EE550
Computational Biology

Week 13 Course Notes

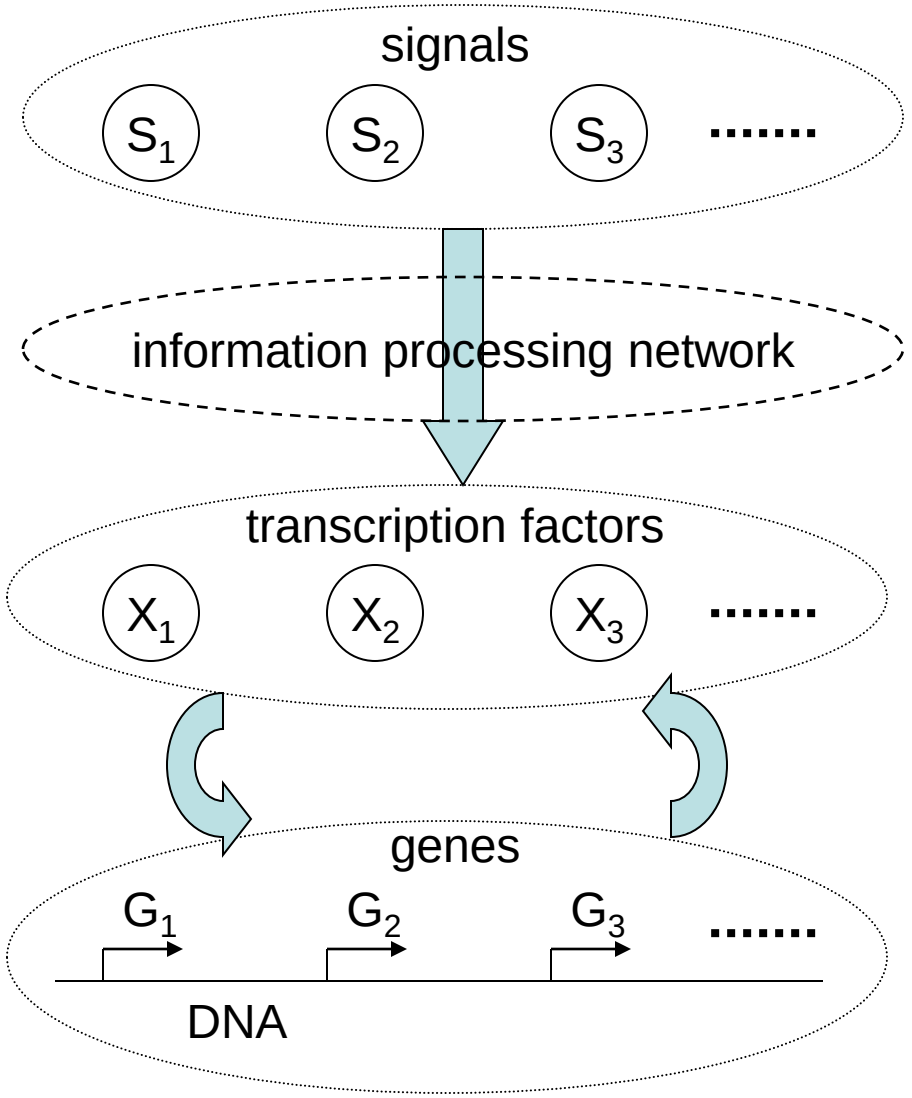
Instructor: Bilge Karaçalı, PhD

Topics

- Gene transcription networks
  - Nodes and edges
  - Activation and repression
  - Production rates
    - Hill function
    - Logic approximation
  - Gene transcription transients

Gene Transcription in a Network

- Extracellular signals are received by the membrane
- These signals activate or inhibit certain transcription factors
- The activation or inhibition of the transcription factors trigger or stop the expression of associated genes
- Some gene products go on to affect the activity of transcription factors or act themselves to induce transcription of additional genes



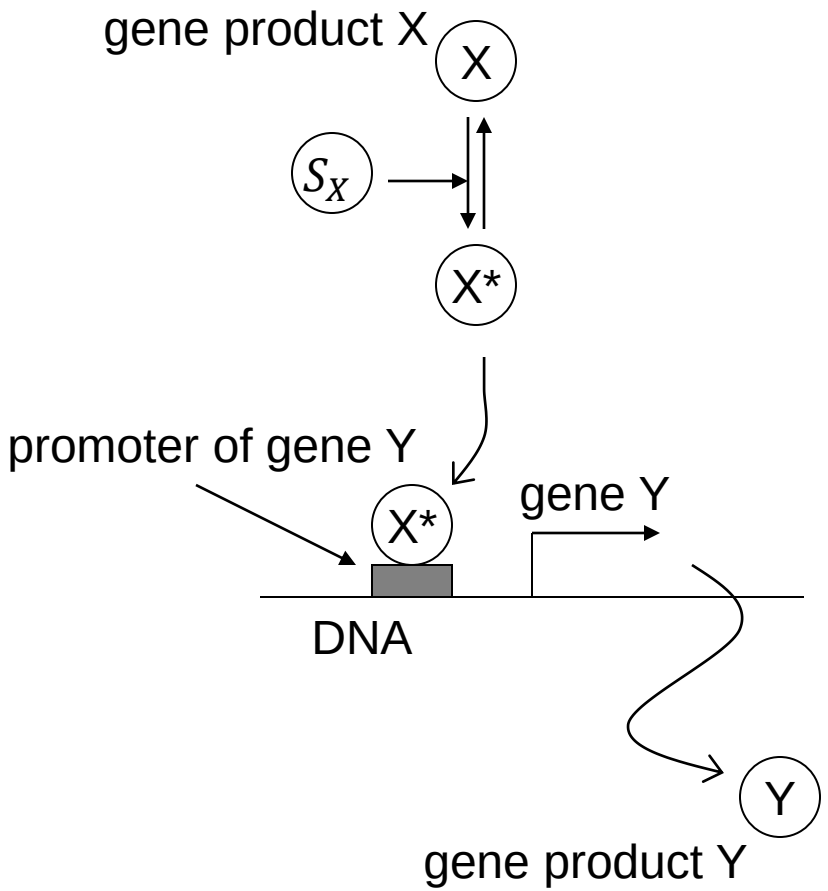
Signal-Based Gene Transcription

- The signal S_X activates the transcription factor X
- The activated transcription factor X* binds the promoter of the gene Y
- Binding of X* to the promoter triggers the expression of gene Y
- The gene product Y is synthesized via the translation of gene Y

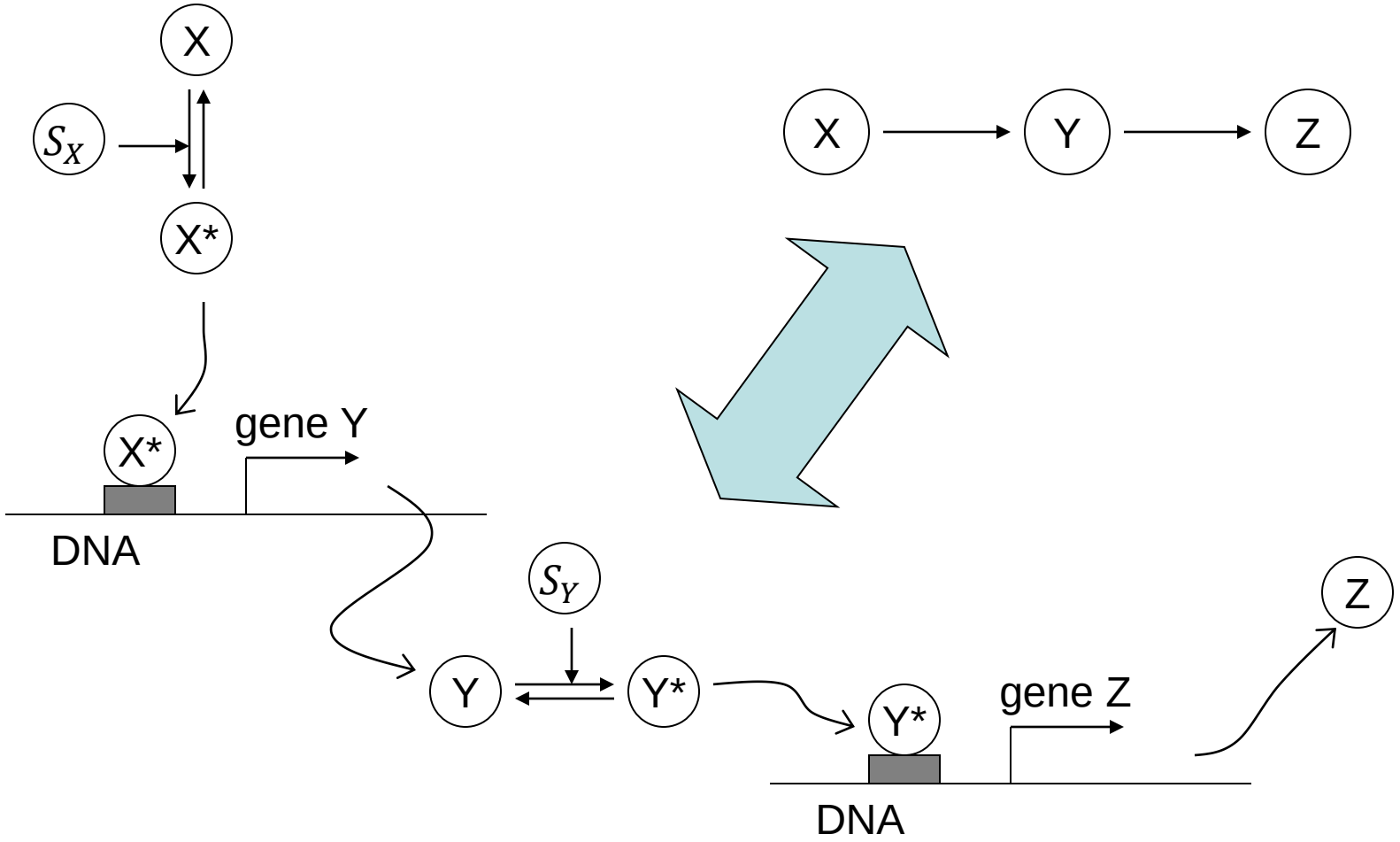
→ **X regulates Y**

Notes:

1. This statement refers to a *topological property* of the transcription network
   → describing the relationship between genes X and Y regardless of whether X has been expressed or not
2. The *transient behavior* of Y depends on the presence and activity of X at a given point in time

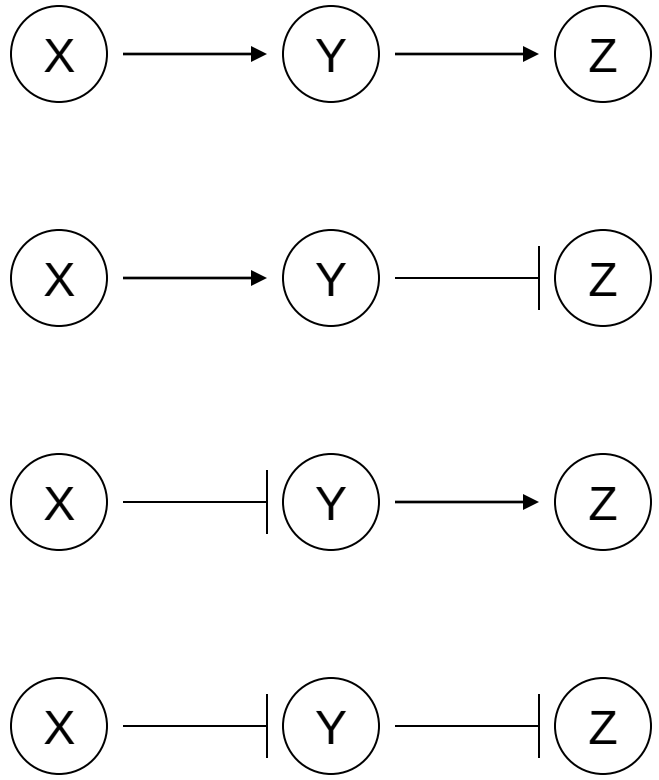


Graph Representation of Gene Expression Regulation



Activation or Repression of Gene Transcription

- Transcription factors may activate or repress the transcription of a gene
  - In activation, the activity of the regulator leads to the activity of the regulatee
    - Activation is indicated by a pointed arrow
  - In repression, the activity of the regulator leads to the inactivity of the regulatee
    - Repression is indicated by a block
- Chains of activation and repression effects can achieve net activation or net repression of the final gene product by the initial transcription factor



Production Rates of Gene Products

- The production rates of gene products can be expressed as a function of active transcription factor concentration

$$\text{rate of production of } Y = f([X^*])$$

$f(\cdot)$: the input function

  - The input function is monotonic
    - increasing if X* activates Y
    - decreasing if X* represses Y
  - It is also limited from above and below
    - The rate of production represents an average response of a stochastic process:

X* binds the promoter of gene Y with some probability

    - This probability cannot be higher than 1 (limiting factor in activation) or lower than 0 (limiting factor in repression)
      - If the probability is 1, then all promoters of gene Y are continually occupied by X*, and the protein synthesis machinery is at full capacity
      - If the probability is 0, then all promoters of gene Y are free of X*, and no synthesis of the protein Y is taking place

The Hill Function

- For activation: $f^{act}([X^*]) = \beta \frac{[X^*]^n}{\kappa^n + [X^*]^n}$
- For repression: $f^{rep}([X^*]) = \beta \frac{1}{1 + \left(\frac{[X^*]}{\kappa}\right)^n}$

κ: activation coefficient

β: maximal expression rate

n: Hill coefficient

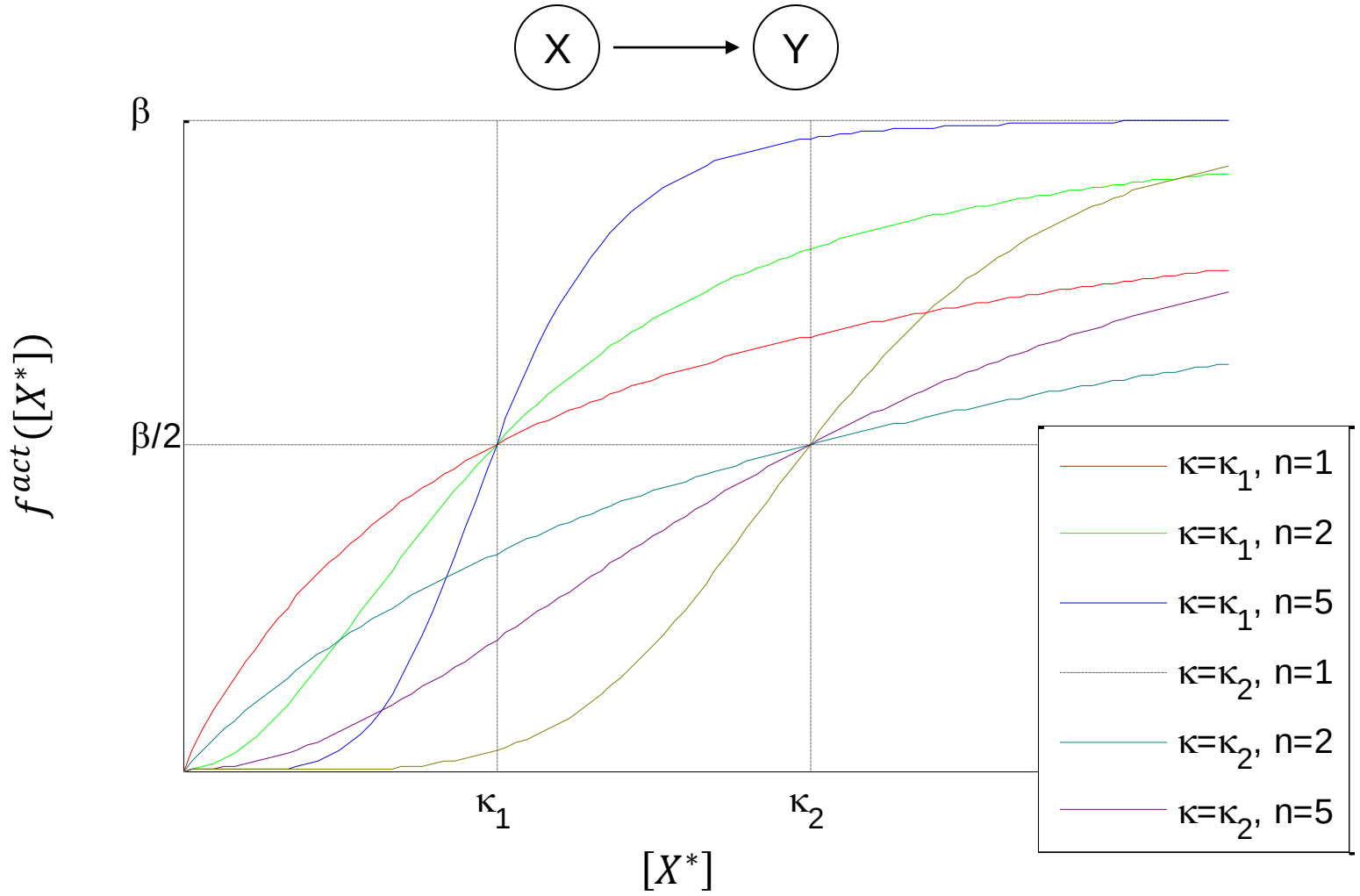
- Note that $[X^*] = \kappa$ provides 50% activity/repression

$$f(\kappa) = 0.5 \max_{[X^*]} f([X^*])$$

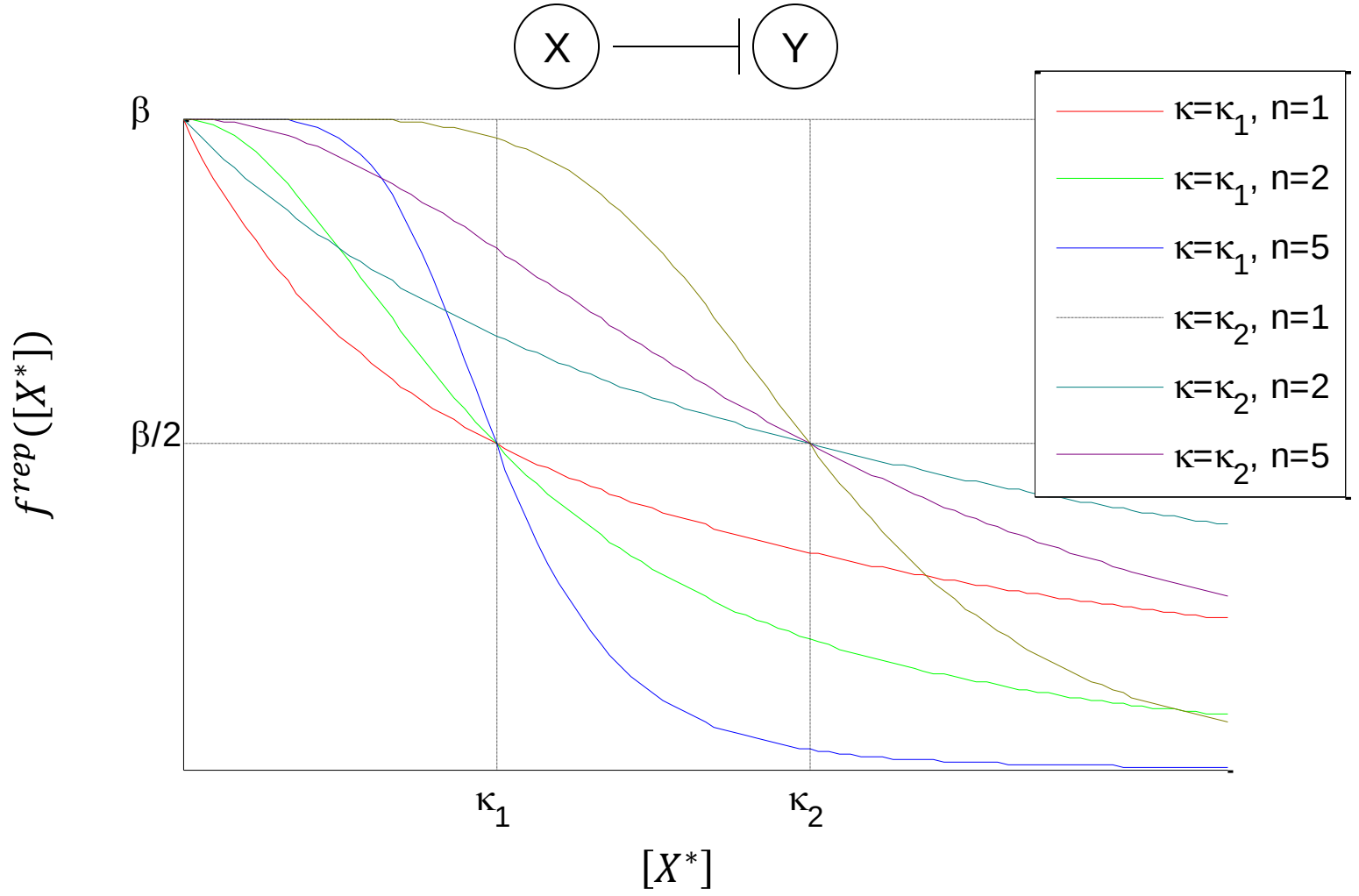
- For genes with non-zero basal expression rate,

$$f([X^*]) \leftarrow f([X^*]) + \beta_0$$

The Hill Function for Activation



The Hill Function for Repression



Logic Approximation to the Hill Function

- In both activation and repression, the input function crosses the $\beta/2$ level at $[X^*] = \kappa$
- As the function becomes steeper (with increasing n) the Hill function approximates a step function:

$$\lim_{n\to\infty} \frac{\beta[X^*]^n}{\kappa^n + [X^*]^n} = \beta u([X^*] - \kappa) = \beta\ \mathbf{1}([X^*] \geq \kappa)$$

$$\lim_{n\to\infty} \frac{\beta}{1 + \left(\frac{[X^*]}{\kappa}\right)^n} = \beta u(\kappa - [X^*]) = \beta\ \mathbf{1}([X^*] \leq \kappa)$$

where $u(\cdot)$ is the unit step function.

Logic Approximation to the Hill Function

- This logic approximation becomes useful for figuring out how a gene will be regulated by the associated transcription factors
  - For activation:

$$f^{act-app}([X^*]) = \begin{cases} \beta & \text{if } [X^*] \geq \kappa \\ 0 & \text{otherwise} \end{cases}$$

  - For repression:

$$f^{rep-app}([X^*]) = \begin{cases} \beta & \text{if } [X^*] \leq \kappa \\ 0 & \text{otherwise} \end{cases}$$

Multivariate Input Functions

- Genes may be regulated simultaneously by several transcription factors
- This links the rate of production of a gene product to the concentrations of the associated transcription factors at a given point in time
  - Hence, multivariate input functions:

$$\text{production rate of } Y = f([X_1^*], [X_2^*], \dots)$$

- Depending on the precise relationship, these multivariate input functions can be approximated using mathematical expressions
  - AND, OR, SUM, ...
- The precise shape and form of the input function for each gene continually evolves under selection pressure
  - Neutral or advantageous modifications achieved by the altered molecular machinery
  - Variety ensures adaptability under varying environmental conditions

Multivariate Input Functions

- **AND**: The expression of gene Y requires the binding of both X_1* and X_2* to its promoter

$$\text{production rate of } Y \approx \beta u([X_1^*] - \kappa_1) u([X_2^*] - \kappa_2)$$

- **OR**: The binding of either X_1* or X_2* to its promoter suffices to trigger the expression of gene Y

$$\text{production rate of } Y \approx \beta \max\{u([X_1^*] - \kappa_1), u([X_2^*] - \kappa_2)\}$$

- **SUM**: The rate of expression of gene Y is related to a linear combination of X_1* and X_2* concentrations in the nuclear environment

$$\text{production rate of } Y \approx f(\beta_1 [X_1^*] + \beta_2 [X_2^*])$$

Gene Expression Transients

- The concentration of a given protein in the cell is regulated jointly by
  - the expression of its gene via the associated transcription factors and subsequent synthesis (at rate β)
  - the degradation of the protein (at rate $\alpha_{deg}[Y]$)
  - the dilution due to cell growth (at rate $\alpha_{dil}[Y]$)
- Assuming that the protein product of gene Y is synthesized at full capacity, the rate of change of $[Y]$ in time is given by

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

where $\alpha = \alpha_{deg} + \alpha_{dil}$.

- At steady state, the opposing forces balance each other out at a stable concentration $[Y]$

$$\frac{d}{dt}([Y])(t) = 0 \;\Rightarrow\; \beta - \alpha[Y]_{st} = 0 \;\Rightarrow\; [Y]_{st} = \frac{\beta}{\alpha}$$

Dynamics of Gene Activation

- In gene activation, at $t = 0$,
  - the initial value of $[Y]$ is 0
  - the activating transcription factor
- The transient behavior of $[Y]$ is then given by the function

$$([Y])(t) = [Y]_{st}(1 - e^{-\alpha t}) \ .$$

for $t \geq 0$.

  - This can be seen easily by computing the derivative of $([Y])(t)$ with respect to t:

$$\begin{aligned}\frac{d}{dt}([Y])(t) &= \frac{d}{dt}\left([Y]_{st}(1 - e^{-\alpha t})\right) \\ &= -[Y]_{st}(-\alpha)e^{-\alpha t} \\ &= \alpha\left([Y]_{st} - ([Y])(t)\right) \\ &= \alpha\left(\frac{\beta}{\alpha} - ([Y])(t)\right) \\ &= \beta - \alpha([Y])(t)\end{aligned}$$

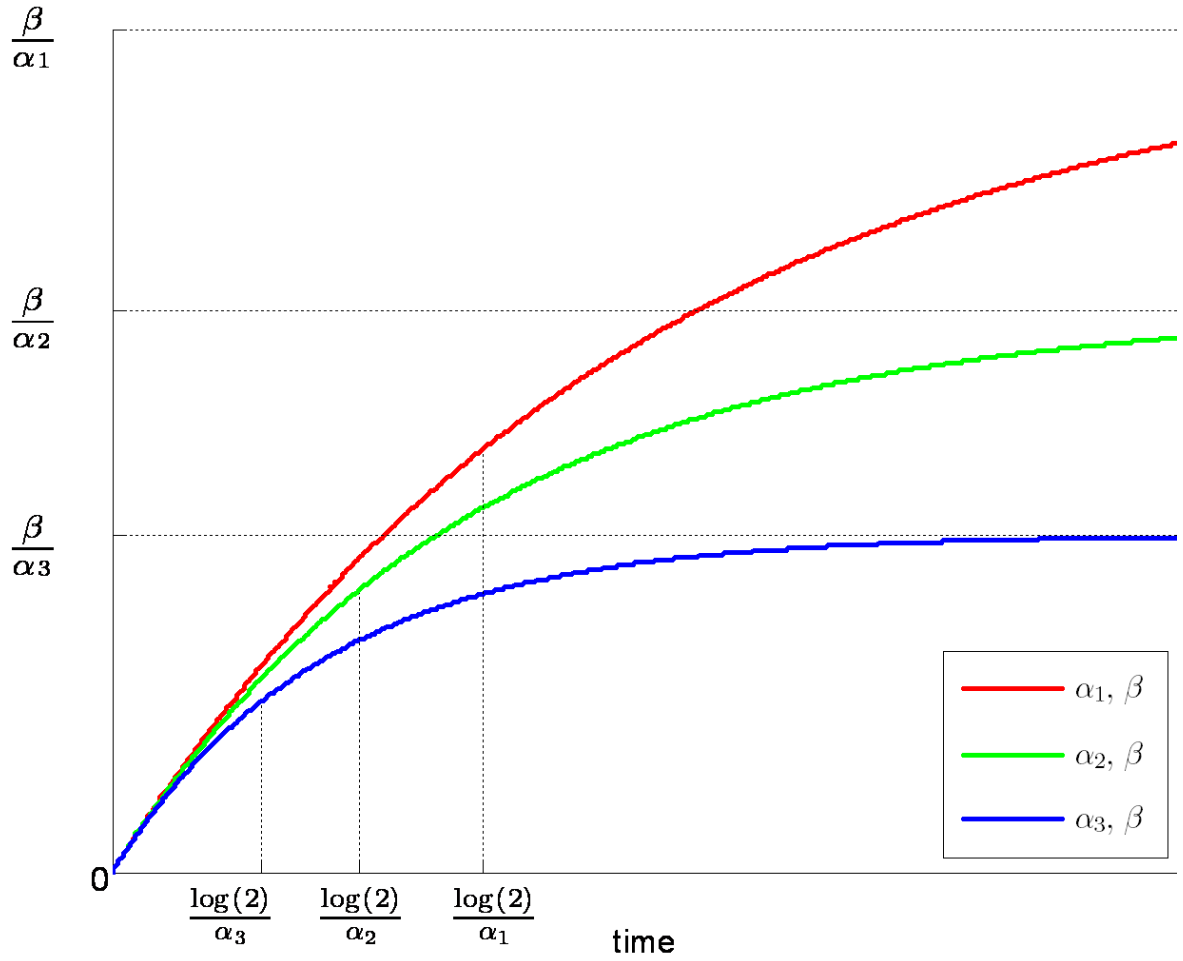
Dynamics of Gene Activation

- This suggests:
  - With the activation of gene expression at full capacity, $[Y]$ rises from 0 to the steady state value of $[Y]_{st}$
  - The time it takes to reach half the steady state level is

$$[Y]_{st}\left(1 - e^{-\alpha T_{1/2}}\right) = \frac{[Y]_{st}}{2} \Rightarrow T_{1/2} = \log(2)/\alpha$$

  - Hence, the delay (along with the time constant of the exponential rise) is inversely proportional to α (**not** β)
    - Hence the maximal production rate has no bearing on the speed of gene activation
    - Conversely, increasing the production rate will not reduce the delay to the half the steady state level

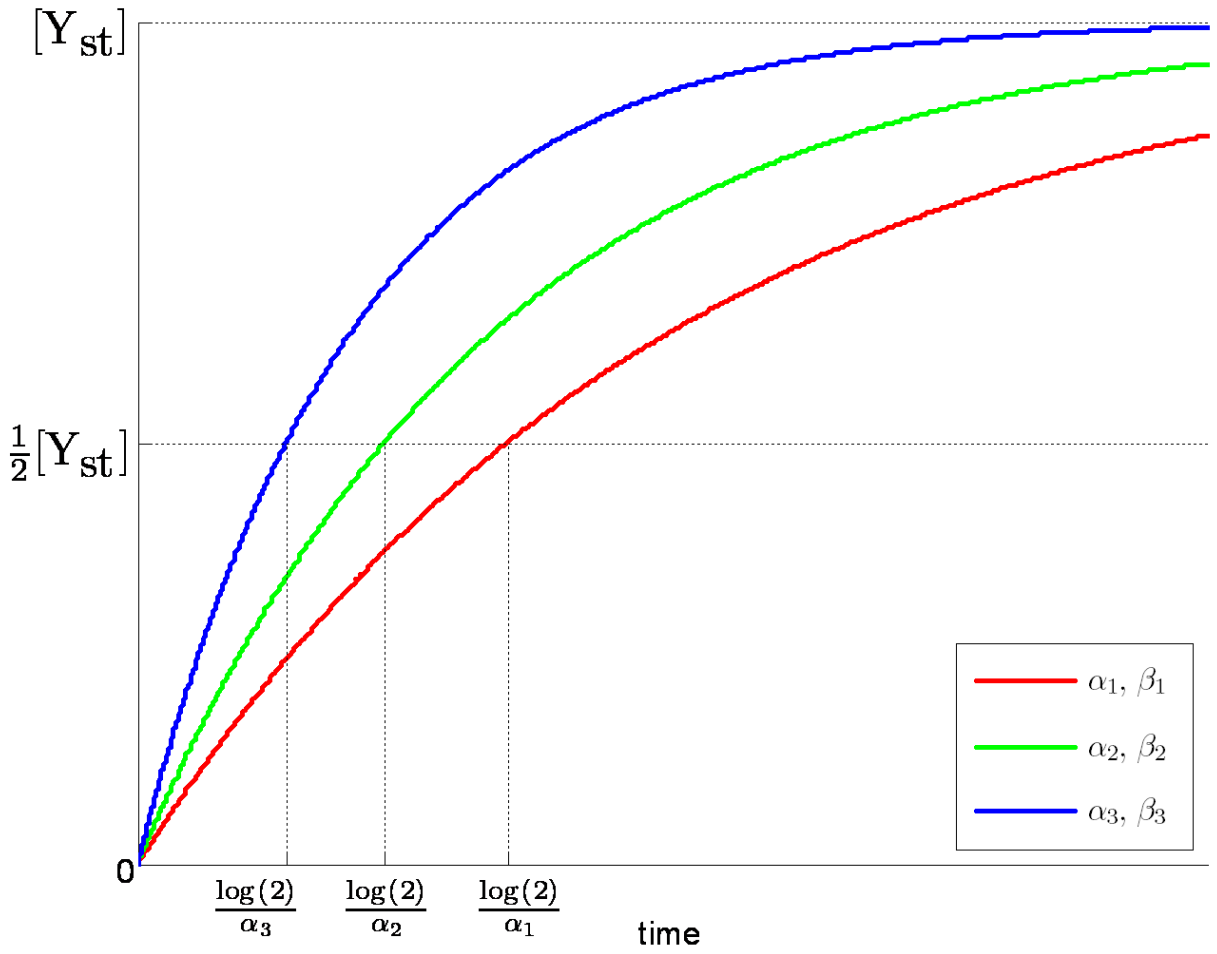
Dynamics of Gene Activation



- Transient graphs for $\alpha_1 < \alpha_2 < \alpha_3$
- Notes:
  - Larger α implies faster rise to the steady state level
  - To maintain the same steady state level with a larger α requires larger β
  - Larger β implies greater investment from the cell's part to the molecular machinery of $[Y]$ protein synthesis

→ **Insensible cycle of rapid protein synthesis and degradation**

Dynamics of Gene Activation



- $\alpha_1 < \alpha_2 < \alpha_3$
- $\beta_1 < \beta_2 < \beta_3$ such that

$$\frac{\beta_1}{\alpha_1} = \frac{\beta_2}{\alpha_2} = \frac{\beta_3}{\alpha_3}$$

- Note that **faster response is achieved at the expense of greater production and degradation rates**
  - So that the same steady state level is maintained

Dynamics of Protein Decay

- Now, suppose the protein product $[Y]$ is initially at the steady state $[Y]_{st}$
- The transient behavior of $[Y]$ after the moment when the gene transcription is turned off (at time $t = 0$) is given by the function

$$([Y])(t) = [Y]_{st}e^{-\alpha t}$$

for $t \geq 0$.

  - Note that turning the gene off means no more production of $[Y]$
  - The differential equation now becomes

$$\frac{d}{dt}([Y]) = -\alpha[Y]$$

  - Taking the derivative of $([Y])(t)$ with respect to t validates the expression

$$\begin{aligned}\frac{d}{dt}([Y]_{st}e^{-\alpha t}) &= [Y]_{st}(-\alpha)e^{-\alpha t} \\ &= -\alpha[Y]_{st}e^{-\alpha t} \\ &= -\alpha([Y])(t)\end{aligned}$$

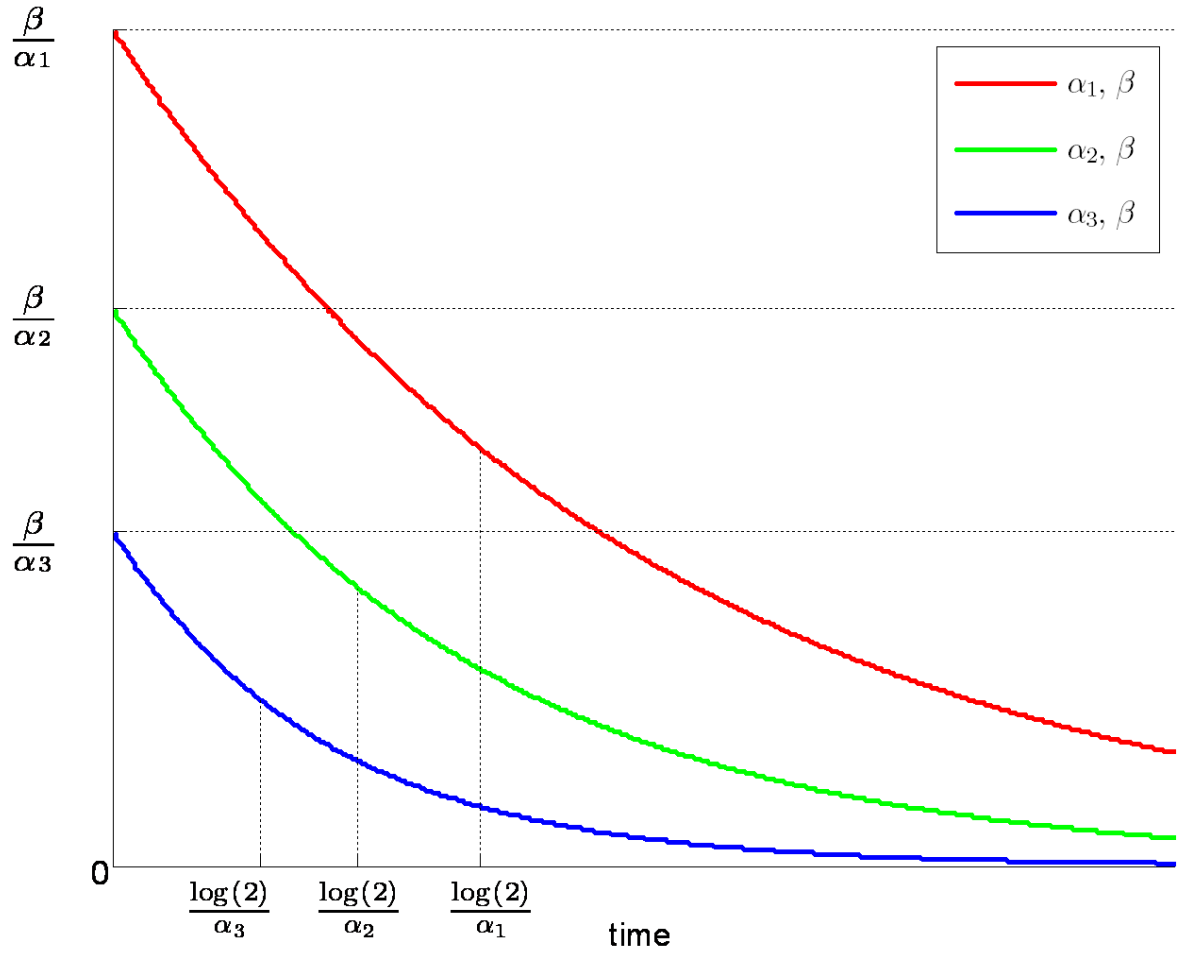
Dynamics of Protein Decay

- Remarks:
  - The delay in falling from the steady state level to zero is again related to α (and **not** β)
  - The time it takes for the concentration to fall to half the steady state level is again

$$[Y]_{st} e^{-\alpha T_{1/2}} = \frac{[Y]_{st}}{2}$$

$$\Rightarrow T_{1/2} = \log(2)\,/\alpha$$

Dynamics of Protein Decay



Incorporating the mRNA Dynamics into the Transient Analysis

- The differential equation linking the rate of change of protein Y concentration in time assumes that gene expression directly leads to protein synthesis
- In fact, it bears its own dynamics due to the mRNA transients
  - mRNA is synthesized and degraded at its own rate
  - The synthesis of the protein Y is linked to the concentration of Y mRNA in the cytoplasm

Incorporating the mRNA Dynamics into the Transient Analysis

- This suggests the alternative rate equation

$$\frac{d}{dt}([Y])(t) = \beta'([Y_{mRNA}])(t) - \alpha([Y])(t)$$

with $[Y_{mRNA}]$ denoting the concentration of Y mRNA, governed by

$$\frac{d}{dt}([Y_{mRNA}])(t) = \beta_{mRNA} - \alpha_{mRNA}([Y_{mRNA}])(t)$$

  - At the steady state of Y mRNA,

$$\frac{d}{dt}([Y_{mRNA}])(t) = 0 \Rightarrow [Y_{mRNA}]_{st} = \frac{\beta_{mRNA}}{\alpha_{mRNA}}$$

  - Therefore, a more accurate expression for the maximal production rate β of the protein Y is

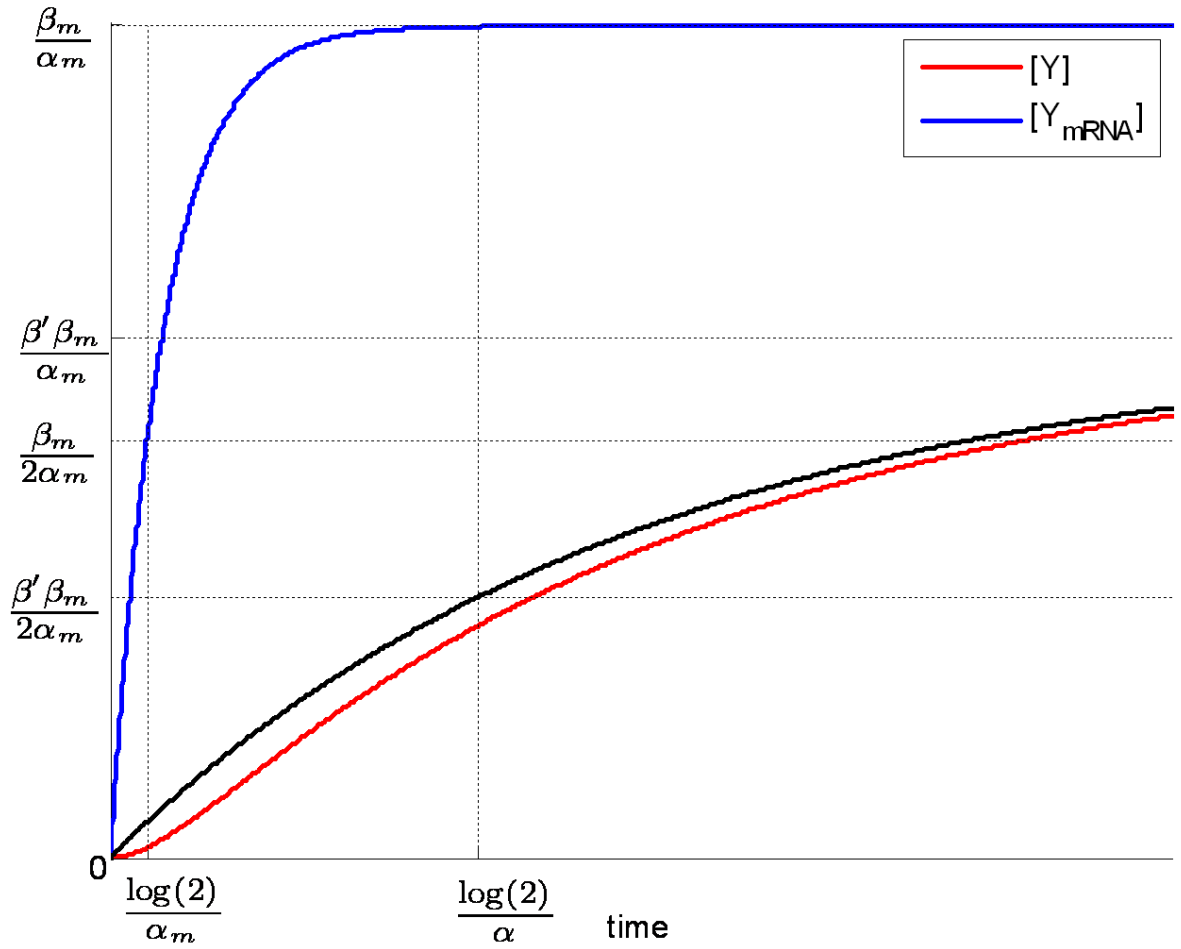
$$\beta = \frac{\beta'\beta_{mRNA}}{\alpha_{mRNA}}$$

  - This holds for $\alpha_{mRNA} \gg \alpha$
    - Y mRNA reaches steady state much faster than Y protein

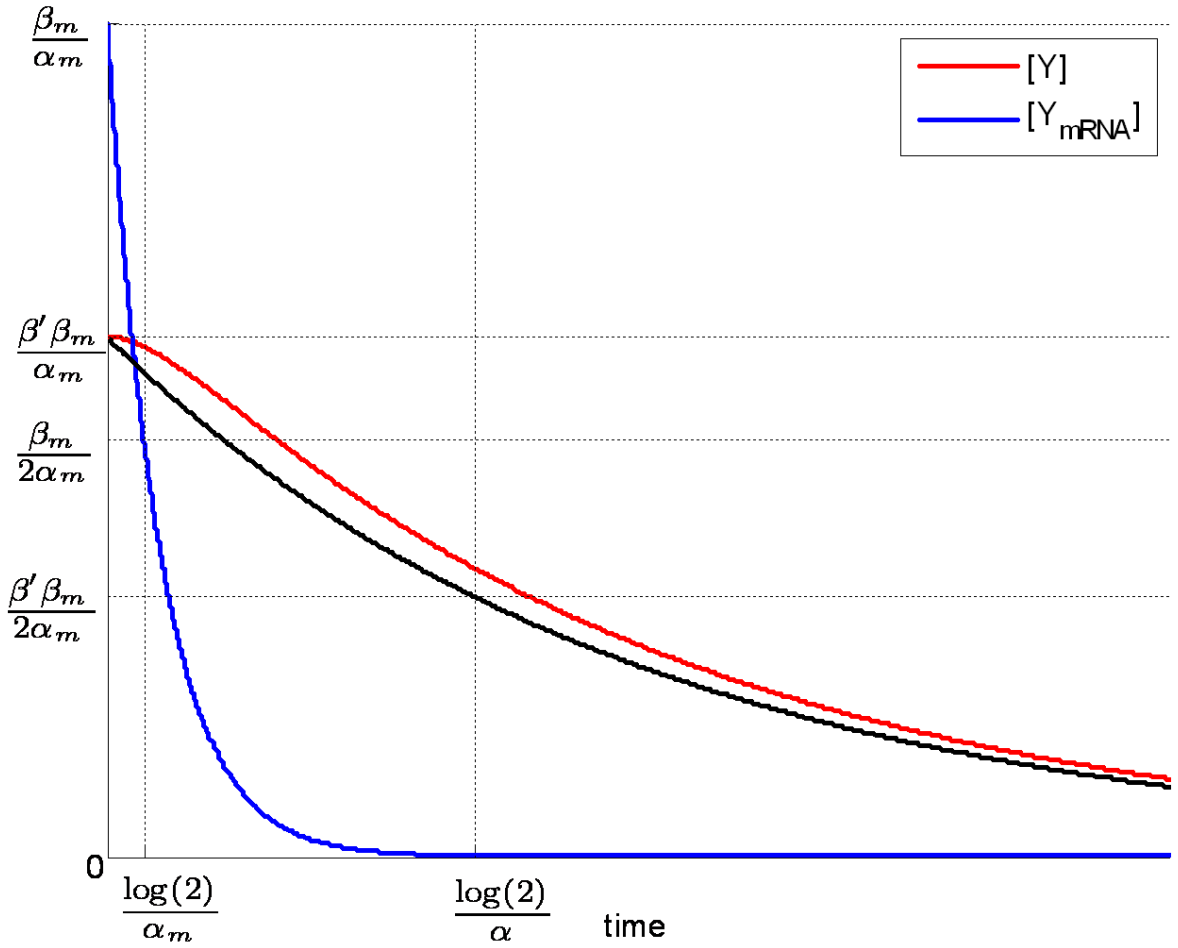
Incorporating the mRNA Dynamics into the Transient Analysis

- Pseudo code for a numeric solution of the corresponding system of ordinary differential equations
  - Initialization
    - $([Y])(0) = ([Y_{mRNA}])(0) = 0$
    - $\Delta t \ll 1$
  - For $t = 0, \Delta t, 2\Delta t, 3\Delta t, \ldots, t_{\max}$:
    - Calculate $\frac{d}{dt}([Y])(t)$ using $([Y])(t)$ and $([Y_{mRNA}])(t)$
    - Calculate $\frac{d}{dt}([Y_{mRNA}])(t)$ using $([Y])(t)$ and $([Y_{mRNA}])(t)$
    - Set $([Y])(t + \Delta t) = ([Y])(t) + \Delta t \cdot \frac{d}{dt}([Y])(t)$
    - Set $([Y_{mRNA}])(t + \Delta t) = ([Y_{mRNA}])(t) + \Delta t \cdot \frac{d}{dt}([Y_{mRNA}])(t)$

mRNA Dynamics in the Transient Analysis: Activation



mRNA Dynamics in the Transient Analysis: Repression



Remarks

- Transient analysis of protein concentrations can be carried out using a variety of dynamic models
  - Rise or decay for protein concentration through decaying exponentials
  - Incorporating the dynamics of mRNA concentration
  - Incorporating the travel time of mRNA from the nucleus to the Endoplasmic Reticulum
  - Incorporating tRNA concentration changes in the cytoplasm
  - …
- In general, each dynamic model possesses its own set of advantages and disadvantages
  - Simplistic models allow easy prediction of system behavior but are not necessarily very accurate
  - Model complexity can be increased by incorporating additional factors for more accurate predictions, at the expense of intuitive understanding

Dynamics of Stable Proteins

- Stable proteins are those that are not actively degraded by the cell

$$\alpha_{\mathrm{deg}} = 0 \Rightarrow \alpha = \alpha_{\mathrm{dil}}$$

- In the absence of degradation, the concentration drops to its half level at the time of cell division

$$T_{1/2} = \frac{\log(2)}{\alpha_{\mathrm{dil}}} = T_{cc}$$

$$\Rightarrow \alpha_{\mathrm{dil}} = \frac{\log(2)}{T_{cc}}$$

T_{cc}: The period of one cell cycle (cell generation time)

- This suggests that response time can be a limiting constraint in the evolutionary design of gene transcription circuits
  - In a case where the response is to be given through non-degraded proteins
  - If the response time exceeds one cell cycle, then the cell cannot hope to respond to the environmental changes in time
  - The faster the response, the better
- Gene transcription networks that develop additional mechanisms to speed up the response time would then gain important selective advantage

Summary

- The response time in simple gene regulation circuits are determined by degradation and dilution rates of proteins
  - This remains true even when the dynamic model complexity is increased
- These response times are generally too large to be practical when responding to environmental inputs such as those on nutrient levels
  - Impossible for stable proteins
    - Cell generation time is about 30min for some bacteria
    - Human cell generation time is about one day
- In actuality, gene transcription networks possess additional mechanisms to reduce the response time without embarking on a futile cycle of rapid protein synthesis and degradation

➔ Network motifs in gene transcription networks