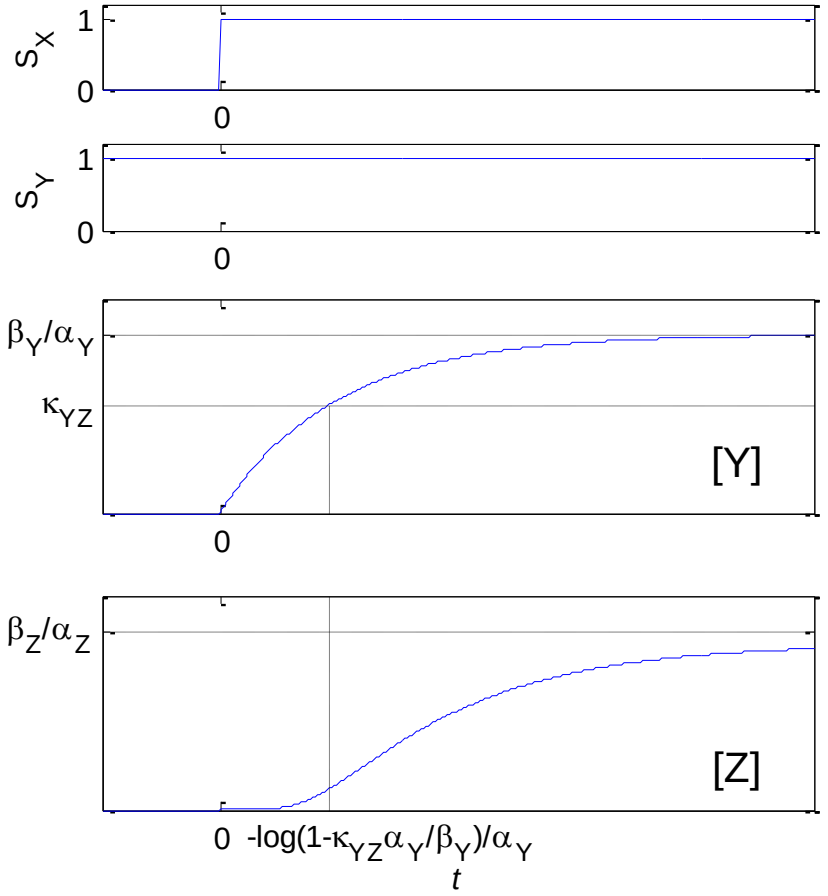
$$\text{production rate of } Z = \frac{\beta_Z [Y]^{n_{YZ}} \cdot S_Y(t)}{\kappa_{YZ}^{n_{YZ}} + [Y]^{n_{YZ}}} \cdot S_X(t)$$

    - The dynamics are thus governed by

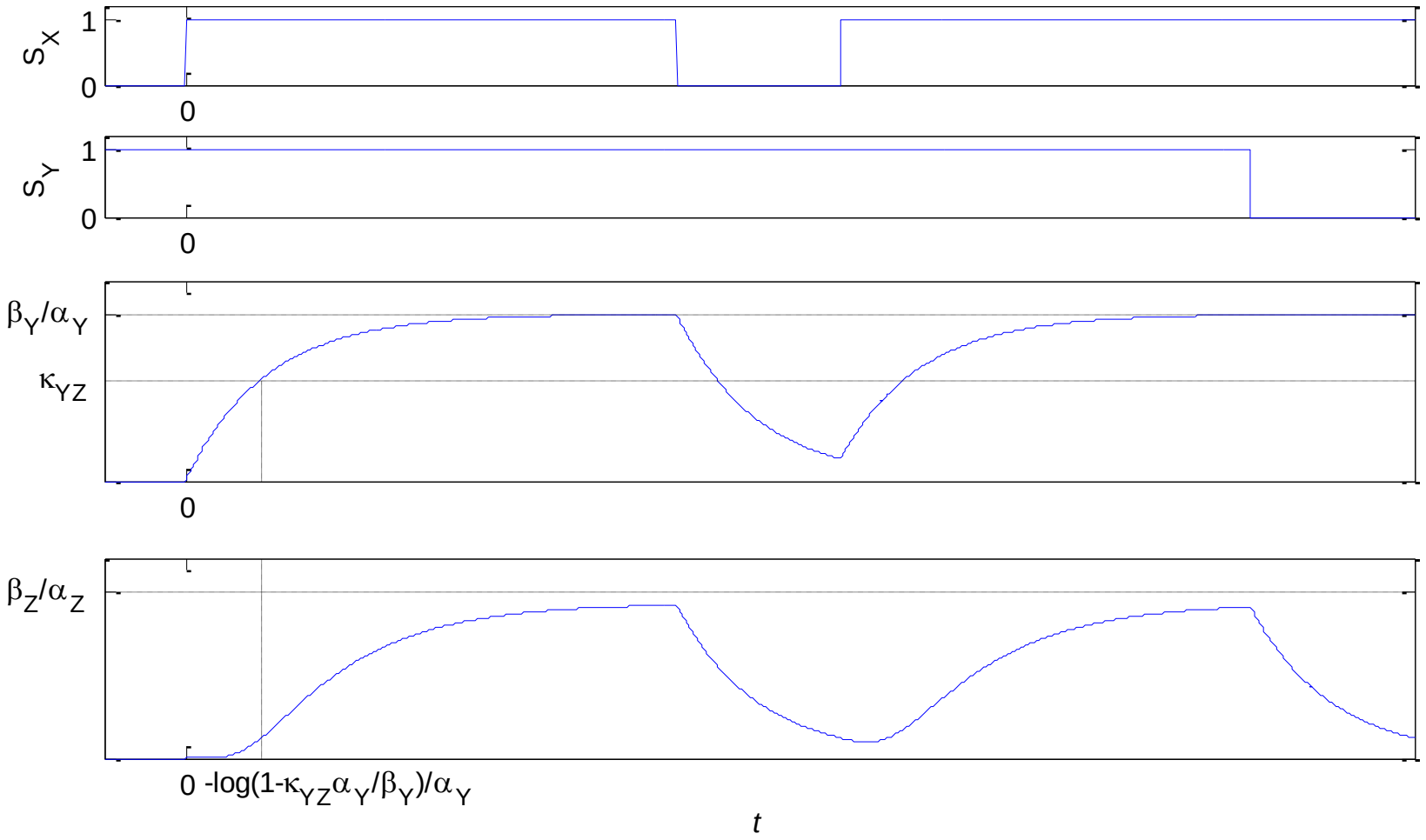
$$\frac{d}{dt}([Z])(t) = \frac{\beta_Z([Y])^{n_{YZ}}(t) \cdot S_Y(t)}{\kappa_{YZ}^{n_{YZ}} + ([Y])^{n_{YZ}}(t)} \cdot S_X(t) - \alpha_Z([Z])(t)$$

Coherent Type-1 Feed Forward Loop with AND Integration

- Dynamic evaluation:
  - The expression of Y is turned on when S_X is switched on at time $t = 0$
    - The activated transcription factor X* binds the promoters of Y and Z
    - $[Y]$ (and hence $[Y^*]$) starts to build up toward its steady state value following an exponential rise
  - As activated $[Y]$ crosses the threshold κ_{YZ}, it starts binding the promoter of Z in large amounts, initiating the transcription of Z



Coherent Type-1 Feed Forward Loop with AND Integration



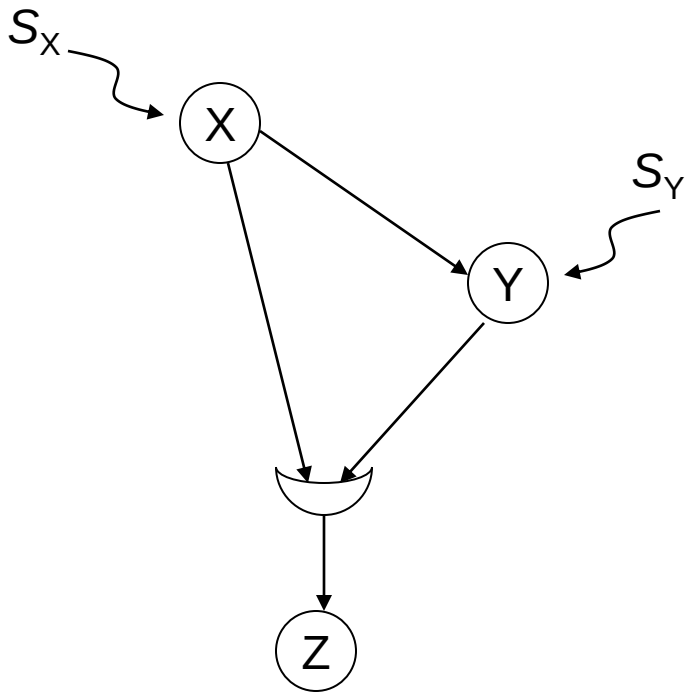
Coherent Type-1 Feed Forward Loop with AND Integration

- The coherent type-1 FFL network element with AND integration acts as a **sign-sensitive delay element**
  - A delay of $-\log(1-\kappa_{YZ}\alpha_Y/\beta_Y)/\alpha_Y$ is present at the initiation of the Z transcription
  - No such delay exists when either S_X or S_Y is turned off
- This mechanism protects the gene transcription against spurious activations
  - Spurious activations cause the cell both energy and raw materials
    - Hence, there is no reason to start Z transcription unless it really is required
  - In C1-FFL w/ AND, the Z transcription is activated only when the signal S_X persists for a sufficiently long time
    - Indicating that Z transcription really is required

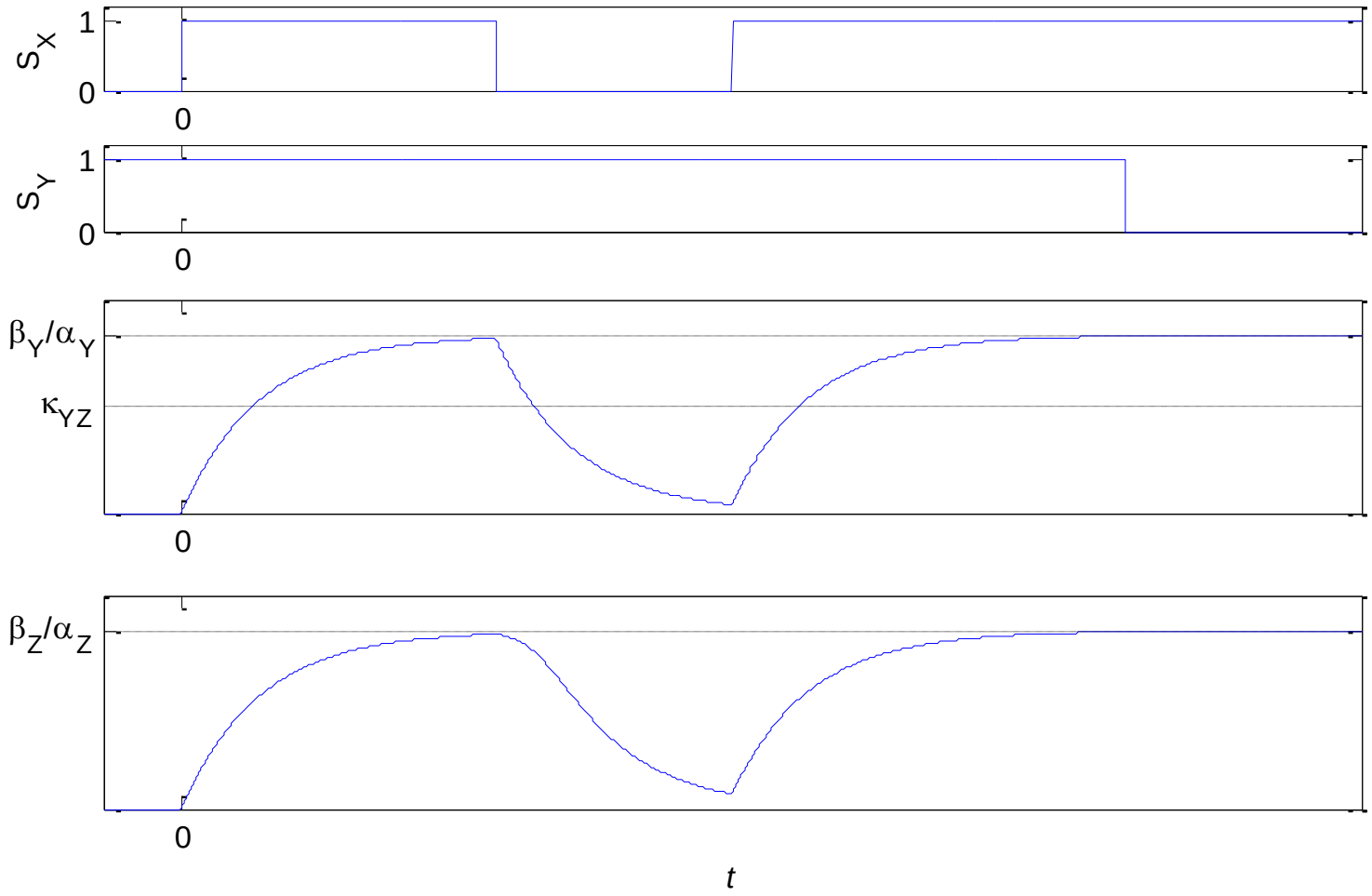
Coherent Type-1 Feed Forward Loop with OR Integration

- Premises:
  - S_Y is present
  - S_X becomes present at time $t = 0$
  - $[X]$ is constant at steady state, $[Y]$ and $[Z]$ are initially zero

    $([X])(0^-) = [X]_{st}$, $([Y])(0^-) = ([Z])(0^-) = 0$
- Dynamic evaluation:
  - As soon as S_X becomes present, the transcriptions of both Y and Z begin
    - Only one of X or Y is sufficient to initiate Z transcription
  - When S_X is turned off again, the transcription of Y ceases and the $[Y]$ level drop exponentially
  - The transcription of Z ceases only when the $[Y]$ level is below κ_{YZ}



Coherent Type-1 Feed Forward Loop with OR Integration

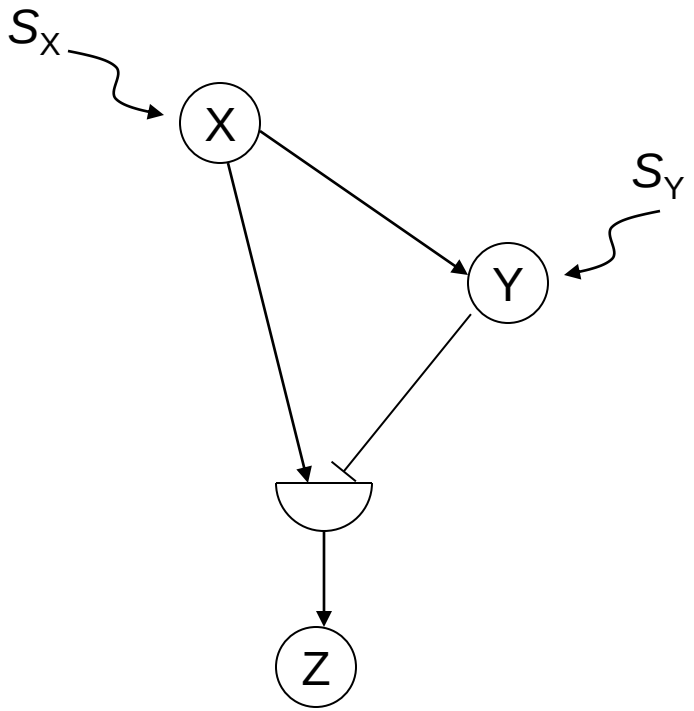


Coherent Type-1 Feed Forward Loop with OR Integration

- The coherent type-1 FFL network element with OR integration also acts as a **sign-sensitive delay element**
- However, in contrast with the same element with AND integration, the delay is observed at the cessation of the gene transcription
- This mechanism thus protects the transcription of gene Z against spurious loss of signal S_X
  - The process requiring Z should not be shut off accidentally due to a noise in S_X
    - Accidental shut-off's are also costly

Incoherent Type-1 Feed Forward Loop with AND Integration

- In this feed forward loop, the two paths are antagonistic
  - X directly activates Z
  - X also represses Z indirectly through Y
- Dynamic evaluation:
  - Premises:
    - S_Y is present
    - S_X becomes present at time $t = 0$
    - $[X]$ is constant at steady state, $[Y]$ and $[Z]$ are initially zero
  - Immediately as S_X is turned on, the transcriptions of both Y and Z begin following the exponential curve
  - Gradually as $[Y]$ builds up, it crosses the threshold κ_{YZ}, causing Y to repress Z
  - As Z is repressed, $[Z]$ decreases



Incoherent Type-1 Feed Forward Loop with AND Integration

- Kinetic model:
  - With the activation of X → X* at time $t = 0$, $[Y]$ increases via

$$\frac{d}{dt}([Y])(t) = S_X(t)\beta_Y - \alpha_Y([Y])(t)$$

    toward its steady state level $[Y]_{st} = \beta_Y/\alpha_Y$
  - The transcription of Z follows the transient equation

$$\frac{d}{dt}([Z])(t) = \frac{\beta_Z}{1 + \left(\frac{S_Y(t)([Y])(t)}{\kappa_{YZ}}\right)^{n_{YZ}}} S_X(t) - \alpha_Z([Z])(t)$$

  - Initially, $[Z]$ rises according to the exponential curve of simple regulation towards $[Z]_{st} = \beta_Z/\alpha_Z$
  - Around time $t \simeq -\log(1 - \kappa_{YZ}\alpha_Y/\beta_Y)/\alpha_Y$, increasing $[Y]$ starts to repress the Z transcription

Incoherent Type-1 Feed Forward Loop with AND Integration

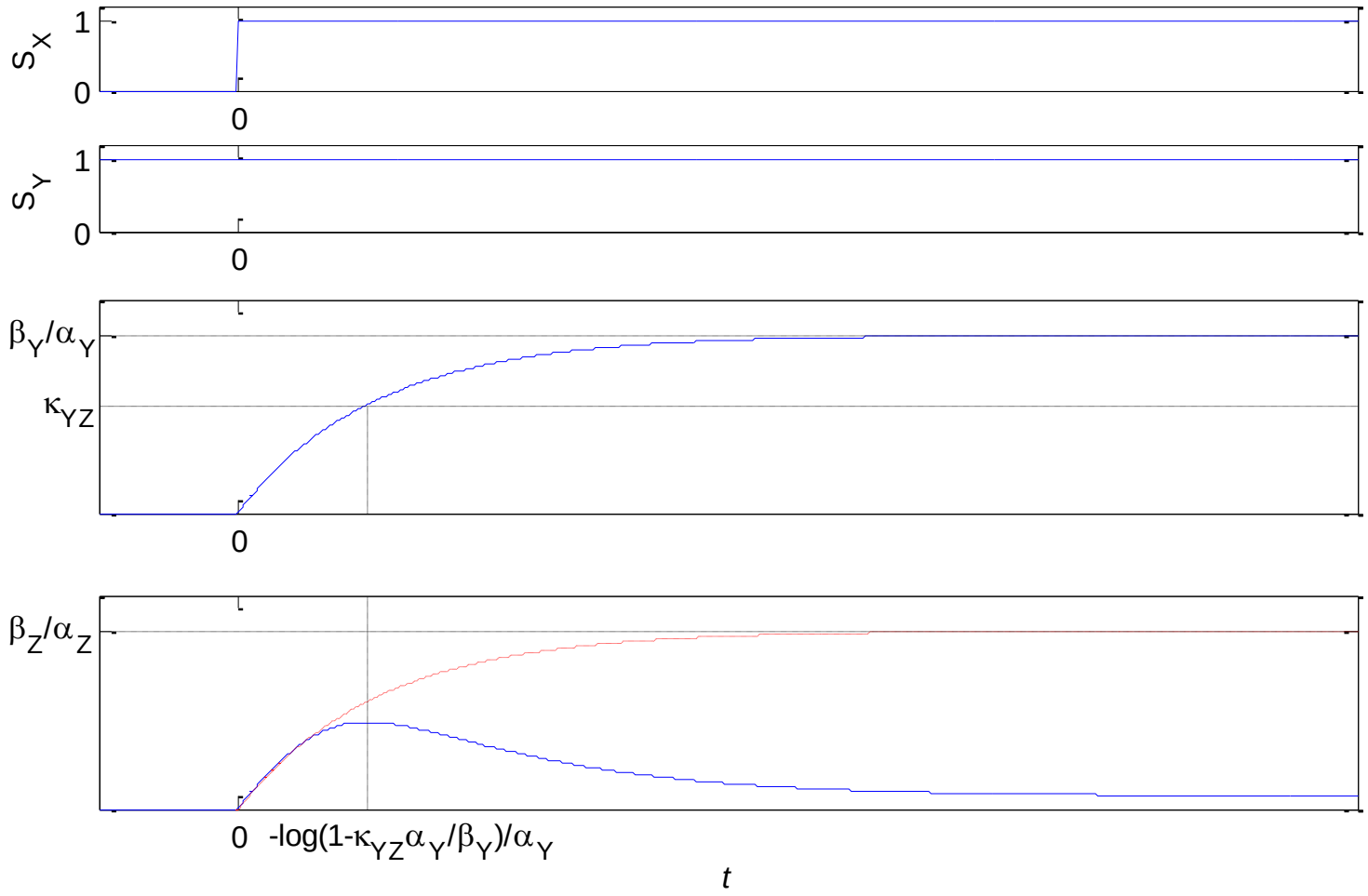
- Kinetic model (continued):
  - Eventually, $[Y]$ attains its steady state level and $[Z]$ decays toward a different steady state level $[Z]'_{st}$

$$[Z]'_{st} = \frac{\beta_Z}{\alpha_Z \left(1 + \left(\frac{[Y]_{st}}{\kappa_{YZ}}\right)^{n_{YZ}}\right)}$$

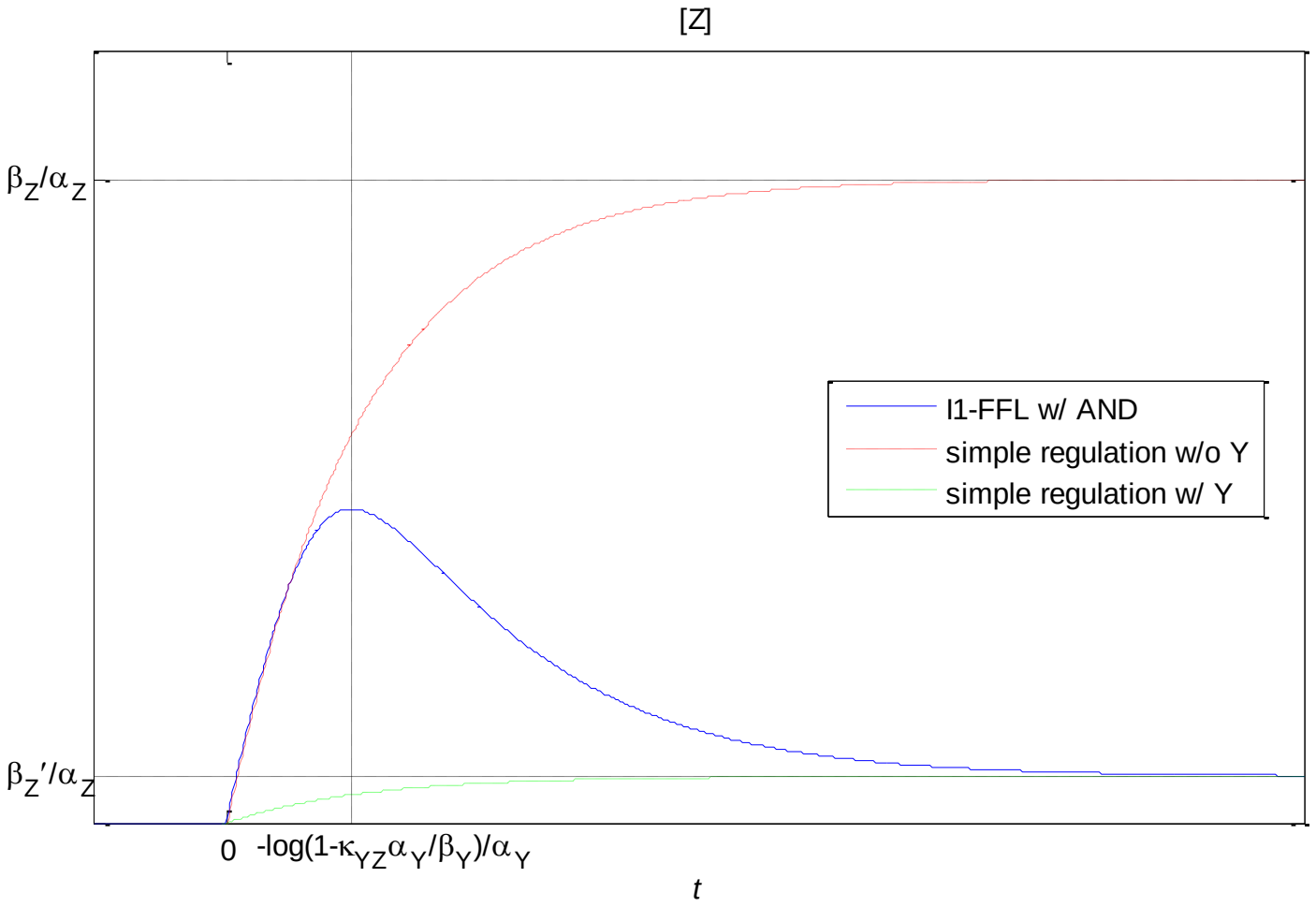
  - The repression coefficient F is defined as the ratio of the two levels:

$$F = \frac{[Z]_{st}}{[Z]'_{st}} = 1 + \left(\frac{[Y]_{st}}{\kappa_{YZ}}\right)^{n_{YZ}}$$

Incoherent Type-1 Feed Forward Loop with AND Integration



Incoherent Type-1 Feed Forward Loop with AND Integration



Incoherent Type-1 Feed Forward Loop with AND Integration

- The incoherent type-1 feed forward loop with AND integration acts as a pulse generator
  - In the absence of repression from Y, Z undergoes a rapid rise towards $[Z]_{st}$
  - Eventually $[Y]$ rises sufficiently and begins to repress $[Z]$
  - Under repression, $[Z]$ declines toward $[Z]'_{st}$
- The response time of $[Z]$ is dramatically improved as well (assuming $[Z]'_{st}$ is the desired steady-state level)
  - Instead of rising towards $[Z]'_{st}$ via simple regulation, $[Z]$ is shot up towards $[Z]_{st} \gg [Z]'_{st}$ and brought back down to $[Z]'_{st}$ later
  - Rise towards $[Z]_{st}$ is much faster than towards $[Z]'_{st}$ via simple regulation and crosses the $[Z]'_{st}$ level much sooner

Summary

- Gene transcription networks are endowed with specific network elements that carry out critical functions
  - Autoregulation
    - Negative autoregulation: Rapid response
    - Positive autoregulation: Delayed response and bi-stability
  - Feed-forward loop
    - C1-FFL: Sign-sensitive delay for protection against spurious signals (with AND integration) and signal losses (with OR integration)
    - I1-FFL: Pulse generation and rapid response
- Such critical network elements are observed “abundantly” in gene transcription networks
- The statistical significance of this “abundance” is crucial to derive a functional understanding of gene transcription regulation