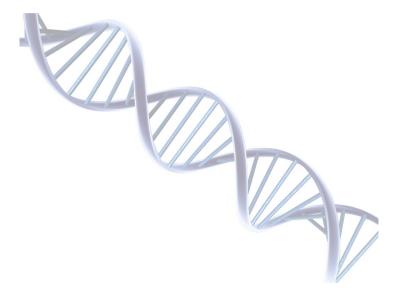




# **Genetic Circuit Engineering Meets Control Theory**



### **Domitilla Del Vecchio**

Mechanical Engineering
MIT





### A journey towards modular composition

**Engineering biology: Why and how** 

Modular composition: A grand challenge

Inter-module loads and the *load driver* 

Disturbance attenuation via time scale separation

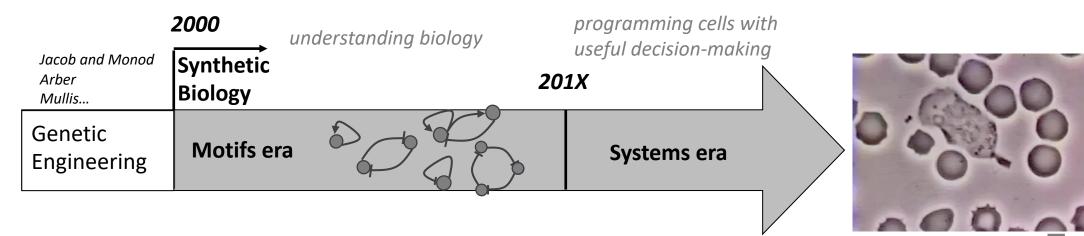
Resource loading and the resource decoupler

Disturbance rejection despite leaky integral actions

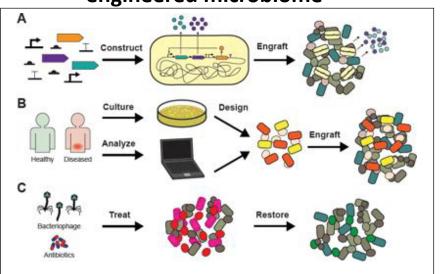
Decentralized implementation

Outlook

# **Engineering biology: historical view**



#### engineered microbiome



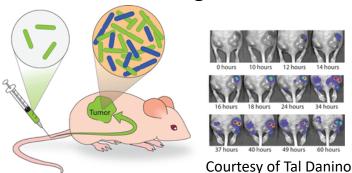
control of cell fate

regenerative

medicine

Regenerative

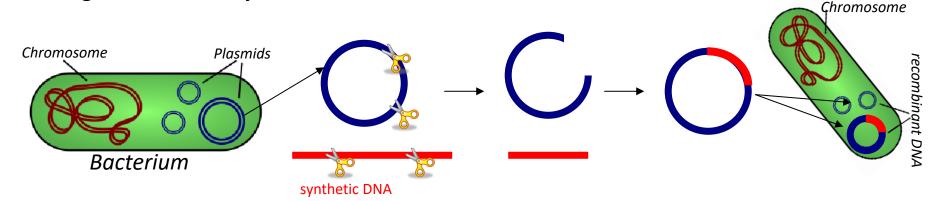
# curing disease



tracking, recognizing, killing cancer cells

### How do we engineer cells with de novo decision making

#### Decision making is encoded in synthetic DNA

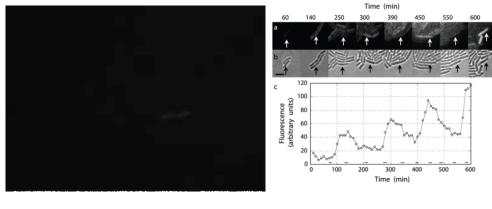


Synthetic DNA encodes genes that express proteins regulating expression of other genes (synthetic/endogenous)

- regulatory interactions create circuits
- interactions can be externally controlled by chemical signals

#### 

#### A biomolecular oscillator



(Elowitz and Leibler, Nature 2000)

### A journey towards modular composition

**Engineering biology: Why and how** 

Modular composition: A grand challenge

Inter-module loads and the *load driver* 

Disturbance attenuation via time scale separation

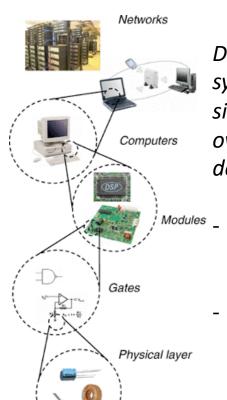
Resource loading and the resource decoupler

Disturbance rejection despite leaky integral actions

Decentralized implementation

Outlook

# Modular design in engineering biology

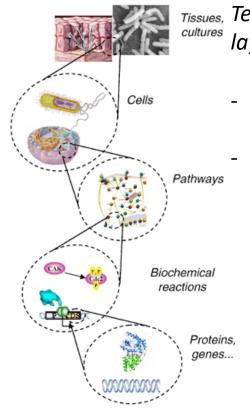


Describing a sophisticated system as the composition of simpler subsystems helps overcoming the complexity of design:

- we "forget" the details within subsystems when we compose them
- feedback can maintain I/O properties providing simplified abstractions for layered design

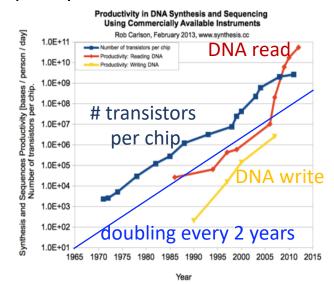
increased scale is becoming possible

modularity is critical to manage complexity/time the I/O behavior of a "module" should not change upon composition



Tissues, Tempting in engineering biology: cultures layered and modular composition

- we have large libraries of genetic parts
- we can synthesize DNA very quickly

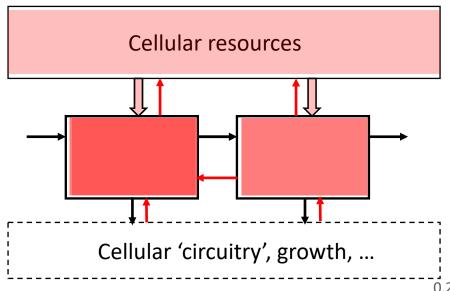


in practice, modularity fails

⇒ need to re-design "modules" after composition

for a circuit with 11 genes it takes one PhD thesis of 5-6 years

### Some reasons why modularity is a challenge



these issues can be viewed as *lack* of robustness to perturbations

can we "insulate" desired I/O behaviors from these perturbations?

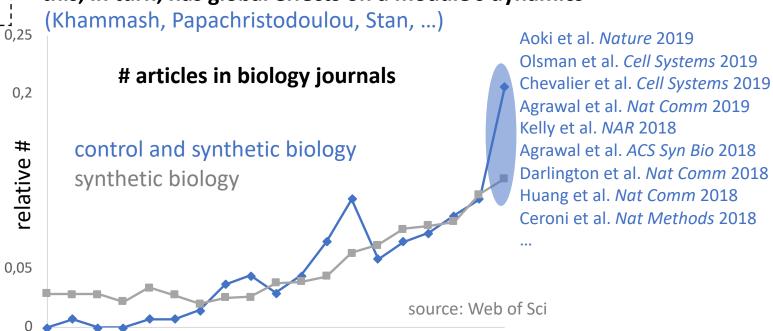
→ this is a Control System Design problem

Loads applied by downstream modules change the behavior of upstream systems

(Del Vecchio, Hespanha, Klavins, Papachristodoulou, Sontag, ...)

Modules apply a load the cellular resources: creates subtle couplings (Bates, Del Vecchio, Murray, Stan,...)

Modules often have "off-target" interactions, affect growth rate, and this, in turn, has global effects on a module's dynamics



### A journey towards modular composition

**Engineering biology: Why and how** 

Modular composition: A grand challenge

Inter-module loads and the load driver

Disturbance attenuation via time scale separation

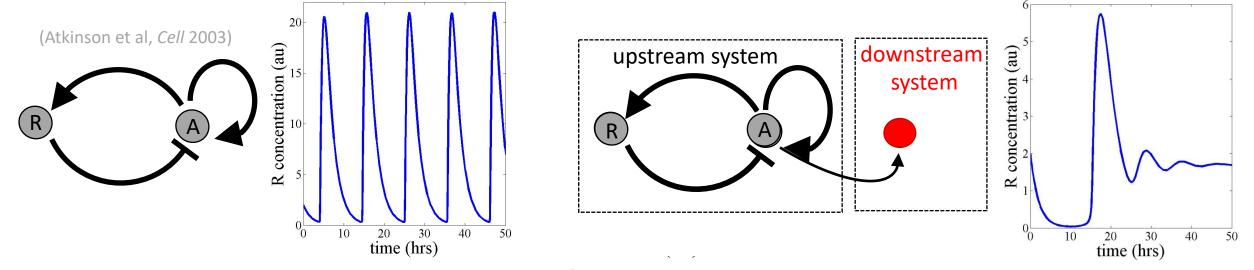
Resource loading and the resource decoupler

Disturbance rejection despite leaky integral actions

Decentralized implementation

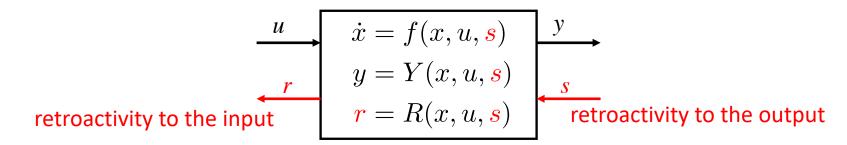
Outlook

### Inter-module loading changes upstream system's dynamics



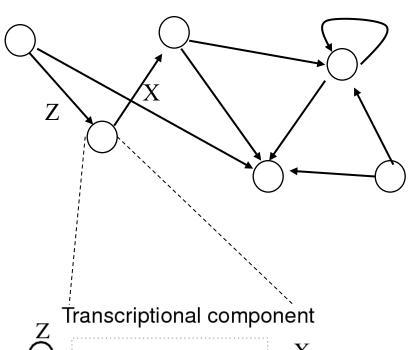
loads change the behavior of the upstream system → we fail to transmit the signal to the downstream system

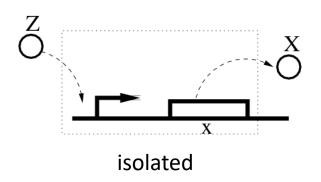
systems & signals representation of loads: retroactivity



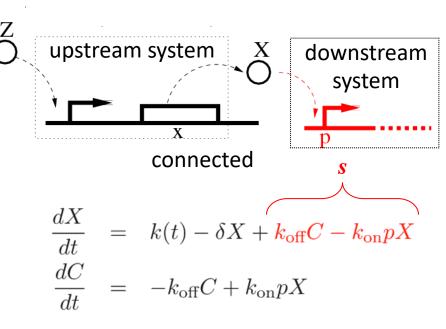
the I/O model of the **isolated system** is obtained when s=0

### Retroactivity is a reaction flux affecting upstream system's output



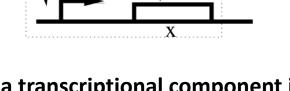


$$\frac{dX}{dt} = k(t) - \delta X$$



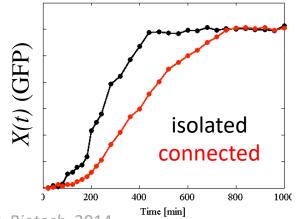
experiments in transcriptional components in yeast

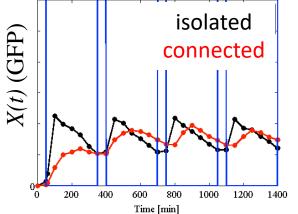
retroactivity reduces bandwidth of upstream system

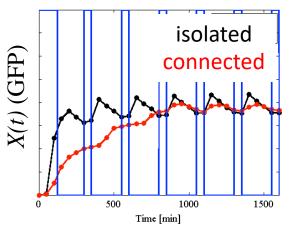


a transcriptional component is an input/output module

how does its input/output response change upon interconnection?



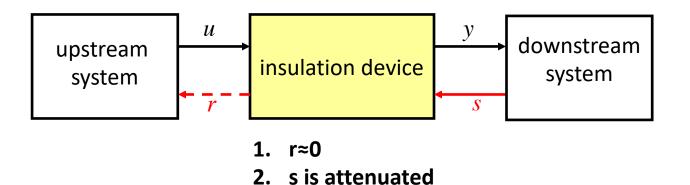




Jayanthi et al. ACS Syn Bio 2013; Mishra et al, Nat. Biotech, 2014

### Insulation devices to mitigate retroactivity

consider retroactivity as a state-dependent disturbance



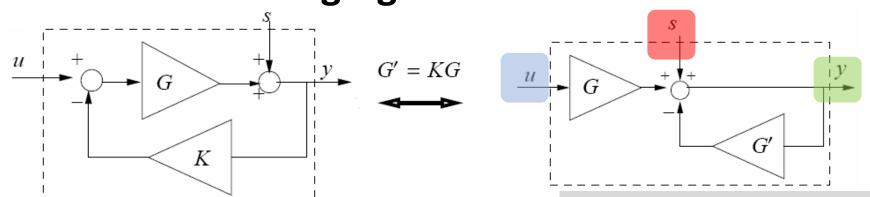
Principle 1 high-gain feedback

requires an explicit negative feedback

Principle 2 time scale separation

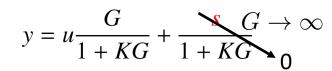
no explicit feedback required

# From high-gain feedback to time-scale separation

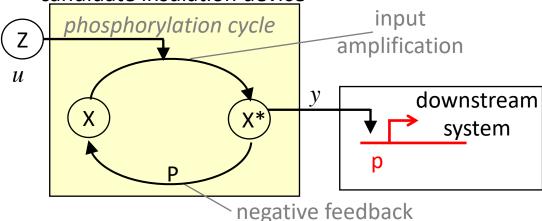


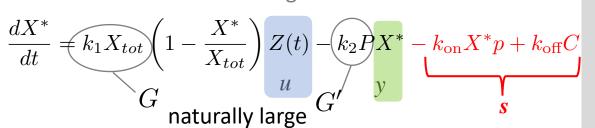
large input amplification and a large negative feedback

what biomolecular systems can realize it?



candidate insulation device





input  $u \rightarrow Z = Sln1$  W = Ypd1 X = Skn7 most naturally-occurring systems are not that "simple" and involve multiple modifications need a different scheme for retroactivity attenuation:

- no explicit feedback

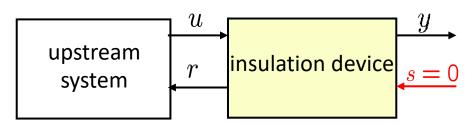
what is common between these two systems?

- speed

output

# Retroactivity attenuation via time-scale separation

#### isolated system



upstream system 
$$\dot{u}=f_0(u,t)+r(u,y)$$
 insulation device  $\dot{y}=G_1f_1(u,y)$  large  $G_1\gg 1$ 

#### connected system



upstream system 
$$\dot{\bar{u}} = f_0(\bar{u},t) + r(\bar{u},\bar{y})$$

insulation device 
$$\dot{\bar{y}} = G_1 f_1(\bar{u}, \bar{y}) + G_2 M_s(\bar{y}, v)$$

downstream system 
$$\,\dot{v} = -G_2 N_{\color{red} s}(\overline{y},v)\,$$
 very large rates

 $G_2 \gg G_1$ 

**Fact:** There are a matrix T and a non-singular matrix P such that  $P \cdot M - T \cdot N = 0$  (closed system)

**Theorem:** If 
$$\left.\frac{\partial f_1(u,y)}{\partial y}\right|_{y=h(u)}$$
 is Hurwitz with  $y=h(u) \Rightarrow f_1(u,y)=0$ , then:

$$||y(t) - \bar{y}(t)|| = \mathcal{O}(1/G_1), \ \ \forall t \in [t_b, t_f] \ \ \text{independent of} \ \ G_2Ms$$

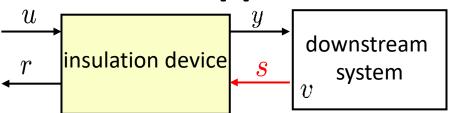
Proof: use singular perturbation and nested application of Tikhonov's theorem

change of cords: 
$$x = Py + Tv$$

$$\begin{cases}
\epsilon_1 \dot{\bar{x}} = P f_1(\bar{u}, \bar{y}), & \epsilon_1 = 1/G_1 \\
\epsilon_2 \dot{v} = -s(\bar{y}, v), & \epsilon_2 = 1/G_2
\end{cases}$$

$$\epsilon_1 = 0 \Rightarrow \bar{y} = h(\bar{u})$$

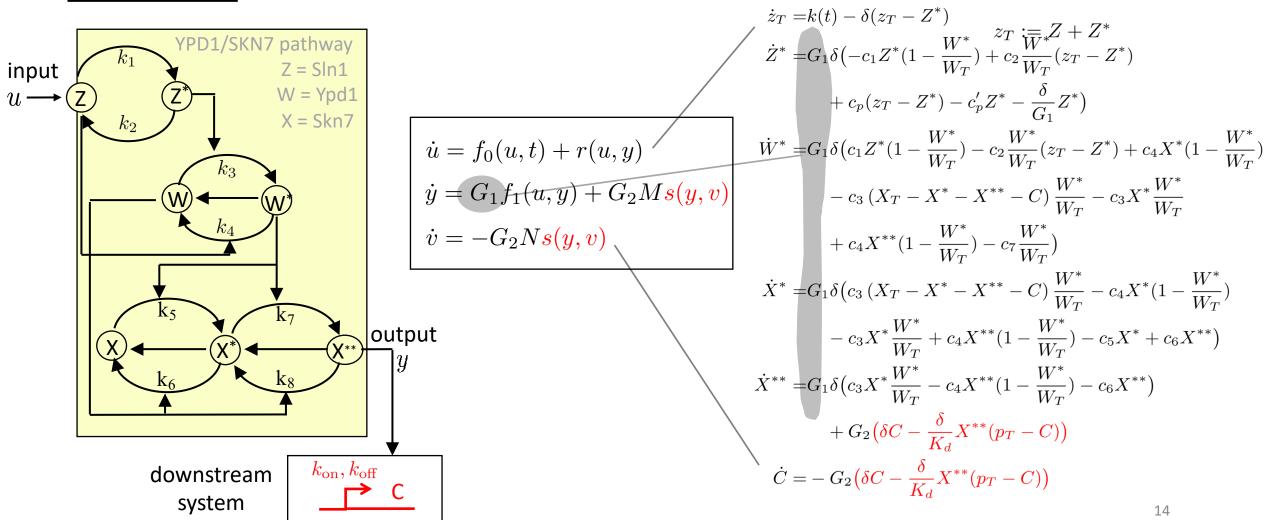
### **Application to signal transduction networks**



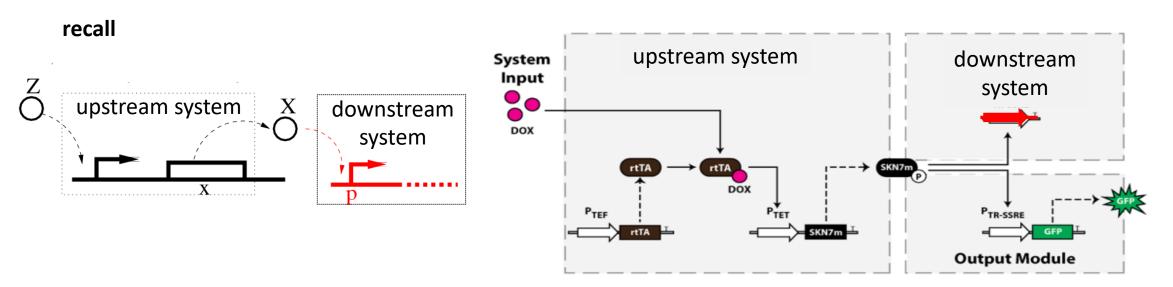
 $\delta \in [0.001, 0.01] \text{ min}^{-1}$   $k_i W_T \in [1, 100] \text{ min}^{-1}$   $k_{\text{off}} \in [0.1, 10^4] \text{min}^{-1}$ 

Z: gene expression time scale W/X: time scale of signal transduction  $G_1=(k_iW_T)/\delta\gg 1$ 

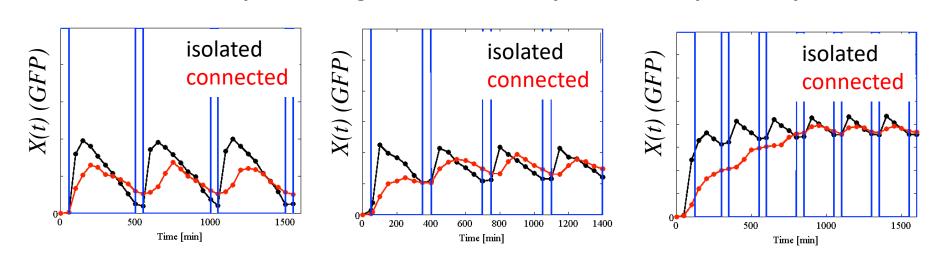
C: time scale of reversible binding to DNA  $G_2 = k_{\rm off}/\delta \gg 1$ 



### The Load Driver: Insulation by time scale separation

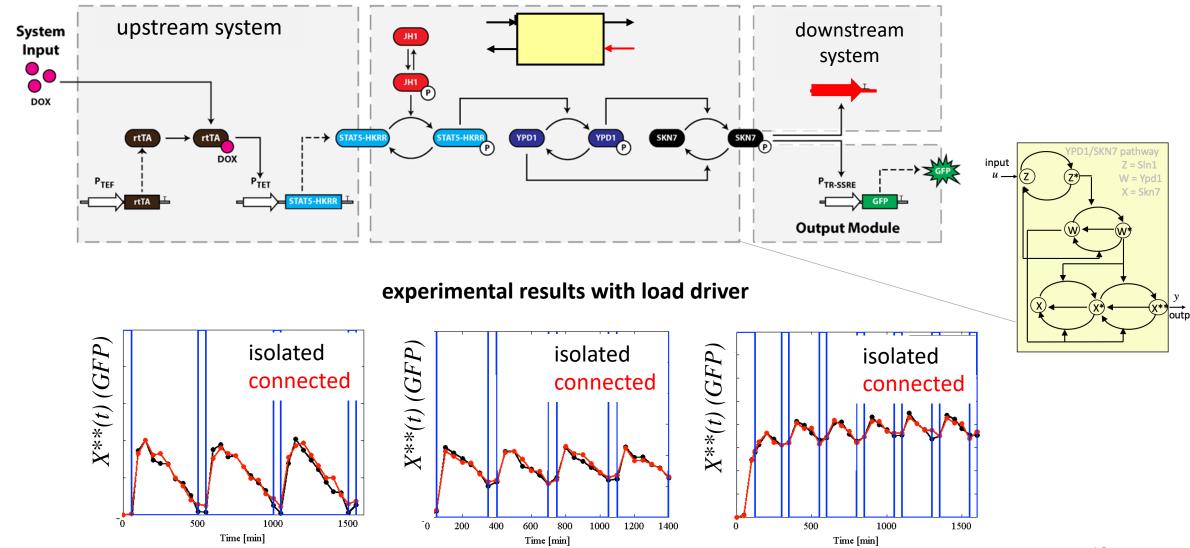


#### directly connecting the downstream system to the upstream system



### The Load Driver: Insulation by time scale separation





### The Load Driver: Insulation by time scale separation

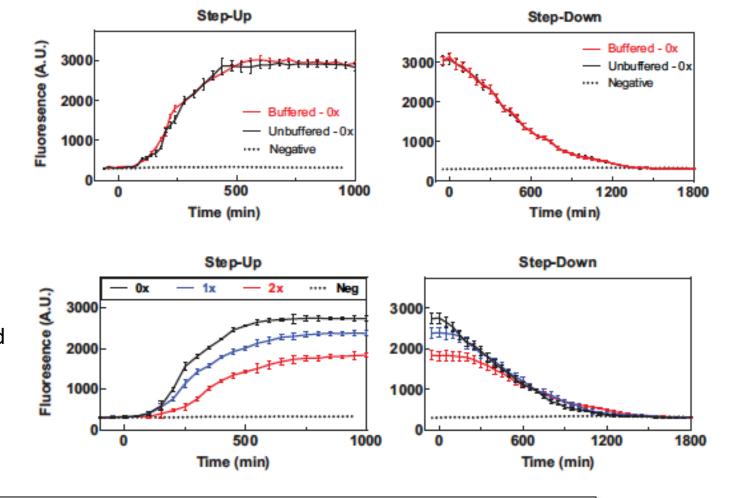
#### **Experiments with fast time scales**

(endogenous amounts of YPD1 and SKN7)

#### **Experiments with slow time scales**

(reduced amounts of YPD1 and SKN7 obtained through weaker constitutive promoters)

insulation property is lost



Slow/fast/slow pattern allow to reliably transmit signals to large loads: the synergy between slow transcription and fast signal transduction is likely to be used by natural systems to insulate signals from downstream loads

### A journey towards modular composition

**Engineering biology: Why and how** 

Modular composition: A grand challenge

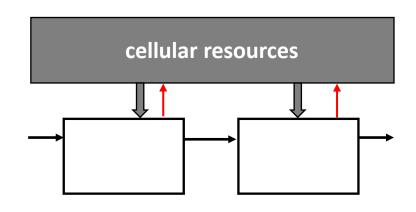
Inter-module loads and the load driver → upstream system of the load driver → upstream system

### Resource loading and the resource decoupler

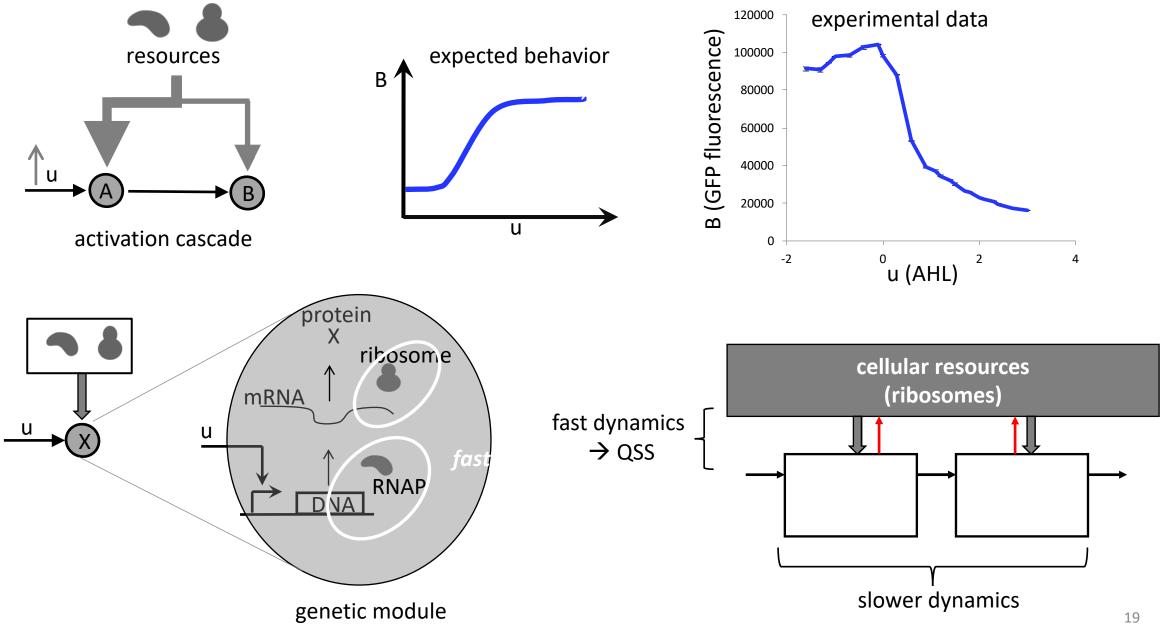
Disturbance rejection despite leaky integral actions

Decentralized implementation

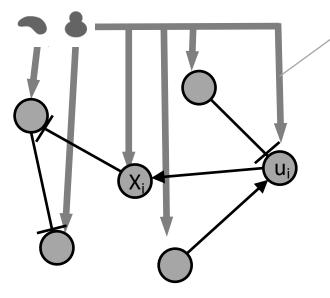
Outlook



# Modules become coupled by loading cellular resources



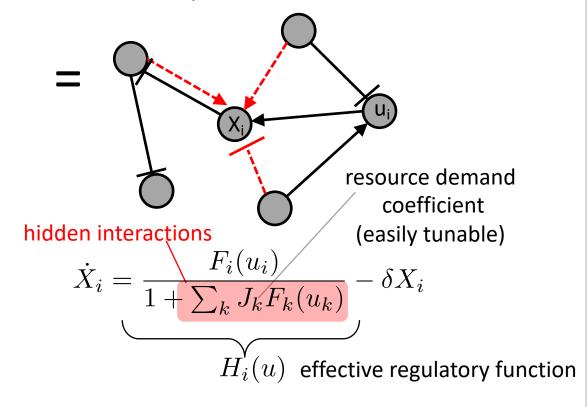
# Coupling can be mathematically captured by "hidden" graphs



$$\dot{X}_i = F_i(u_i) - \delta X_i$$

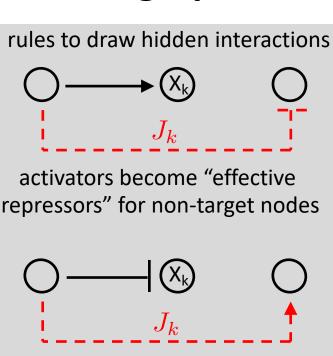
intended regulatory function

use time scale separation and conservation of resources

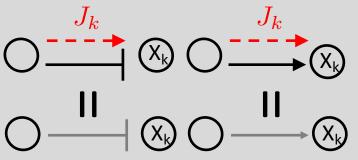


effective interaction graph

$$\frac{\partial H_i(u)}{\partial u} = \underbrace{\frac{\partial F_i/\partial u}{(1+\sum_j J_j F_j)^2}}_{\text{re-scaling of intended regulatory links}} - \underbrace{\frac{F_i\sum_j J_j \partial F_j/\partial u}{(1+\sum_j J_j F_j)^2}}_{\text{"hidden" interaction graph}}$$

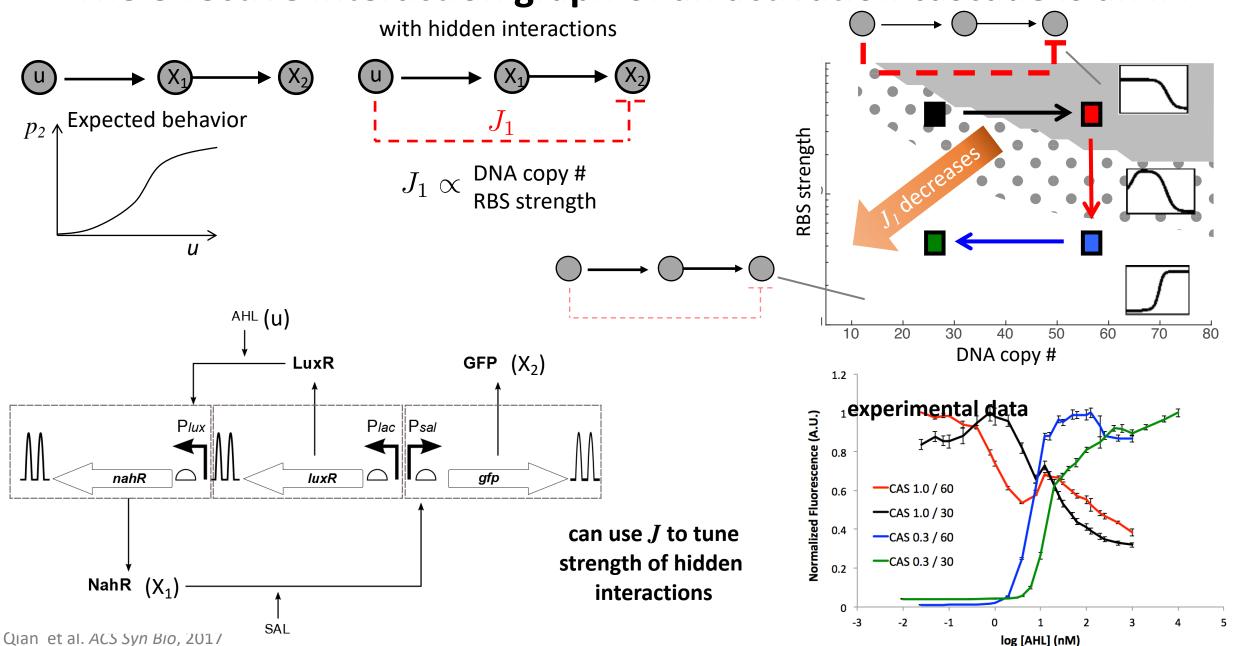


repressors become "effective activators" for non-target nodes

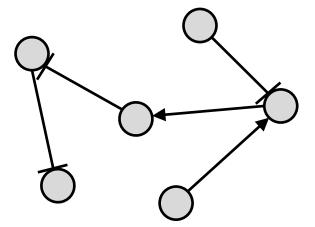


the effect on target nodes is weaker

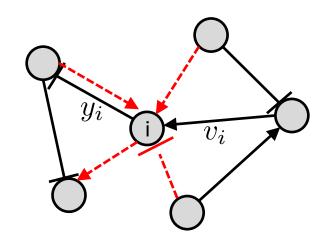
The effective interaction graph of an activation cascade is an iFFL



#### system without hidden interactions

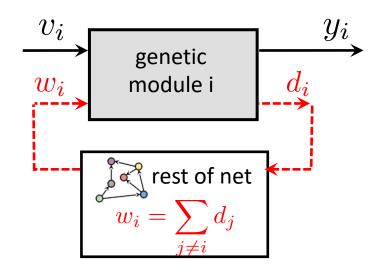


#### system with hidden interactions

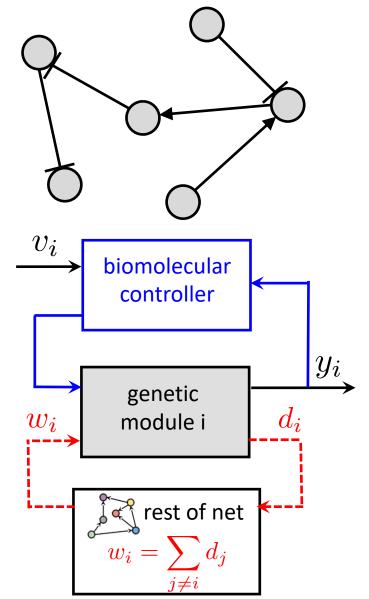


 $d_i \propto J_i$  resource demand at node i

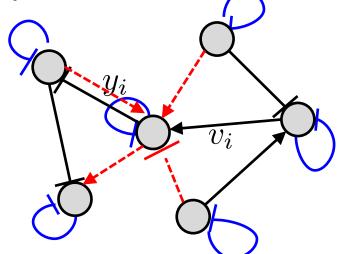
$$w_i = \sum_{j \neq i} d_j$$



#### system without hidden interactions



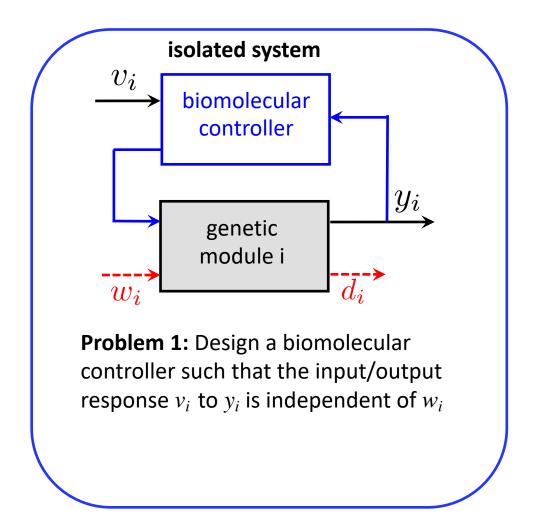
system with hidden interactions

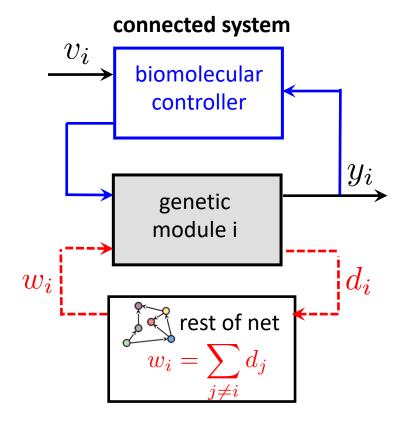


 $rac{d_i \propto J_i}{ ext{resource demand}}$  at node i

$$w_i = \sum_{j \neq i} d_j$$

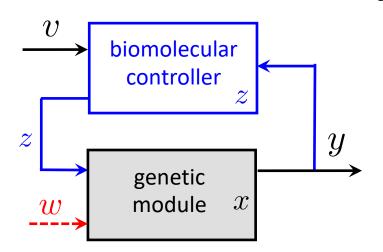
**Problem:** Design a local feedback controller such that  $y_i$  depends only on  $v_i$  and it is independent of  $w_i$ 





**Problem 2:** Determine conditions such that the biomolecular controller can still attenuate the effect of  $w_i$  on  $y_i$  (i.e., ensure network stability)

### Disturbance rejection despite leaky integrators



**Approach:** For *v* and *w* constants, use integral control, e.g.

$$\dot{x} = f(x, v, z, \mathbf{w}), \quad y = g(x)$$
 $\dot{z} = k(v - y)$ 

under stability conditions, y is independent of w at steady state

**Challenge:** molecular decay is unavoidable *in vivo* due to cell growth  $\rightarrow$  integrator leakiness

$$\dot{x} = f(x, z, \mathbf{w}), \quad y = g(x)$$

$$\dot{z} = k(v - y) - \gamma z$$

cannot send growth to zero

→ increase speed of all controller's reactions

$$\dot{x}=f(x,z, extbf{w}), \ y=g(x)$$
 quasi-integral control  $\dot{z}=rac{1}{\epsilon}(v-y)-\gamma z$  (QIC) structure

Theorem: 
$$\dot{x}=f(x,z, \textcolor{red}{w})$$
 
$$\dot{z}_1=\frac{1}{\epsilon}h(v,z,y)-\gamma z_1$$
 
$$\dot{z}_2=\frac{k}{\epsilon}(v-y)-\gamma z_2$$
 
$$y=g(x)$$

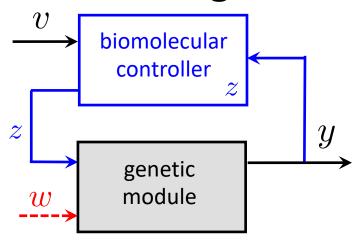
If this closed loop system with  $\gamma=0$  is LES for small  $\epsilon>0$ 

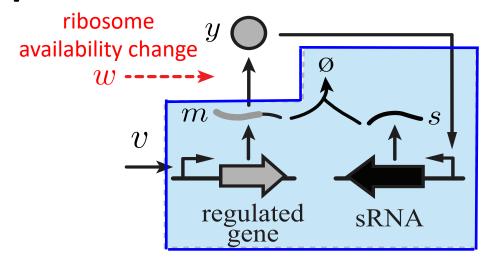
Then:  $y(\epsilon) \to v \text{ as } \epsilon \to 0$  independent of w

#### biomolecular implementation:

all fast controller reactions compute the difference and integrate

### Quasi-integral control implementation via sRNA silencing





Briat et al. (2016) sequestration-based feedback

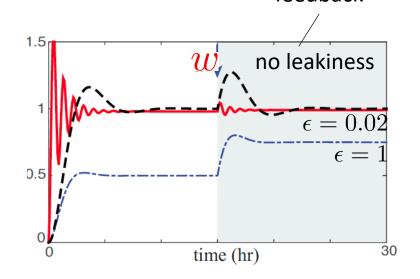
$$\dot{y} = R(\mathbf{w})m - \delta y$$

$$\dot{m} = \begin{cases} \mathbf{v} & \theta \\ -ms & \gamma m \end{cases}$$
fast RNA interactions
$$\dot{s} = \begin{cases} \mathbf{v} & \theta \\ \theta & \gamma s \end{cases} + \gamma s$$
(for "free")

$$z=m-s$$
 
$$\epsilon\dot{z}=(v-y)-\epsilon\gamma z \qquad \text{(QIC)}$$

closed loop system when  $\gamma=0$  is LES

then  $y(\epsilon) \to v \text{ as } \epsilon \to 0$  independent of disturbance



# Tracking performance of quasi-integral control: a SSP problem

**Problem:** With time-varying inputs, can we still attenuate effect of disturbance as timescale separation between

controller and plant increases ( $\epsilon \to 0$ ) ?

#### Not obvious - tempting observation:

$$\dot{y} = R(\mathbf{w})m - \delta y$$

$$\epsilon \dot{m} = v(t) - \theta ms - \epsilon \gamma m$$

$$\epsilon \dot{s} = y - \theta ms - \epsilon \gamma s$$

$$\epsilon = 0 \Rightarrow y(t) = v(t)$$

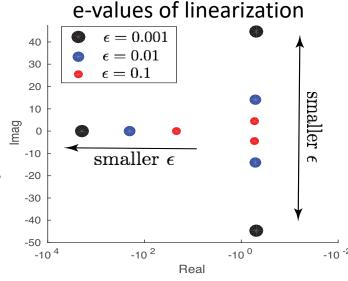
independent of w  $s' = y - \theta ms$ 

Boundary layer dynamics

$$m' = v(t) - \theta ms$$

$$s' = y - \theta ms$$

Jacobian is singular everywhere



#### → Singular singular perturbation (SSP) problem

$$\dot{y} = f(y, x, t), \ y \in \mathbb{R}^{q}$$

$$\epsilon \dot{x} = g(y, x, \epsilon), \ x \in \mathbb{R}^{p}$$

$$y' = \epsilon f(y, x, t)$$

$$x' = g(y, x, \epsilon)$$

$$J = \begin{pmatrix} 0 & 0 \\ \frac{\partial g}{\partial y} & \frac{\partial g}{\partial x} \end{pmatrix} \bigg|_{\epsilon=0}$$

 $\boldsymbol{J}$  has more than q zero e-values

If zero e-value of J has algebraic multiciplity (am) = geometric multiciplity (gm)  $\rightarrow$  system can be taken to standard SP form by  $\epsilon$ -independent coordinate change with less than p fast variables (Gu, Nefedov and O'Malley (1989); Sobolev (2005))

- not applicable here since gm=1 and am=2 (am<gm)
- *Marino & Kokotovic (1988): there is no*  $\epsilon$ -independent diffeomorphism to standard SP form

Asymptotic expansion (with fractional exponents) can address some SSP problems assuming a limiting solution exists as  $\epsilon \to 0$  (O'Malley & Jameson (1975); O'Malley (1979))

- not applicable – no limiting solution exists as  $\epsilon \to 0$ 

### Solving the SSP problem

$$\dot{y} = A_1 \begin{bmatrix} y \\ x \end{bmatrix} + B_1 w(t), \ y \in \mathbb{R}^q$$
 assume  $J = \begin{pmatrix} 0 & 0 \\ A_{21}^0 & A_{22}^0 \end{pmatrix} \begin{cases} -\text{ zero e-value has am} = q+1 \text{ and gm} = q \\ -\text{ all other e-values have negative real part} \end{cases}$  
$$\epsilon \dot{x} = A_2 \epsilon \begin{bmatrix} y \\ x \end{bmatrix} + B_2 v(t), \ x \in \mathbb{R}^p$$

assume 
$$J=\left(egin{array}{cc} 0 & 0 \ A_{21}^0 & A_{22}^0 \end{array}
ight) \left\{egin{array}{cc} ext{-} \end{array}
ight.$$

 $\rightarrow$  There is an  $\epsilon$ -independent coordinate change with S Hurwitz

$$\dot{y} = A_1 \begin{bmatrix} y \\ x \end{bmatrix} + B_1 w(t) \quad slow$$

$$\dot{\epsilon}\dot{z}_1 = Ry + B_2 v(t) + \epsilon Dz \quad \text{one-dimensional} \implies \begin{cases} \text{set } \epsilon = 0 \\ \text{in the fast dynamics} \end{cases}$$

$$\dot{\epsilon}\dot{z}_2 = Sz_2 + B + 3v(t) + \epsilon E \begin{bmatrix} y \\ z_1 \end{bmatrix} \quad \textit{fast}$$

**€**-dependent reduced system

$$\dot{\bar{y}} = \bar{A}_{11}\bar{y} + \bar{A}_{12}\bar{z}_1 + \bar{B}_1 \mathbf{w}(t) + \bar{B}_4 v(t)$$

$$\epsilon \dot{\bar{z}}_1 = R\bar{y} + \bar{B}_2 v(t) + \epsilon \bar{D}\bar{z}_1$$

#### **Theorem (SSP):** Assume that:

- A2. the reduced system is such that  $\ \bar{D} < 0, \ (\bar{A}_{11}, \bar{A}_{12})$  controllable,  $R\bar{A}_{12} > 0$  Then:  $\limsup_{t \to \infty} \|y(t) \bar{y}(t)\| = \mathcal{O}(\sqrt{\epsilon})$

Proof: - decompose the error system into a slow and a fast subsystem

- S Hurwitz  $\rightarrow$  fast subsystem is ISS with gain  $\mathcal{O}(\epsilon)$
- A2.  $\rightarrow$  slow subsystem is ISS with gain  $\mathcal{O}(1/\sqrt{\epsilon})$

result follows from ISS small gain theorem for € sufficiently small

# Solving the robust tracking problem

original system

$$\dot{y} = A_1 \begin{bmatrix} y \\ x \end{bmatrix} + B_1 w(t), \ y \in \mathbb{R}^q$$

$$\epsilon \dot{x} = A_2^{\epsilon} \begin{bmatrix} y \\ x \end{bmatrix} + B_2 v(t), \ x \in \mathbb{R}^p$$

**€**-dependent reduced system

$$\dot{\bar{y}} = \bar{A}_{11}\bar{y} + \bar{A}_{12}\bar{z}_1 + \bar{B}_1 w(t) + \bar{B}_4 v(t) \qquad \limsup_{t \to \infty} ||y(t) - \bar{y}(t)|| = \mathcal{O}(\sqrt{\epsilon})$$

$$\epsilon \dot{\bar{z}}_1 = R\bar{y} + \bar{B}_2 v(t) + \epsilon \bar{D}\bar{z}_1$$

#### Theorem (robust tracking for reduced system):

If, in addition, all input derivatives are bounded, then

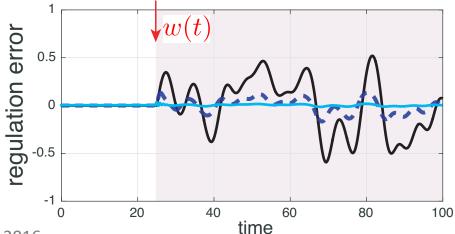
$$\limsup_{t \to \infty} ||R\bar{y}(t) + \bar{B}_2 v(t)|| = \mathcal{O}(\sqrt{\epsilon})$$

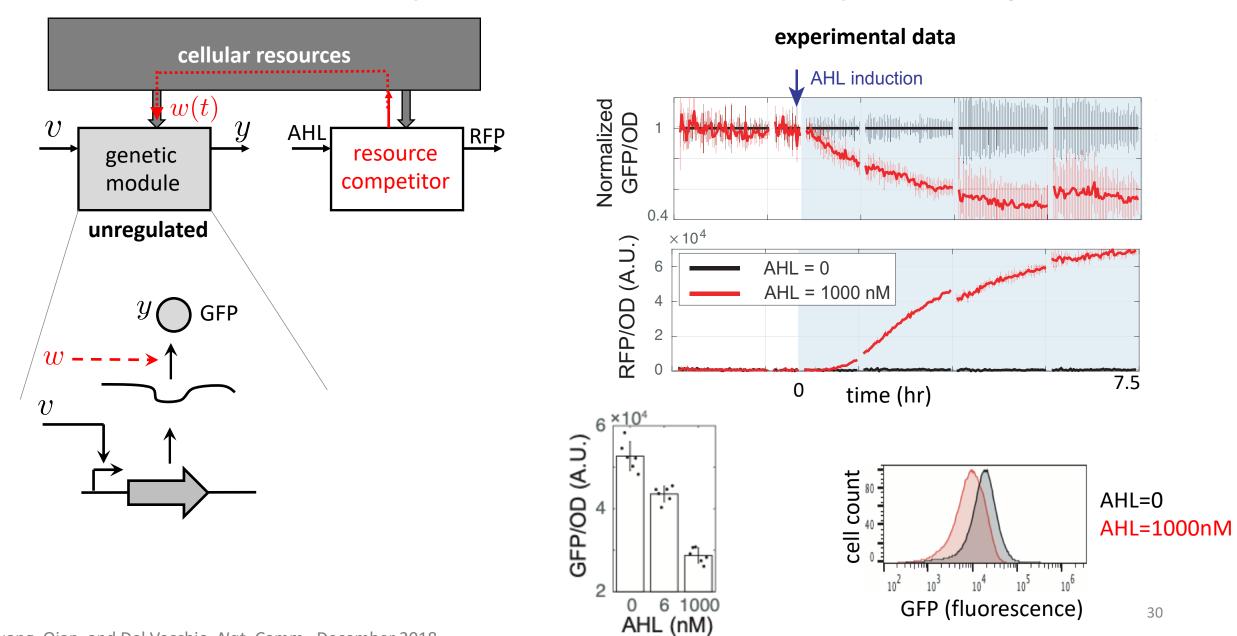
$$\rightarrow \limsup_{t \to \infty} \|y(t) + R^{-1}\bar{B}_2v(t)\| = \mathcal{O}(\sqrt{\epsilon}) \qquad \Rightarrow y(t) \text{ independent of } \frac{w(t)}{v(t)} \text{ as } \epsilon \to 0$$

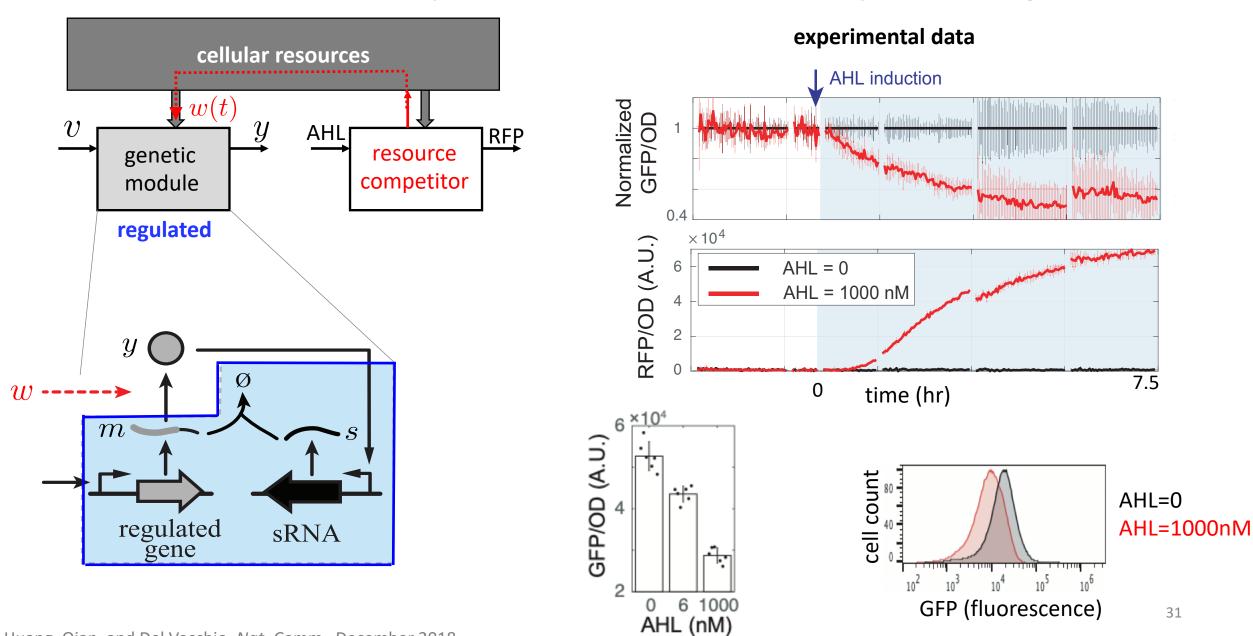
$$\dot{y} = R(\mathbf{w})m - \delta y$$

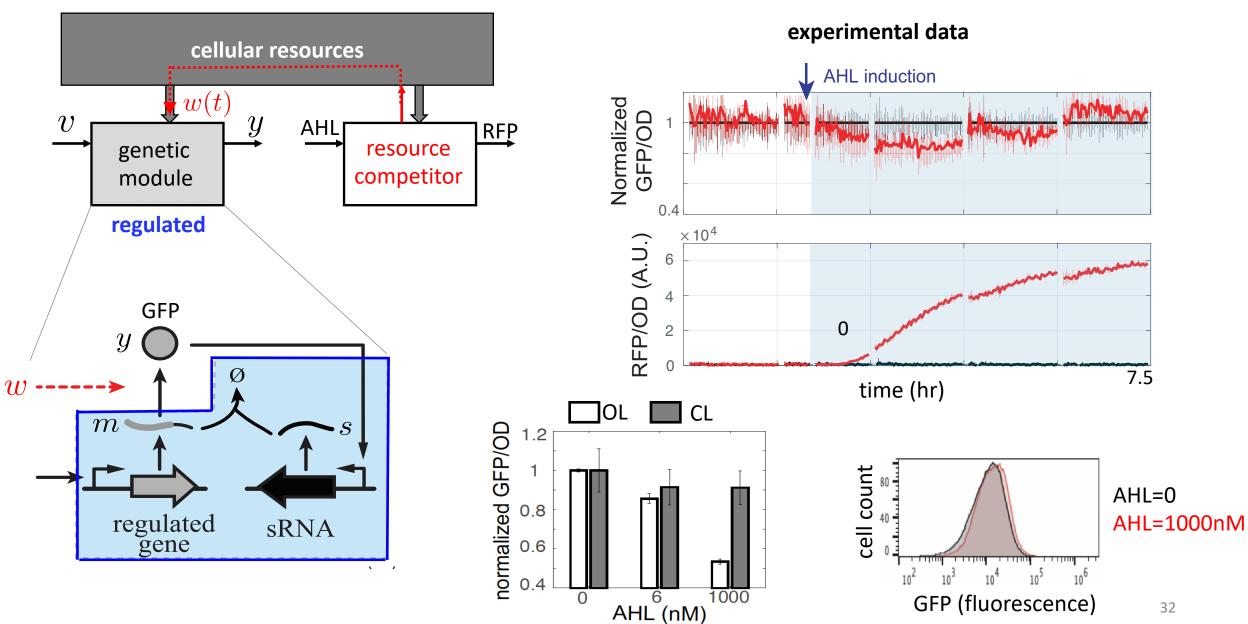
$$\epsilon \dot{m} = v(t) - \theta ms - \epsilon \gamma m$$

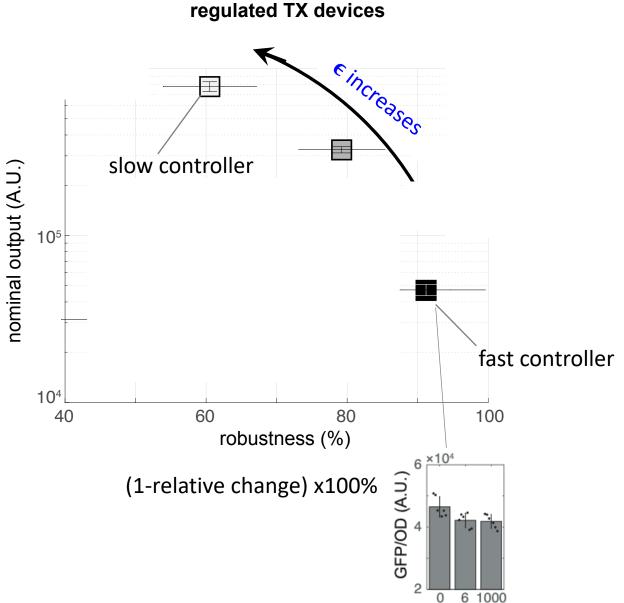
$$\epsilon \dot{s} = y - \theta ms - \epsilon \gamma s$$



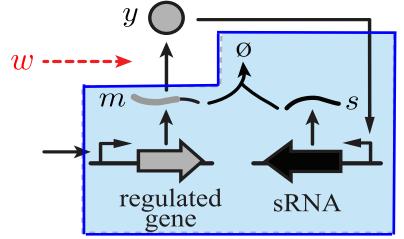


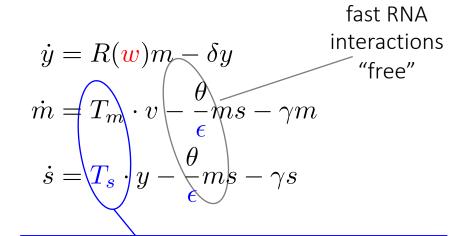






AHL (nM)





high RNA TX rates

 $T_s$  easily tunable by sRNA promoter  $T_m$  GFP promoter strength

### A journey towards modular composition

**Engineering biology: Why and how** 

Modular composition: A grand challenge

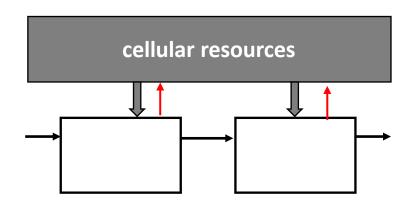
Inter-module loads and the load driver → upstream system odownstream system

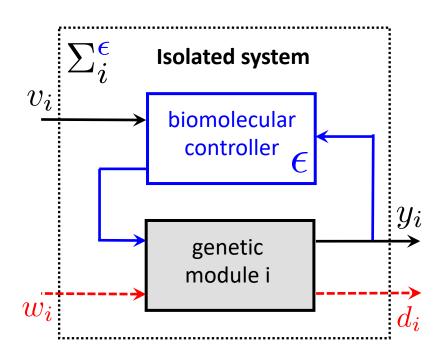
### Resource loading and the resource decoupler

Disturbance rejection despite leaky integral actions

Decentralized implementation

Outlook

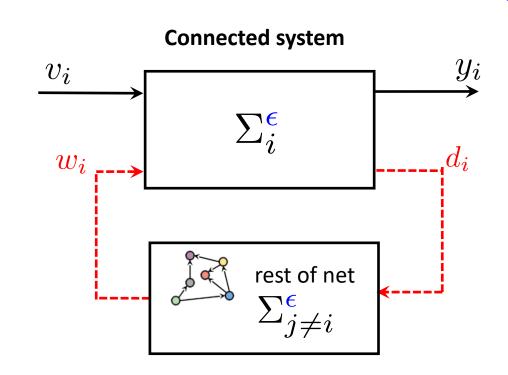




We have: for constant  $w_i$  bounded independent of  $\epsilon$ 

$$\lim_{t \to \infty} \|y_i(t) - h(v_i)\| = \mathcal{O}(\sqrt{\epsilon}) \|w_i\|$$

 $\rightarrow y_i$  independent of  $w_i$ 

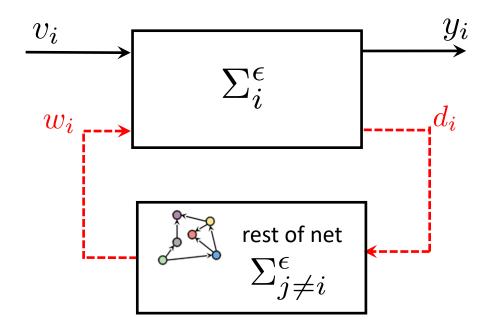


**Problem:** Can we guarantee that

$$\lim_{t \to \infty} \|y_i(t) - h(v_i)\| = \mathcal{O}(\sqrt{\epsilon})$$

- ensure that  $w_i$  steady state has  $\epsilon$ -independent bound
- ensure closed loop system approaches steady state

#### ensure that $w_i$ has an $\epsilon$ -independent steady state bound



physics leads to steady state relationships:

(i) system 
$$i$$
  $d_i = g_i(v_i) + \hat{g}_i(v_i) \cdot w_i + \tilde{g}_i(w_i) \cdot O(\epsilon)$  
$$/$$
 Innear term in  $w_i$  HOT in  $w_i$ 

(ii) interconnection 
$$w_i = \sum_{j \neq i} d_j$$

(iii) system (i)+(ii) 
$$\rightarrow A(v) \cdot w = g(v) + \tilde{g}(w)\mathcal{O}(\epsilon)$$

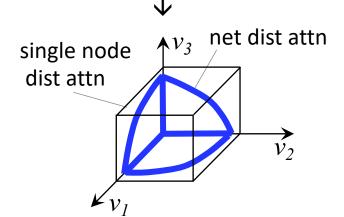
$$A(v) = \begin{bmatrix} 1 & -\hat{g}_2 & \cdot & \cdot & -\hat{g}_n \\ -\hat{g}_1 & 1 & -\hat{g}_2 & \cdot & -\hat{g}_n \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ -\hat{g}_1 & -\hat{g}_2 & \cdot & -\hat{g}_{n-1} & 1 \end{bmatrix}$$

ightarrow If A(v) is invertible, then w has  $\epsilon$ -independent bound

sufficient check: diagonal dominance  $\sum_{j \neq i} \hat{g}_j(v_j) < 1, \ \ \forall \ i$ 

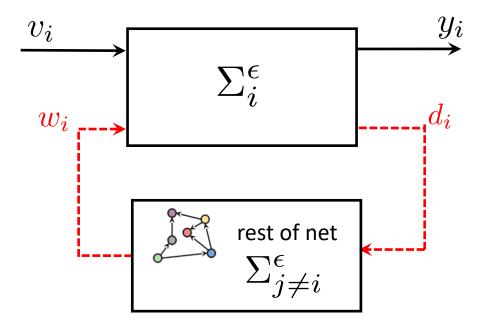
$$\sum_{j \neq i} \hat{g}_j(v_j) < 1, \ \forall i$$

 $\hat{g}_j(v_j)$  increasing function of  $v_j$ 



obtain constraints on tunable parameters as a function of the number of nodes

#### ensure closed loop system approaches steady state



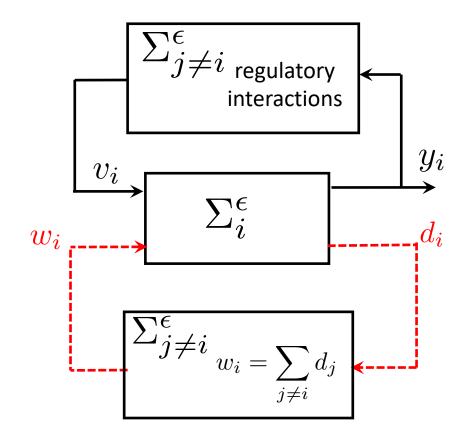
**Assumption:**  $\Sigma_i^{\epsilon}$  is input-to-state/output monotone

→ use Small Gain Theorem for Monotone Systems (Angeli & Sontag, IEEE TAC 2003)

A(v) diagonally dominant  $\rightarrow$  unique/globally attractive equilibrium

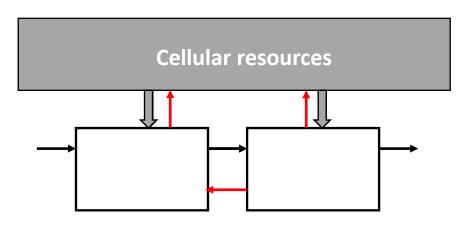
Note: we can prove the controller makes  $\sum_i^{\epsilon}$  SP - monotone if its dynamics are much faster than the plant (Grunberg and Del Vecchio, IEEE CDC 2019)

**on-going:** time-varying inputs and regulatory interactions



In preparation – ingredients: small-gain theorem for singularly perturbed monotone systems (Christofides & Teel, IEEE TAC 1995; Angeli & Sontag, IEEE TAC 2003)

### **Summary**

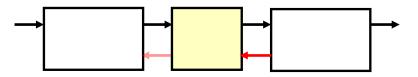


Loads applied by downstream modules change the behavior of upstream systems

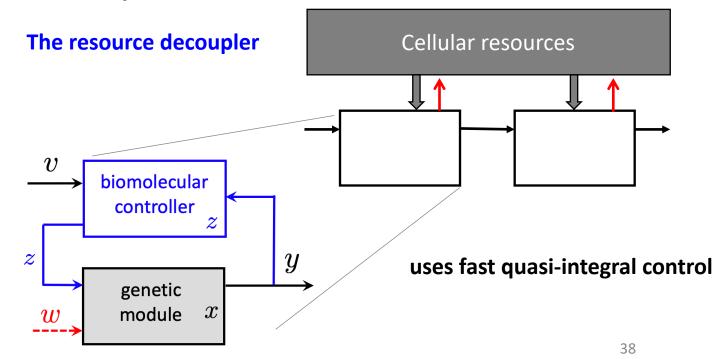
Loads that modules apply to cellular resources cause subtle couplings among theoretically independent modules

An engineering framework for insulating genetic modules from perturbations

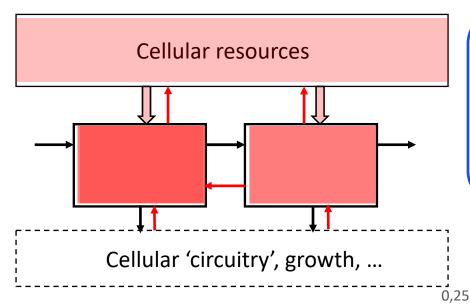




uses time scale separation in place of high-gain negative feedback



### Some reasons why modularity is a challenge



Loads applied by downstream modules change the behavior of upstream systems

(Del Vecchio, Hespanha, Klavins, Papachristodoulou, Sontag, ...)

Modules apply a load the cellular resources: creates subtle couplings (Bates, Del Vecchio, Murray, Stan,...)

Modules often have "off-target" interactions, affect growth rate, and this, in turn, has global effects on a module's dynamics

source: Web of Sci

(Khammash, Papachristodoulou, Stan, ...)

# articles in biology journals

- lab conditions: temperature, nutrients, Ph,...

- cell type/strain

- growth phase

# articles in biology journals

control and synthetic biology
synthetic biology

Aoki et al. *Nature* 2019
Olsman et al. *Cell Systems* 2019
Chevalier et al. *Cell Systems* 2019
Agrawal et al. *Nat Comm* 2019
Kelly et al. *NAR* 2018
Agrawal et al. *ACS Syn Bio* 2018
Darlington et al. *Nat Comm* 2018
Huang et al. *Nat Comm* 2018
Ceroni et al. *Nat Methods* 2018

- mutations

39

Former Students/ Post-docs:

Andras Gyorgy (NYU, Abu Dhabi)

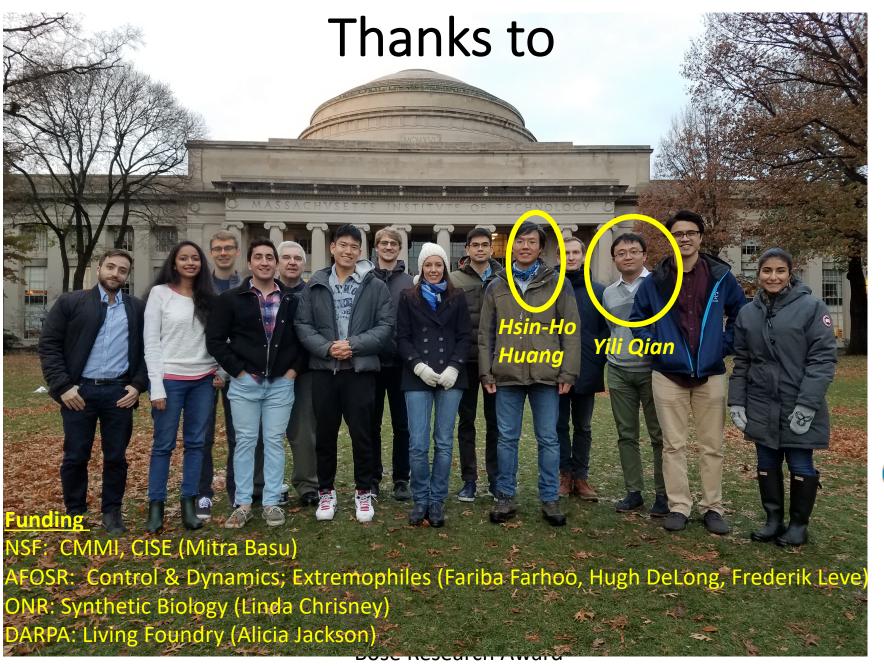
Hattie Chang (Harvard)

Jose Jimenez (U. of Surrey)

John Yazbek







**Collaborators:** 

Eduardo Sontag (NEU)

Ron Weiss (MIT)

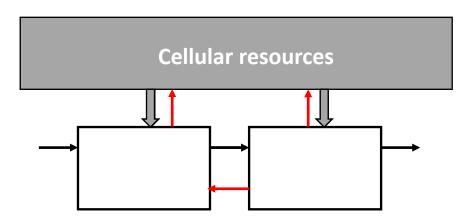
Jim Collins (MIT)

Richard Murray (Caltech)





### **Summary**

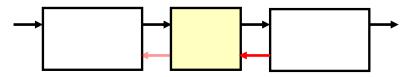


Loads applied by downstream modules change the behavior of upstream systems

Loads that modules apply to cellular resources cause subtle couplings among theoretically independent modules

An engineering framework for insulating genetic modules from perturbations





uses time scale separation in place of high-gain negative feedback

