Robustness of reprogramming strategies to small perturbations*

Rushina Shah and Domitilla Del Vecchio

Abstract— **Cell reprogramming is the process of using external inputs to induce a transition from one cell type to another. This problem is analyzed theoretically through the study of multistable gene regulatory network motifs that control cell fate decision making. Here, cell phenotypes are modeled as different stable steady states of the network, and inputs for reprogramming are selected to trigger a transition between one stable steady state to another. These core motifs are embedded in larger gene regulatory networks that can be viewed as adding a perturbation to the dynamics of the motifs, thus potentially changing the qualitative response to inputs. There is therefore a need to determine whether strategies developed for these core motifs are robust to perturbations applied to the motif by the rest of the network. In this work, we determine conditions under which reprogramming strategies that reprogram a core motif are robust to small perturbations. Under these conditions, these strategies are still guaranteed to reprogram the motif when it is embedded in the larger network. Thus, reprogramming strategies designed by the analysis of core motifs serve as a practical way to reprogram the motifs when they are part of a larger network.**

I. Introduction

Cell differentiation is the process by which a cell changes from one cell type to another, for instance, the process by which embryonic cells give rise to specialized, adult cells such as skin cells, blood cells or muscle cells [1]. In 2006, Shinya Yamanaka and co-workers showed that this process could be reversed by the induction of external inputs [2]. Cell fate reprogramming refers to the artificial process in which external inputs, such as the overexpression and/or enhanced degradation of key proteins, are introduced to convert cells to a desired cell type. In the past two decades, several successful attempts at cell reprogramming have been seen [3]-[8], albeit with very low efficiency [9]. Selecting inputs for these attempts have relied chiefly on trial-and-error experimentation, after identifying the key proteins that control cell fate.

Recently, advances have been made to approach this problem theoretically and/or computationally in conjunction with experiments. This is typically done by analyzing the core gene regulatory motifs that control cell fate decision-making processes, several of which have been experimentally characterized such that at least their structure is known [10]-[13]. Examples include the Oct-Sox-Nanog motif that controls the maintenance of pluripotency [11] and the PU-GATA motif that controls the differentiation of the hematopoetic stem cell [13].

Multistability is a common feature of core network motifs implicated in cell fate decisions [14]. Here, each stable steady state represents a different cell phenotype. Under this framework, selecting inputs for cell reprogramming involves choosing inputs to induce a transition in the system state from one stable steady state to another. Theoretical studies to design inputs to reprogram such multistable systems include [15]-[17]. In [15], a feedback controller is proposed to induce pluripotency, by analyzing a model of the core motif controlling pluripotency. In [16], general strategies are provided to reprogram core motifs that are cooperative. In [17], bistable monotone systems are studied to design pulse-like inputs to switch states. Computational approaches to study this problem typically involve using bifurcation tools and parameter sampling-based methods to analyze core motifs [18], [19]. These studies provide systematic strategies to approach the reprogramming problem, and potentially reduce the time and effort needed to design inputs for the same.

However, while these works rely on the structure and key proteins in the core motifs, this core network is typically embedded in a larger gene regulatory network (GRN) [20], whose structure is not always known, and further, might not share the properties, for example cooperativitiy as used in [16], of the core motif. The rest of the GRN applies a perturbation to the dynamics of the core motif, potentially changing its response to the inputs designed by the techniques described above. Thus, there is a need to evaluate the robustness of reprogramming strategies to the perturbations applied to the core motif by the rest of the GRN. In this work, we analyze reprogramming strategies designed to reprogram the core motif in the following way. A constant input is applied to the system, causing the stimulated system to converge within the region of attraction of the desired steady state. Once this constant input is removed, the unstimulated system then converges to the desired state. This type of transient input has been used in [15], [16] to reprogram core motifs to desired stable steady states. We provide conditions under which, for sufficiently small perturbations, inputs designed to reprogram the core motif in such a fashion are guaranteed to also reprogram the large network close to the desired stable steady state.

This paper is organized as follows. Section II describes the system and states the problem definition. Section III states the results of this paper, and these are demonstrated using the pluripotency network in Section IV.

*This work was supported in part by NIH Grant number 1-R01-EB024591-01.

Rushina Shah and Domitilla Del Vecchio are with the Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. Email addresses: rushina@mit.edu (R. Shah) ddv@mit.edu (D. Del Vecchio)

II. System description and problem definition

We consider a dynamical system Σ_w of the form:

$$\Sigma_w : \dot{x} = f(x, w), \tag{1}$$

where state $x \in X \subset \mathbb{R}^n_+$ and a constant input vector $w \in W \subset \mathbb{R}^q$. Let $\mathbb{S}$ be the set of stable steady states of the system $\Sigma_0 : \dot{x} = f(x, 0)$. Further, we let $\mathcal{R}_w(S)$ denote the region of attraction of a stable steady state S of system Σ_w. Trajectories of this system starting at $x_0 \in X$ at time $t \geq 0$ are represented by $\phi_w(t, x_0)$. The ω-limit set for a point $x_0 \in X$, is denoted by $\omega_w(x_0)$.

This dynamical system Σ_w represents the isolated dynamics of a core gene regulatory network motif that controls cell fate. Previous works [15], [16], [17] have devised strategies to reprogram such core motifs, that is, design transient stimulations to trigger a transition from the inital state of the system to a new stable steady state. Although the study of the dynamics of the motif is indicative of the effects of external inputs on the system's steady states, this motif is always embedded in a larger network of interactions. It is therefore of interest to investigate the extent to which inputs that reprogram the core motif to a desired state when it is in isolation, also reprogram the motif to a vicinity of this desired state once the motif is embedded into a larger network. We therefore represent the dynamics of this extended GRN by:

$$\Sigma^p_w : \begin{aligned} \dot{x} &= f(x, w) + ph(x, y), \\ \dot{y} &= g(x, y), \end{aligned} \tag{2}$$

where $y \in Y \subset \mathbb{R}^m_+$ represents the state of the species of the GRN not included in the core motif. The term $ph(x, y)$ represents the perturbation to the dynamics of x due to the rest of the network, where the function $h : \mathbb{R}^{n+m}_+ \to \mathbb{R}^n_+$, and $p \in P \subset \mathbb{R}_+$. When the function $h(x, y)$ is bounded, the scalar p modulates the strength of the perturbation applied to the dynamics of x. Trajectories of the system Σ^p_w starting at $r' = (x', y') \in X \times Y$ at time $t \geq 0$ are represented by $\phi^p_w(t, r')$. The ω-limit set for a point $r' \in X \times Y$, is denoted by $\omega^p_w(r')$.

We denote by Σ^0_w the system Σ^p_w with $p = 0$, that is $\Sigma^0_w : (\dot{x}, \dot{y}) = (f(x, w), g(x, y))$. This system represents a cascade, where the core motif drives the rest of the network, but the rest of the network has no effect on the core motif. Trajectories of this system starting at $z' \in X \times Y$ at $t \geq 0$ are represented by $\phi^0_w(t, z')$. The region of attraction of a stable steady state S' of Σ^0_w is represented by $\mathcal{R}^0_w(S')$.

In this paper, we wish to investigate conditions under which inputs selected based on the analysis of a core motif Σ_w are effective in reprogramming the extended network Σ^p_w under sufficiently small perturbation p. For this, we first formally define the following notions of reprogramming the core motif Σ_0.

Definition 1: We say that the system Σ_0 is *strongly reprogrammable* under a constant input w^* to a steady state $S^* \in \mathbb{S}$ provided that the system Σ_{w^*} is such that, for all initial conditions $x_0 \in X$, $\omega_{w^*}(x_0) \subset \mathcal{R}_0(S^*)$.

We call the triplet (w^*, Σ_0, S^*) a strong reprogramming strategy when the input w^* can strongly reprogram system Σ_0 to steady state S^*.

Definition 2: We say that the system Σ_0 is *weakly reprogrammable* from a given steady state $S^*_1 \in \mathbb{S}$ to a steady state $S^*_2 \in \mathbb{S}$, under a constant input w^*, provided that the omega-limit set of S^*_1, $\omega_{w^*}(S^*_1) \subset \mathcal{R}_0(S^*_2)$.

We call the quadruplet $(w^*, \Sigma_0, S^*_1, S^*_2)$ a weak reprogramming strategy when the input w^* can weakly reprogram system Σ_0 from steady state S^*_1 to steady state S^*_2.

Consider an input w^* that can reprogram the system Σ_0 to some steady state $S^* \in \mathbb{S}$. In this work, we wish to determine conditions under which the same input w^* can reprogram system Σ^p_0 such that it remains "close" to S^*. Since we are comparing vectors in $\mathbb{R}^n$ (state of Σ_w) to vectors in $\mathbb{R}^{n+m}$ (state of Σ^p_w), we define the following projections. The projection $proj_X : \mathbb{R}^{n+m} \to \mathbb{R}^n$ is such that $proj_X z = [\mathbb{I}_{n\times n}\mathbf{0}_{n\times m}]z$, and the projection $proj_Y : \mathbb{R}^{n+m} \to \mathbb{R}^m$ is such that $proj_Y z = [\mathbf{0}_{m\times n}\mathbb{I}_{m\times m}]z$. Using this, we formally define the notion of "close" for states of these two systems as follows:

Definition 3: State $z \in X \times Y \subset \mathbb{R}^{n+m}_+$ is said to be in the ϵ-neighborhood of $x \in X \subset \mathbb{R}^n_+$ (represented by $z \in N_\epsilon(x)$) if $||proj_X z - x|| < \epsilon$.

Then, we define the notion of robustness for a reprogramming strategy that reprograms Σ_0 to a steady state S^*.

Definition 4: A strong reprogramming strategy (w^*, Σ_0, S^*) is *robust to small perturbations* if, for every $\epsilon > 0$, there exists $p^* > 0$ such that for all $p \leq p^*$, and for all initial conditions $r' \in X \times Y$, the ω-limit set $\omega^p_{w^*}(r')$ is such that for any $r \in \omega^p_{w^*}(r')$, we have that $\omega^p_0(r) \subseteq N_\epsilon(S^*)$.

In other words, for the strategy (w^*, Σ_0, S^*) to be robust to small perturbations, the following must be true. For Σ^p_0 starting at any initial condition $r' \in X \times Y$, the ω-limit set when input w^* is applied is such that, sufficiently long after the input is removed, all trajectories starting from this set must remain ϵ-close to S^*.

Definition 5: A weak reprogramming strategy $(w^*, \Sigma_0, S^*_1, S^*_2)$ is *robust to small perturbations* if, for every $\epsilon_2 > 0$, there exists $\epsilon_1 > 0$ and $p^* > 0$ such that for all $p \leq p^*$, and for all $r' \in N_{\epsilon_1}(S^*_1)$, the ω-limit set $\omega^p_{w^*}(r')$ is such that for any $r \in \omega^p_{w^*}(r')$, we have that $\omega^p_0(r) \subseteq N_{\epsilon_2}(S^*_2)$.

In other words, for $(w^*, \Sigma_0, S^*_1, S^*_2)$ to be robust to small perturbations, the following must be true. For system Σ^p_0 starting sufficiently (ϵ_1) close to S^*_1, and for a sufficiently small perturbation ($p \leq p^*$), the ω-limit set when input w^* is applied is such that, all trajectories starting from this ω-limit set remain ϵ_2 close to S^*_2 sufficiently long after the input is removed.

Finally, we make the following assumptions about systems Σ_w and Σ^p_w.

Assumption 1: Functions $f(x), g(x, y), h(x, y)$ are continuously differentiable on all $x, y \in X \times Y$.

Assumption 2: For fixed $x = x^* \in X$, system $\dot{y} = g(x^*, y)$ has globally exponentially stable steady state, $\bar{y}(x^*)$.

Assumption 3: The trajectories of the systems Σ_w, Σ_w^p are bounded for $t \geq 0$ for any given $w \in W \cup \{0\}$ and $p \in P$.

Assumption 4: The Jacobian at the stable steady states for the systems Σ_w, Σ_w^0 for any fixed $w \in W \cup \{0\}$ is non-singular.

Assumption 5: The perturbation function $h(x, y)$ is bounded, i.e., $||h(x, y)|| \leq h_B$, for all $x, y \in X \times Y$.

III. Results

In this section, we present results demonstrating that for systems satisfying Assumptions 1-4, inputs selected to reprogram Σ_0 are robust to small perturbations. The main results of this paper are Theorems 1 and 2.

Lemma 1 provides conditions on the system Σ_w such that the system Σ_w^0, the connected system with $p = 0$, has certain stability properties.

Lemma 1: For systems Σ_w and Σ_w^0, the following hold under Assumptions 1-4:

(i) Consider $w^* \in W \cup \{0\}$ such that Σ_{w^*} has a (globally) exponentially stable steady state x_w^*. Then, system $\Sigma_{w^*}^0$ has a (globally) exponentially stable steady state $z_w^* = (x_w^*, \bar{y}(x_w^*))$.

(ii) Consider $x' \in X$ such that $x' \in \mathcal{R}_w(x_w^*)$ where x_w^* is some steady state of Σ_w. Then, $z' = (x', y') \in \mathcal{R}_w^0(z_w^*)$ for any $y' \in Y$, where $z_w^* = (x_w^*, \bar{y}(x_w^*))$.

Proof: (i) The claim of (global) asymptotic stability of z_w^* follows from Corollaries 10.3.2 and 10.3.3 of [21]. Since the Jacobian of Σ_w^0 at the steady state z_w^* is non-singular under Assumption 4, all eigenvalues of this Jacobian have negative real parts, and therefore z_w^* is also (globally) exponentially stable.

(ii) This follows from Theorem 10.3.1 of [21], since $\bar{y}(x_w^*)$ is a globally exponentially stable steady state of the system $\dot{y} = g(x_w^*, y)$ under Assumption 2. ∎

Lemma 2: Let Assumptions 1 and 5 hold. Consider $z' \in \mathcal{R}_w^0(z_w^*)$, where z_w^* is an exponentially stable steady state of Σ_w^0 for some $w \in W \cup \{0\}$. For any such z', and any $\epsilon > 0$, there exist $\delta > 0$ and $p^* > 0$ such that for any $r' \in B_\delta(z')$, for all $p \leq p^*$, $\omega_w^p(r') \subseteq B_\epsilon(z_w^*)$.

Proof: Consider the system Σ_w^0, and the change of variables $\tilde{z} = (\tilde{x}, \tilde{y}) = ((x, y) - z_w^*)$. The dynamics of this transformed system are $\dot{\tilde{z}} = F(\tilde{z}) = (\tilde{f}(\tilde{x}, w), \tilde{g}(\tilde{x}, \tilde{y}))$, where $\tilde{f}(\tilde{x}, w) = f(\tilde{x} + proj_X z_w^*, w)$ and $\tilde{g}(\tilde{x}, \tilde{y}) = g(\tilde{x} + proj_X z_w^*, \tilde{y} + proj_Y z_w^*)$. For this transformed system, $\tilde{z} = 0$ is an exponentially stable steady state, and $\tilde{z}' = z' - z_w^*$ lies in the region of attraction $\tilde{\mathcal{R}}_w^0(0)$. By Theorem 4.17 from [22] (valid under Assumption 1), there exists a positive definite function $V(\tilde{z})$, and a continuous positive definite function $W(\tilde{z})$, both defined for all $\tilde{z} \in \tilde{\mathcal{R}}_w^0(0)$, such that $V(\tilde{z}) \to \infty$ as $\tilde{z} \to \partial\tilde{\mathcal{R}}_w^0(0)$, $\frac{\partial V}{\partial \tilde{z}} F(\tilde{z}) \leq -W(\tilde{z})$, $\forall \tilde{z} \in \tilde{\mathcal{R}}_w^0(0)$, and for any $c > 0$, $\{V(\tilde{z}) \leq c\}$ is a compact subset of $\tilde{\mathcal{R}}_w^0(0)$.

Consider the perturbed system Σ_w^p under the same change of coordinates, and we refer to its state as $\tilde{r}(t) = ((x, y) - z_w^*)$. The dynamics of this system are given by $\dot{\tilde{r}} = F(\tilde{r}) + G(\tilde{r})$, where $G(\tilde{r}) = (p\tilde{h}(\tilde{r}), 0)$, and under Assumption 5 $||G(\tilde{r})|| \leq ph_B$. Then, the hypotheses of Theorem 9.1 of [22] are satisfied. Then, by this theorem, for every compact set $\Omega \subset \{V(\tilde{z}) \leq \rho c, 0 < \rho < 1\}$, there exist constants β, γ, η, μ and k independent of ph_B such that if $\tilde{z}(0) = \tilde{z}' \in \Omega$, $ph_B < \eta$, and $||\tilde{r}(0) - \tilde{z}(0)|| < \mu$, then $||\tilde{r}(t) - \tilde{z}(t)|| \leq ke^{-\gamma t}||\tilde{r}(0) - \tilde{z}(0)|| + \beta ph_B$ for all $t \geq 0$. Choose $\delta = \mu$, $p^* < \min\{\frac{\eta}{h_B}, \frac{\epsilon}{\beta h_B}\}$, $\rho c > V(\tilde{z}')$, and $\Omega = \{V(\tilde{z}) \leq V(\tilde{z}')\}$. Then, $\tilde{z}(0) = \tilde{z}' \in \Omega \subset \{V(\tilde{z}) \leq \rho c\}$, and $ph_B < \eta$ for all $p \leq p^*$. Then, for all $\tilde{r}' \in B_\mu(\tilde{z}') = B_\delta(\tilde{z}')$, we have that $||\tilde{r}(t) - \tilde{z}(t)|| \leq ke^{-\gamma t}\delta + \beta ph_B$ for all $t \geq 0$. Considering the systems in the original coordinates, we have that for all $r' \in B_\delta(z')$, $||\phi_w^p(t, r') - \phi_w^0(t, z')|| \leq ke^{-\gamma t}\delta + \beta ph_B < ke^{-\gamma t}\delta + \epsilon$, since $p \leq p^* < \frac{\epsilon}{\beta h_B}$.

Consider the ω-limit set $\omega_w^p(r')$. By definition of $\omega_w^p(r')$, for any point $r_\omega \in \omega_w^p(r')$, there exists a sequence $(t_i)_{i \in \mathbb{N}}$ such that $\lim_{i \to \infty} t_i \to \infty$, and $\lim_{i \to \infty} \phi_w^p(t_i, r') = r_\omega$. We know that there exist $p^* > 0$ and $\delta > 0$ such that for $p \leq p^*$ and $r' \in B_\delta(z')$, for all $t_i \geq 0$, $||\phi_w^p(t_i, r') - \phi_w^0(t_i, z')|| \leq ke^{-\gamma t_i}\delta + \epsilon$. Taking the limit as $i \to \infty$, we have $\lim_{i \to \infty} \phi_w^p(t_i, r') = r_\omega$, $\lim_{i \to \infty} \phi_w^0(t_i, z') = z_w^*$ since $z' \in \mathcal{R}_w^0(z_w^*)$. Then, for any $r_\omega \in \omega_w^p(r')$, where $r' \in B_\delta(z')$, we have that $||r_\omega - z_w^*|| < \epsilon$. Thus, for any $\epsilon > 0$, there exist a $p^* > 0$ and a $\delta > 0$ such that for $p \leq p^*$ and $r' \in B_\delta(z')$, $\omega_w^p(r') \subseteq B_\epsilon(z_w^*)$. ∎

Theorem 1: Let Assumptions 1-5 hold. Consider stable steady state $S^* \in \mathbb{S}$ of system Σ_0. Supppose that there is an input $w^* \in W$ such that Σ_{w^*} has a globally exponentially stable steady state $x_w^* \in \mathcal{R}_0(S^*)$. Then, the strong reprogramming strategy (w^*, Σ_0, S^*) is robust to small perturbations in the sense of Definition 4.

Proof: Consider a globally exponentially stable steady state x_w^* of system Σ_{w^*} such that $x_w^* \in \mathcal{R}_0(S^*)$. From Lemma 1, there exists $S' = (S^*, \bar{y}(S^*))$ that is an exponentially stable steady state of Σ_0^0, and there exists a globally exponentially stable steady state $z_w^* = (x_w^*, \bar{y}(x_w^*))$ of $\Sigma_{w^*}^0$ such that $z_w^* \in \mathcal{R}_0^0(S')$. So, any $z' \in X \times Y$ is such that $z' \in \mathcal{R}_w^0(x_w^*)$.

By Lemma 2, for any $\epsilon_1 > 0$, there exists a $p_1^* > 0$, such that for any $r' = z' \in X \times Y$, we have that $\omega_{w^*}^p(r') \subseteq B_{\epsilon_1}(z_w^*)$. Consider an $r \in \omega_{w^*}^p(r')$. By Lemma 1, we have $z_w^* \in \mathcal{R}_0^0(S')$, where $S' = (S^*, \bar{y}(S^*))$. Then, under Lemma 2, for any $\epsilon > 0$, there exists a $\delta > 0$ and a $p_2^* > 0$ such that for any $r \in B_\delta(z_w^*)$ and for all $p \leq p_2^*$, we have $\omega_0^p(r) \subseteq B_\epsilon(S')$. Choose $\epsilon_1 < \delta$ so that $r \in B_{\epsilon_1}(z_w^*) \subset B_\delta(z_w^*)$, and choose $p^* = \min\{p_1^*, p_2^*\}$. Then, for any $\epsilon > 0$, there exists a $p^* > 0$ such that for all $p \leq p^*$, the following is true. For all $r' \in X \times Y$, for any $r \in \omega_{w^*}(r')$, we have that $\omega_0^p(r) \subseteq B_\epsilon(S')$. Since $S' = (S^*, \bar{y}(S^*))$, we also have that $B_\epsilon(S') \subseteq N_\epsilon(S^*)$, and thus $\omega_0^p(r) \subseteq N_\epsilon(S^*)$. ∎

Theorem 2: Let Assumptions 1-5 hold. Consider two stable steady states S_1^* and S_2^* of system Σ_0. Suppose that there is an input w^* such that system Σ_{w^*} has an exponentially stable steady state x_w^* such that $S_1^* \in \mathcal{R}_{w^*}(x_w^*)$ and $x_w^* \in \mathcal{R}_0(S_2^*)$. Then the weak reprogramming strategy $(w^*, \Sigma_0, S_1^*, S_2^*)$ is robust to small perturbation in the sense of Definition 5.

Proof: By Lemma 1, we have that the system Σ_0^0 has an exponentially stable steady state $S_2' = (S_2^*, \bar{y}(S_2^*))$, and

the system $\Sigma^0_{w^*}$ has an exponentially stable steady state $z^*_w = (x^*_w, \bar{y}(x^*_w))$. Since $x^*_w \in \mathcal{R}_0(S^*_2)$, by Lemma 1, $z^*_w \in \mathcal{R}^0_0(S'_2)$.

Further, since $S^*_1 \in \mathcal{R}_{w^*}(x^*_w)$, by Lemma 1, for any $y' \in Y$, we have that $z' = (S^*_1, y') \in \mathcal{R}^0_{w^*}(z^*_w)$. Then, by Lemma 2, for any $\epsilon' > 0$, there exists an $\epsilon_1 > 0$ and a $p^*_1 > 0$ such that for any $p \leq p^*_1$, and $r' \in B_{\epsilon_1}(z')$, we have that $\omega^p_{w^*}(r') \in B_{\epsilon'}(z^*_w)$. We show that this holds for any $r' \in N_{\epsilon_1}(S^*_1)$, since this holds for any $r' \in B_{\epsilon_1}(z')$ for any $z' = (S^*_1, y')$. Pick $y' = proj_Y r'$. Then, for $||r' - z'|| < \epsilon_1$ implies that $||proj_X r' - S^*_1|| < \epsilon_1$ must be true since $proj_Y r' = proj_Y z' = y'$. Thus, for any $\epsilon' > 0$, there exists $\epsilon_1 > 0$ and $p^*_1 > 0$ such that for any $p \leq p^*_1$, and $r' \in N_{\epsilon_1}(S^*_1)$, we have that $\omega^p_{w^*}(r') \in B_{\epsilon'}(z^*_w)$.

Next, we have that $z^*_w \in \mathcal{R}^0_0(S'_2)$. By Lemma 2, for any $\epsilon_2 > 0$, there exists $\delta > 0$ and $p^*_2 > 0$ such that for any $p \leq p^*_2$ and $r \in B_\delta(z^*_w)$, we have $\omega^p_0(r) \subseteq B_{\epsilon_2}(S'_2) \subseteq N_{\epsilon_2}(S^*_2)$. Choose $\epsilon' < \delta$ and $p^* = \min\{p^*_1, p^*_2\}$, then for any $\epsilon_2 > 0$, there exists an $\epsilon_1 > 0$ and a p^* as found above such that for all $p \leq p^*$, for any $r \in \omega^p_{w^*}(r')$, where $r' \in N_{\epsilon_1}(S^*_1)$, we have $\omega^p_0(r) \subseteq N_{\epsilon_2}(S^*_2)$. ■

Theorems 1 and 2 prove robustness to small perturbation of reprogramming strategies that are such that, when the constant input w^* is applied to Σ_0, system Σ_{w^*} has an exponentially stable steady state in the region of attraction of S^*. However, more general reprogramming strategies, where the constant input could lead to, for instance, a stable limit cycle in the region of attraction of S^*, are not shown to be robust to small perturbations. Such strategies would still reprogram Σ_0 to S^* in the sense of Definitions 1 or 2, but their robustness to small perturbations does not follow from Theorems 1 and 2.

IV. Application Example

We consider the core network motif that controls pluripotency, the Oct4-Sox2-Nanog network [11]. Oct4 and Sox2 typically act as a heterodimer [23], and so this 3-protein system is modeled as a two-dimensional system, where Oct4 can be seen to represent the Oct4-Sox2 heterodimer [15], [24]. Oct4 and Nanog mutually activate each other, while also undergoing self-activation. This interaction network is shown in Fig. 1a. This core motif can result in three stable steady states: one with low levels of Oct4 and Nanog, which corresponds to the trophectoderm state; one with intermediate levels of Oct4 and Nanog, which corresponds to the pluripotent state; and one with high levels of Oct4 and Nanog, which corresponds to the primitive endoderm state.

We use the following Hill-function based ordinary differential equation (ODE) model to represent this core motif, as done in previous works [15], [24]:

$$\dot{x}_1 = f_1(x) = \frac{\eta_1 + a_1(x_1/k_1)^2 + b_1(x_2/k_2)^2 + c_1(x_1x_2/k_3^2)^2}{1 + (x_1/k_1)^2 + (x_2/k_2)^2 + (x_1x_2/k_3^2)^2} - \gamma_1 x_1,$$
$$\dot{x}_2 = f_2(x) = \frac{\eta_2 + a_2(x_1/k_1)^2 + b_2(x_2/k_2)^2 + c_2(x_1x_2/k_3^2)^2}{1 + (x_1/k_1)^2 + (x_2/k_2)^2 + (x_1x_2/k_3^2)^2} - \gamma_2 x_2. \quad (3)$$

Here, x_1 and x_2 are the concentrations of the two species, Nanog and Oct4, respectively. The state variable is $x = (x_1, x_2)$. Parameters η_1 and η_2 are the leaky expression rates of the two species, a_1 and a_2 are the coefficients of activation by Nanog, b_1 and b_2 are the coefficients of activation by Oct4, c_1 and c_2 are the coefficients of cooperative activation by Oct4 and Nanog, k_1, k_2, k_3 are the apparent dissociation constants, and γ_1 and γ_2 are the decay rate constants of the species.

This ODE model, for certain parameters, gives rise to three stable steady states. The nullclines and stable steady states S^*_1, S^*_2, and S^*_3 for an illustrative parameter set are shown in Fig. 1b. We apply inputs to each of the

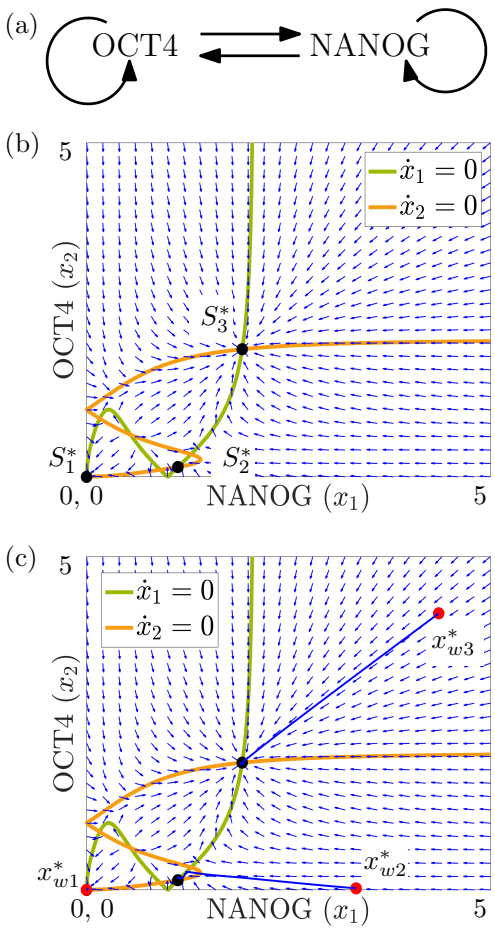

Fig. 1: The core network motif controlling pluripotency, and nullclines of the core motif, modeled as in (3), with blue arrows showing the vector field $(\dot{x}_1, \dot{x}_2)$, normalized to make magnitude 1 for clarity. The illustrative set of parameters used are: $\eta_1 = \eta_2 = 10^{-4}\text{nMs}^{-1}$, $k_1 = k_2 = k_3 = 1\text{nM}$, $a_1 = 2\text{nM}s^{-1}$, $a_2 = 1.8\text{nMs}^{-1}$, $b_1 = 0.25\text{nMs}^{-1}$, $b_2 = 2\text{nMs}^{-1}$, $c_1 = c_2 = 2.5\text{nMs}^{-1}$, $\gamma_1 = \gamma_2 = 1\text{s}^{-1}$. (a) The interaction graph of the core motif. (b) The black dots S^*_1, S^*_2, and S^*_3 represent stable steady states of the system. (c) The red dots x^*_{w1}, x^*_{w2}, and x^*_{w3} represent globally exponentially stable steady states of the system (4) for input values $w_1 = (0, 4\text{nMs}^{-1}, 0, 6\text{nMs}^{-1})$, $w_2 = (1.5\text{nM}, 0, 0, 6\text{nMs}^{-1})$, and $w_3 = (2\text{nM}, 0, 1.8\text{nM}, 0)$, respectively. The blue lines show the trajectories of system (3) once the input is removed. The system's state starting at x^*_{w1} converges to S^*_1, the system's state starting at x^*_{w2} converges to S^*_2, and the system's state starting at x^*_{w3} converges to S^*_3.

ODEs that either increase the production rate of the protein (overexpression) or that increase the protein turn-over rate (enhanced degradation). In the following discussion, we will refer to these types of inputs as positive and negative inputs, respectively. In particular, the modified system with all these inputs applied is:

$$\begin{aligned}\dot{x}_1 &= f_1(x) + u_1 - v_1 x_1,\\ \dot{x}_2 &= f_2(x) + u_2 - v_2 x_2,\end{aligned} \tag{4}$$

where the $+u_i$ term models protein overexpression, with $u_i > 0$, and the $-v_i x_i$ term models enhanced degradation, with $v_i > 0$. These inputs are then used to reprogram the system to a desired steady state S^* as follows. First, a constant input $w = (u_1, v_1, u_2, v_2)$ is applied to the system, as in (4). This input is chosen such that it results in a globally asymptotically stable steady state x^*_w in the region of attraction of S^*. Once the simulated system with this constant input converges to x^*_w, the input is removed, such that the unstimulated system (3) converges to the desired steady state S^*, and is therefore reprogrammed to it. Fig. 1c shows globally asymptotically stable steady states x^*_{w1}, x^*_{w2}, and x^*_{w3} of (4) for three different constant input vectors w_1, w_2 and w_3. These inputs have been selected such that x^*_{w1}, x^*_{w2}, and x^*_{w3} lie in the region of attraction of S^*_1, S^*_2 and S^*_3, respectively. Then, once the system (4) converges to x^*_{w1}, x^*_{w2}, or x^*_{w3} under inputs w_1, w_2 or w_3, these inputs are removed, and the unstimulated system (3) converges to the steady states S^*_1, S^*_2 or S^*_3, respectively.

These inputs w_1, w_2, and w_3 were selected based on strategies outlined in [16], that were developed for networks with the special property of being cooperative monotone, as is the case for the core motif shown in Fig. 1a. However, this motif is infact embedded in a larger network, and interacts with other species. An example of an extended network where this motif is embedded is shown in Figure 2a. Here, we have the additional species Cdx2, which is repressed by Oct4 and represses it in turn, while also undergoing self-activation [25], [26]. Further, the interaction between Oct4 and Nanog is not one of pure activation, since at high concentrations, Oct4 also weakly represses Nanog [27]. In this extended network, the core motif experiences a perturbation in its dynamics due to its interaction with Cdx2 and the repression of Nanog by Oct4. Further, the extended network is no longer cooperative. This perturbation potentially changes the location of the steady states of the system, their regions of attraction, and therefore the effect of the inputs on the system.

The dynamics of this extended GRN are then:

$$\begin{aligned}\dot{x}_1 &= \frac{\eta_1 + a_1(x_1/k_1)^2 + b_1(x_2/k_2)^2 + c_1(x_1x_2/k_3^2)^2}{1+(x_1/k_1)^2+(x_2/k_2)^2+(x_1x_2/k_3^2)^2+\boxed{d_1(x_2/k_2)^4}} - \gamma_1 x_1,\\ \dot{x}_2 &= \frac{\eta_2 + a_2(x_1/k_1)^2 + b_2(x_2/k_2)^2 + c_2(x_1x_2/k_3^2)^2}{1+(x_1/k_1)^2+(x_2/k_2)^2+(x_1x_2/k_3^2)^2+\boxed{d_2(y/k_4)^2}} - \gamma_2 x_2,\\ \dot{y} &= \frac{\eta_3 + a_3(y/k_4)^2}{1+(y/k_4)^2+d_3(x_2/k_2)^2} - \gamma_3 y,\end{aligned} \tag{5}$$

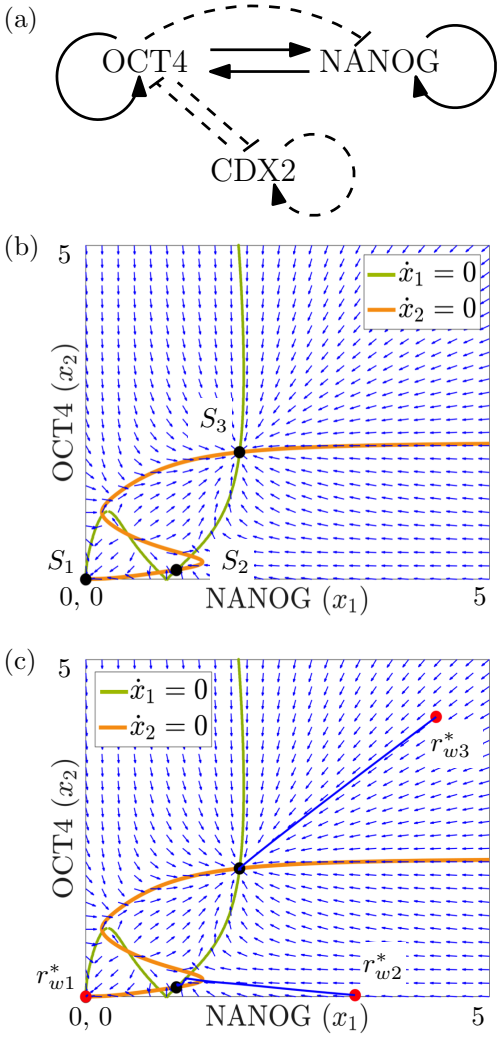

Fig. 2: The core pluripotency network motif embedded in an extended network [25], [26] and nullclines of the same, modeled as in (5), with blue arrows showing the normalized vector field $(\dot{x}_1, \dot{x}_2)$ evaluated at $y = \bar{y}(x)$. Additional parameters of the system are: $\eta_3 = 1\text{nMs}^{-1}$, $a_3 = 1.6\text{nMs}^{-1}$, $d_1 = d_2 = d_3 = 0.01$, $k_4 = 1\text{nM}$, $\gamma_3 = 1\text{nMs}^{-1}$. (a) The interaction graph of the extended GRN. (b) The black dots S_1, S_2, and S_3 represent stable steady states of the perturbed system. (c) The red dots r^*_{w1}, r^*_{w2}, and r^*_{w3} represent globally exponentially stable steady states of the system (6) for input values $w_1 = (0, 4\text{nMs}^{-1}, 0, 6\text{nMs}^{-1})$, $w_2 = (1.5\text{nM}, 0, 0, 6\text{nMs}^{-1})$, and $w_3 = (2\text{nM}, 0, 1.8\text{nM}, 0)$, respectively. The blue lines show the trajectories of system (5) once the input is removed. The system's state starting at r^*_{w1} converges to S_1, starting at r^*_{w2} converges to S_2, and starting at r^*_{w3} converges to S_3.

where y represents the concentration of Cdx2. Parameters $\eta_3, a_3, k_4, \gamma_3$ represent the leaky expression, self-activation and decay of y, and d_3 is the coefficient of repression of Cdx2 by Oct4. These parameters are chosen such that Assumption 2 is satisfied, and for any $x = x^*$, $\dot{y}$ has a globally exponentially stable steady state $\bar{y}(x)$. The terms in the red boxes represent the dynamic perturbations to the core motif shown in (3), where d_1 is the coefficient of repression of Nanog by Oct4, d_2 is the coefficient of repression of Oct4

by Cdx2, and the perturbation strength $p = ||d_1 + d_2||$. The nullclines of this system are shown in Fig. 2b, with $y = \bar{y}(x)$.

Inputs are applied to this system in the same way as in (4). The stimulated GRN is then:

$$\begin{aligned}
\dot{x}_1 &= \frac{\eta_1 + a_1(x_1/k_1)^2 + b_1(x_2/k_2)^2 + c_1(x_1x_2/k_3^2)^2}{1 + (x_1/k_1)^2 + (x_2/k_2)^2 + (x_1x_2/k_3^2)^2 + \boxed{d_1(x_2/k_2)^4}} - \gamma_1 x_1 \\
&\quad + u_1 - v_1 x_1, \\
\dot{x}_2 &= \frac{\eta_2 + a_2(x_1/k_1)^2 + b_2(x_2/k_2)^2 + c_2(x_1x_2/k_3^2)^2}{1 + (x_1/k_1)^2 + (x_2/k_2)^2 + (x_1x_2/k_3^2)^2 + \boxed{d_2(y/k_4)^2}} - \gamma_2 x_2 \\
&\quad + u_2 - v_2 x_2, \\
\dot{y} &= \frac{\eta_3 + a_3(y/k_4)^2}{1 + (y/k_4)^2 + d_3(x_2/k_2)^2} - \gamma_3 y,
\end{aligned} \tag{6}$$

The same inputs w_1, w_2 and w_3 that strongly reprogrammed (3) to S_1^*, S_2^* and S_3^* are applied to the GRN. From Fig. 2c, we see that this reprogramming strategy is robust to small perturbations, as these inputs result in ω-limit sets (in this case globally asymptotically stable steady states) $r_{w1}^*, r_{w2}^*, r_{w3}^*$ such that the corresponding ω-limit sets S_1, S_2, S_3 of the unstimulated systems are close to the original steady states S_1^*, S_2^*, S_3^*, respectively.

V. Conclusions

In this paper, we analyzed strategies developed to reprogram core network motifs, which control cell fate decision making, to a desired steady state S^*. Here, a constant input is applied to the network motif such that the motif's state, under this input, converges to an exponentially steady state in the region of attraction of S^*. Then, once this input is removed, the state of the unstimulated motif converges to S^*, and thus, the motif is reprogrammed to S^*.

When such motifs are embedded in larger networks, the rest of the network applies a perturbation to the dynamics of the motif. From Theorems 1 and 2, we find that for a sufficiently small perturbation, inputs that reprogram the core motif to S^* as described above, are effective in making the larger network remain arbitrarily close to S^*, and are therefore robust to small perturbations in the sense of Defintions 4, 5. These results hold under Assumptions 1-5. While Assumptions 1, 3-5 are physically realistic, the most stringent assumption is Assumption 2, which requires that for any fixed value of the state of the motif, the rest of the network has a globally exponentially stable steady state.

These results are demonstrated via the Oct-Nanog network motif that controls pluripotency. We find that the inputs that were found to reprogram the core motif to a desired steady state, were also able to reprogram the extended network with Cdx2 to a steady state close to the original desired steady state for a small enough perturbation. These results therefore show that the reprogramming strategies described here are practically feasible when the motif is embedded in a larger network. Additional work is required to consider the most complete currently known extended network for the pluripotency motif [20], [28] and to determine whether our results apply.

VI. Acknowledgments

The authors would like to thank Mr. Yili Qian and Mr. Ted Grunberg for valuable discussions and their help with resources and proofreading.

References

[1] J. Slack. *Essential developmental biology*. John Wiley & Sons, 2009.
[2] K. Takahashi and S. Yamanaka. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *cell*, 126(4):663–676, 2006.
[3] S. Ferber, A. Halkin, H. Cohen, I. Ber, Y. Einav, I. Goldberg, I. Barshack, R. Seijffers, J. Kopolovic, N. Kaiser, et al. Pancreatic and duodenal homeobox gene 1 induces expression of insulin genes in liver and ameliorates streptozotocin-induced hyperglycemia. *Nature medicine*, 6(5):568, 2000.
[4] T. Vierbuchen, A. Ostermeier, Z. Pang, Y. Kokubu, T. Südhof, and M. Wernig. Direct conversion of fibroblasts to functional neurons by defined factors. *Nature*, 463(7284):1035, 2010.
[5] X. Liu, F. Li, E. Stubblefield, B. Blanchard, T. Richards, G. Larson, Y. He, Q. Huang, A. Tan, D. Zhang, et al. Direct reprogramming of human fibroblasts into dopaminergic neuron-like cells. *Cell research*, 22(2):321, 2012.
[6] P. Huang, Z. He, S. Ji, H. Sun, D. Xiang, C. Liu, Y. Hu, X. Wang, and L. Hui. Induction of functional hepatocyte-like cells from mouse fibroblasts by defined factors. *nature*, 475(7356):386, 2011.
[7] T. Sapir, K. Shternhall, I. Meivar-Levy, T. Blumenfeld, H. Cohen, E. Skutelsky, S. Eventov-Friedman, I. Barshack, I. Goldberg, S. Pri-Chen, et al. Cell-replacement therapy for diabetes: Generating functional insulin-producing tissue from adult human liver cells. *Proceedings of the National Academy of Sciences*, 102(22):7964–7969, 2005.
[8] P. Hou, C. Chuang, C. Yeh, W. Chiang, H. Liu, T. Lin, and H. Kuo. Direct conversion of human fibroblasts into neural progenitors using transcription factors enriched in human esc-derived neural progenitors. *Stem cell reports*, 8(1):54–68, 2017.
[9] T. Schlaeger and Daheron. A comparison of non-integrating reprogramming methods. *Nat Biotech*, 33(1):58–63, 2015.
[10] T. Graf and T. Enver. Forcing cells to change lineages. *Nature*, 462(7273):587, 2009.
[11] L. Boyer, T. Lee, M. Cole, S. Johnstone, S. Levine, J. Zucker, M. Guenther, R. Kumar, H. Murray, R. Jenner, D. Gifford, D. Melton, R. Jaenisch, and R. Young. Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell*, 122(6):947–956, 2005.
[12] S. Huang. Reprogramming cell fates: Reconciling rarity with robustness. *BioEssays*, 31:546–560, 2009.
[13] S. Huang, Y. Guo, G. May, and T. Enver. Bifurcation dynamics in lineage-commitment in bipotent progenitor cells. *Developmental biology*, 305(2):695–713, 2007.
[14] C. Furusawa and K. Kaneko. A dynamical-systems view of stem cell biology. *Science*, 338, 2012.
[15] D. Del Vecchio, H. Abdallah, Y. Qian, and J. Collins. A blueprint for a synthetic genetic feedback controller to reprogram cell fate. *Cell systems*, 4(1):109–120, 2017.
[16] Rushina Shah and Domitilla Del Vecchio. Reprogramming cooperative monotone dynamical systems. *Submitted to IEEE Conference on Decision and Control*, 2018.
[17] A. Sootla, D. Oyarzún, D. Angeli, and G. Stan. Shaping pulses to control bistable systems: Analysis, computation and counterexamples. *Automatica*, 63:254–264, 2016.
[18] L. Wang, R. Su, Z. Huang, X. Wang, W. Wang, C. Grebogi, and Y. Lai. A geometrical approach to control and controllability of nonlinear dynamical networks. *Nature communications*, 7:11323, 2016.
[19] I. Crespo, T. Perumal, W. Jurkowski, and A. Del Sol. Detecting cellular reprogramming determinants by differential stability analysis of gene regulatory networks. *BMC systems biology*, 7(1):140, 2013.
[20] J. Kim, J. Chu, X. Shen, J. Wang, and S. Orkin. An extended transcriptional network for pluripotency of embryonic stem cells. *Cell*, 132(6):1049–1061, 2008.
[21] A. Isidori. *Nonlinear Control Systems*. Springer, 1995.
[22] H. Khalil. *Nonlinear Systems*. Prentice Hall, 2002.
[23] J. Chew, Y. Loh, W. Zhang, X. Chen, W. Tam, L. Yeap, P. Li, Y. Ang, B. Lim, P. Robson, and H. Ng. Reciprocal transcriptional regulation of pou5f1 and sox2 via the oct4/sox2 complex in embryonic stem cells. *Molecular and Cellular Biology*, 25(14):6031–6046, 2005.

[24] H. Abdallah, Y. Qian, and D. Del Vecchio. A dynamical model for the low efficiency of induced pluripotent stem cell reprogramming. In *Submitted to American Control Conference*, 2015.

[25] D. Strumpf, C. Mao, Y. Yamanaka, A. Ralston, K. Chawengsaksophak, F. Beck, and J. Rossant. Cdx2 is required for correct cell fate specification and differentiation of trophectoderm in the mouse blastocyst. *Development*, 132(9):2093–2102, 04 2005.

[26] V. Chickarmane and C. Peterson. A computational model for understanding stem cell, trophectoderm and endoderm lineage determination. *PLoS ONE*, 3(10), 2008.

[27] T. Kalmar, C. Lim, P. Hayward, S. Muñoz-Descalzo, J. Nichols, J. Garcia-Ojalvo, and A. Martinez Arias. Regulated fluctuations in nanog expression mediate cell fate decisions in embryonic stem cells. *PLoS Biol*, 7(7), 2009.

[28] X. Huang and J. Wang. The extended pluripotency protein interactome and its links to reprogramming. *Current opinion in genetics & development*, 28:16–24, 2014.