Stochastic Differential Contraction in Nonlinear System Analysis

Mardavij Roozbehani, Alexandre Megretski and Munther A Dahleh

Abstract*— By adopting on a notion of** *stochastic differential contraction, the paper presents new results on the** *incremental mean squared gain* **(IMSG) analysis of nonlinear systems with stochastic inputs. The** *relative power* **between two stochastic processes is defined as the asymptotic average (over time) of the second moment of the point-wise distance (in some metric) between the two processes. The IMSG of a system is then defined as the** *relative power* **between two output trajectories driven by two independent instances of i.i.d. inputs with unit relative power. While contracting metrics have been previously used for analysis of nonlinear systems, the formulation and analysis method in this paper lead to new conditions that yield a more accurate upper bound on the system gain. The idea is to introduce a notion of stochastic differential contraction which does not explicitly embed an exponential rate of contraction. This approach is more suitable for analysis of systems with stochastic inputs. In particular, and unlike previous approaches, the standard $\mathcal{H}_2$-norm analysis results for linear systems can be recovered as a special case in this setting.**

I. INTRODUCTION

Contraction theory, which can be seen as an extension of the classical Lyapunov theory to analysis of the behavior of system trajectories with respect to each other, instead of around nominal equilibria, has a long history in the nonlinear systems literature. According to [1] and [2], the earliest references on the subject appear to be in the math circles [3]–[7], with further and independent grounding and theoretical developments contributed by various authors including [8] and [9]. These methods have found applications in various engineering problems including analysis of limit cycles [10], system identification [11], [12], synchronization of nonlinear oscillators [13], process control [14] and general stability analysis [15]. Recently, such methods have been revisited and extended to analysis of system gains for nonlinear systems with stochastic inputs [16], [17].

Previous approaches to contraction analysis of nonlinear systems with stochastic inputs [16], [17], rely on an exponential rate of contraction, which leads to Lyapunov inequalities of the form:

$$u(t+1) - \mu u(t) \leq \bar{\phi}(w_1(t), w_2(t)), \quad \forall t \in \mathbb{Z}_+, \qquad (1)$$

where $\mu < 1$ is the contraction rate and $u(t) = \bar{\psi}(x_1(t), x_2(t))$ is a functional that needs to be bounded from above. Here, $x_1(\cdot)$ and $x_2(\cdot)$ are two trajectories of the system corresponding to two different inputs $w_1(\cdot)$ and $w_2(\cdot)$. Inequalities of the form (1) inevitably lead to upper bounds on $u(\cdot)$ that scale like

$$\frac{1}{1-\mu}\ , \qquad (2)$$

as μ approaches to 1, which is identical in nature to the scaling that arises from a worst-case analysis. This worst-case type scaling is particularly evident when the underlying system is linear, in which case the contraction rate μ is equal to the square of the modulus of the dominant pole.

In contrast to [16] and [17], we define a notion of stochastic differential contraction, i.e., a metric which contracts only in expectation and does not explicitly embed an exponential rate of contraction. Therefore, the associated Lyapunov inequalities do not impose an exponential rate of decay on the distance between trajectories. Instead, they require that the underlying Lyapunov function decreases (in expectation) by an amount that does not explicitly depend on the value of the function itself (unlike the exponential decay rate case). Intuitively, we propose conditions that amount to Lyapunov inequalities of the form:

$$v(t+1) - v(t) \leq \phi(w_1(t), w_2(t)) - \psi(x_1(t), x_2(t))$$

where $\psi(\cdot,\cdot)$ is a suitable function that we wish to bound from above along the system trajectories. This approach allows us to avoid an inverse scaling of the form (2), and obtain sharper bounds than those in the existing literature. Interestingly, with this approach we recover the standard $\mathcal{H}_2$ analysis results for linear systems [18] as a special case. To the best of our knowledge this does not apply to the existing results in the literature, indicating that our results are less conservative.

Finally, our results are presented in terms of implicit nonlinear equations (descriptor form) that define the underlying dynamics. This is an attractive feature because many engineering systems including networks with optimization in the loop, systems with algebraic constraints, and parametrically identified systems from input/output data [12], are naturally represented in such implicit forms.

A. Preliminaries

Notation 1: The sets of non-negative real numbers and integers are denoted by $\mathbb{R}_+$ and $\mathbb{Z}_+$ respectively. For a vector $v \in \mathbb{R}^l$, v_k denotes the k-th element of v, and $|v|$ denotes the standard 2-norm: $|v|^2 \stackrel{\text{def}}{=} \left(\sum_{i=1}^{l} |v_i|^2\right)$. We will use $|v|_P$ to denote the standard quadratic norm (semi-norm) with respect to a positive definite (semidefinie) symmetric matrix P: $|v|_P^2 = v^T P v$. The space of $\mathbb{R}^l$-valued

The authors are with the Laboratory for Information and Decision Systems (LIDS), Massachusetts Institute of Technology (MIT), Cambridge, MA, U.S.A. emails: {mardavij, ameg, dahleh}@mit.edu.

This work was supported by the National Science Foundation (Grant CPS-11358430), the MIT-Skoltech Initiative, and the MIT-Masdar Initiative.

functions $u:\mathbb{Z}_+\mapsto\mathbb{R}^l$ with finite power

$$\|u\|_P^2 \stackrel{\text{def}}{=} \limsup_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T}|u(t)|_P^2$$

is denoted by $\bar{\ell}_2^P$.

With a slight abuse of notation, we denote discrete-time stochastic processes using functions defined over the domain of integers. If $u:\mathbb{Z}_+\mapsto\mathbb{R}^l$ is a stochastic process, then it is said to have finite average power if:

$$\mathcal{E}\|u\|_P^2 \stackrel{\text{def}}{=} \limsup_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T}\mathbf{E}|u(t)|_P^2<\infty.$$

The space of all such sequences is denoted by $\mathcal{E}\bar{\ell}_2^P$. For i.i.d. sequences the above definition is equivalent to the underlying distribution having finite second moment.When the standard 2-norm is applied (i.e., $P=I_l$) we drop P from all the introduced notations. For a differentiable function $f:\mathbb{R}^n\mapsto\mathbb{R}^m$, we use $\dot{f}_x(\cdot)$ or simply $\dot{f}(\cdot)$ to denote the Jacobian: $\dot{f}(x)=J_x(f(x))=df(x)/dx$. Since throughout the paper time is discrete, this notation would not be confused with derivative with respect to time. The cone of positive semi-definite (PSD) matrices in $\mathbb{R}^{n\times n}$ is denoted by $\mathcal{S}_+^n$. Finally, $\mathbf{Tr}(X)$ denotes the trace of a square matrix X.

Definition 1: Incremental Mean Squared Gain (IMSG): Let $\mathcal{W}$ denote the class of admissible input signals $w(\cdot)$ for a discrete-time dynamical system Σ. Let $\mathcal{Y}$ denote the signal space for the output $y(\cdot)$. Given a pair of positive semidefinite matrices P and Q of appropriate dimensions, the *(P,Q)-scaled incremental mean squared gain* of Σ is denoted by

$$\mathcal{G}_{P,Q}(w\to y),$$

and is defined to be the minimal $\gamma\geq 0$ such that the inequality

$$\inf_{\gamma}:\quad \gamma^2\mathcal{E}\|w-\tilde{w}\|_P^2-\mathcal{E}\|y-\tilde{y}\|_Q^2\geq 0 \tag{3}$$

is satisfied for all input/output pairs $(w,y)\in\mathcal{W}\times\mathcal{Y}$ and $(\tilde{w},\tilde{y})\in\mathcal{W}\times\mathcal{Y}$ such that

$$w-\tilde{w}\in\mathcal{E}\bar{\ell}_2^P.$$

II. MAIN RESULTS

A. Stochastic Differential Contraction Analysis of Implicit Dynamical Systems

Consider discrete-time dynamical system Σ specified by the implicit state-space model:

$$\Sigma\ :\quad \psi(x(t+1),x(t),w(t))=0 \tag{4}$$

$$y(t)\ =\ h(x(t)) \tag{5}$$

where $\psi:\mathbb{R}^n\times\mathbb{R}^n\times\mathbb{R}^m\mapsto\mathbf{0}_n$, and $h:\mathbb{R}^n\mapsto\mathbb{R}^p$ are continuously differentiable functions.

Assumption 1: The function ψ is invertible in the sense that the equation $\psi(a,x,w)=0$ has a unique solution $a\in\mathbb{R}^n$ for all $x\in\mathbb{R}^n$, and $w\in\mathbb{R}^m$.

Assumption 2: The class of admissible input signals $\mathcal{W}$ is restricted to i.i.d. sequences $w:\mathbb{Z}_+\mapsto\mathbb{R}^n$ with bounded covariance matrix $W\stackrel{\text{def}}{=}\mathbf{E}[ww^T]\in\mathbb{R}^{n\times n}$. Furthermore, w is independent of the initial condition $x(0)$ of Σ.

The following theorem presents our main result.

Theorem 1: Consider the system Σ defined in (4)–(5). Suppose that Assumptions 1 and 2 are satisfied. Let $g:\mathbb{R}^n\mapsto\mathbb{R}^n$ and $f:\mathbb{R}^n\mapsto\mathbb{R}^n$ be a pair of continuously differentiable functions such that the equation

$$g(x(t+1))=f(x(t))+w(t) \tag{6}$$

$$y(t)=h(x(t)) \tag{7}$$

is satisfied for every pair of sequences $x:\mathbb{Z}_+\mapsto\mathbb{R}^n$ and $w:\mathbb{Z}_+\mapsto\mathbb{R}^n$ that satisfy (4). Suppose that there exist positive semidefinite matrices $P\in\mathcal{S}_+^n$ and $Q\in\mathcal{S}_+^n$ such that the following inequality holds:

$$|\dot{g}(x)\Delta|_P^2\geq\left|\dot{f}(x)\Delta\right|_P^2+\left|\dot{h}(x)\Delta\right|_Q^2,\quad\forall x,\Delta \tag{8}$$

Then, the (P,Q)-scaled IMSG of Σ is upper bounded by 1:

$$\mathcal{G}_{P,Q}(w\to y)\leq 1. \tag{9}$$

Proof: Let (x_1,y_1) and (x_2,y_2) be two solutions of Σ corresponding to two different inputs $w_1(\cdot)$ and $w_2(\cdot)$, and initial conditions $x_1(0)$ and $x_2(0)$. Let (x_γ,y_γ) be the solution corresponding to the input

$$w_\gamma(t)=\gamma w_1(t)+(1-\gamma)w_2(t),$$

and the initial condition:

$$x_\gamma(0)=\gamma x_1(0)+(1-\gamma)x_2(0),$$

Then $\Delta_\gamma(t)=\partial x_\gamma(t)/\partial\gamma$ and $\Gamma_\gamma(t)=\partial y_\gamma(t)/\partial\gamma$ are well-defined continuous functions of γ. Thus, considering (6) and (7) with $x=x_\gamma$ and $y=y_\gamma$, and differentiating with respect to γ we obtain the following linear dynamics for $\Delta_\gamma(t)$:

$$\dot{g}(x_\gamma(t+1))\Delta_\gamma(t+1)=\dot{f}(x_\gamma(t))\Delta_\gamma(t)+w_1(t)-w_2(t) \tag{10}$$

$$\Gamma_\gamma(t)=\dot{h}(x_\gamma(t))\Delta_\gamma(t) \tag{11}$$

Define the functional $S:\mathbb{R}^{2n}\mapsto\mathbb{R}_+$ according to:

$$S:(x,\Delta)\mapsto|\dot{g}(x)\Delta|_P^2$$

For every fixed $\gamma\in[0,1]$, let $V_\gamma:\mathbb{Z}_+\mapsto\mathbb{R}_+$ be the function that is defined as follows:

$$V_\gamma(t)=S(x_\gamma(t),\Delta_\gamma(t)).$$

It then follows from (10) that:

$$\mathbf{E}V_\gamma(t+1)=\mathbf{E}|\dot{g}(x_\gamma(t+1))\Delta_\gamma(t+1)|_P^2 \tag{12}$$

$$=\mathbf{E}\left|\dot{f}(x_\gamma(t))\Delta_\gamma(t)\right|_P^2+\mathbf{E}|w_1(t)-w_2(t)|_P^2 \tag{13}$$

where, the equality between (12) and (13) follows from independence of noise from the state, which in turn follows

from Assumption 2. It then follows from equations (8) and (11) – (13) that:

$$\mathbf{E}V_\gamma(t+1) - \mathbf{E}V_\gamma(t) \leq \mathbf{E}\left|w_1(t) - w_2(t)\right|_P^2 - \mathbf{E}\left|\dot{h}(x_\gamma)\Delta_\gamma(t)\right|_Q^2 \quad (14)$$

Integrating with respect to γ and summing up both sides of the above inequality yields:

$$\frac{1}{T}\sum_{t=0}^{T-1}\int_0^1 \mathbf{E}\left|\dot{h}(x_\gamma)\Delta_\gamma(t)\right|_Q^2 d\gamma \leq \mathbf{E}\left|w_1(\hat{t}) - w_2(\hat{t})\right|_P^2 + \frac{1}{T}\int_0^1 \mathbf{E}V_\gamma(0)d\gamma \quad (15)$$

where $\hat{t} \in \{0, \ldots, T-1\}$ is an arbitrary time. Since

$$|y_1(t) - y_2(t)|_Q^2 = \left|\int_0^1 \dot{h}(x_\gamma)\Delta_\gamma(t)d\gamma\right|_Q^2 \leq \int_0^1 \left|\dot{h}(x_\gamma)\Delta_\gamma(t)\right|_Q^2 d\gamma, \quad (16)$$

it follows from (15) and (16) that

$$\limsup_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T-1}\mathbf{E}\left|y_1(t) - y_2(t)\right|_Q^2 \leq \mathbf{E}\left|w_1(\hat{t}) - w_2(\hat{t})\right|_P^2 = 2\mathbf{Tr}(PW) \quad (17)$$

This completes the proof. ∎

Remark 1: Note that condition (8) is equivalent to the following matrix inequality constraint:

$$\dot{g}(x)^T P \dot{g}(x) \succeq \dot{f}(x)^T P \dot{f}(x) + \dot{h}(x)^T Q \dot{h}(x), \quad \forall x.$$

In the following corollary, the standard $\mathcal{H}_2$ analysis results for linear systems (see, e.g., [18]) are recovered.

Corollary 1: Consider the linear system

$$x(t+1) = Ax(t) + w(t)$$
$$y(t) = Cx(t)$$

where $w(\cdot)$ is a zero-mean i.i.d. process. If there exists a PSD matrix P such that

$$P \succeq A^T P A + C^T C, \quad (18)$$

then,

$$\limsup_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T-1}\mathbf{E}\left|y(t)\right|^2 \leq \mathbf{Tr}(PW) \quad (19)$$

Proof: Apply Theorem 1 with $g(x) = x$, $f(x) = Ax$, and $h(x) = Cx$. Then (8) reduces to (18), and (19) follows from (17) and the fact that $(w(\cdot), x(\cdot), y(\cdot)) = (0, 0, 0)$ is an admissible solution. ∎

III. DISCUSSION AND NUMERICAL EXAMPLES

A. Linear Systems

In this section we briefly discuss our results and compare with the existing literature. For the moment let us consider the simple case where (6) is of the form:

$$x(t+1) = ax(t) + w(t)$$

where all variable are scalars, $a \in (-1, 1)$, and $W = 1$. With $Q = 1$, (8) reduces to:

$$P - a^2 P \geq 1$$

and we have

$$\mathcal{E}\left\|x - \tilde{x}\right\|^2 \leq \inf_P 2\mathbf{Tr}(PW) = 2(1-a^2)^{-1}.$$

This is in agreement with the bound that can be obtained from the results of [17] with $\mu = a^2,\ D = 1,\ \beta = 1$ and $M = 1$.

Next, let us consider a linear system with two states, defined via the following state space model:

$$A = \begin{bmatrix} a & 0 \\ 1 & 0.9 \end{bmatrix}, \quad B = I_2, \quad C = I_2 \quad D = 0_2,$$

and $W = I_2$. From Corollary 1, our results can be written as the following optimization problem:

$$\min_P \quad \mathbf{Tr}(P) \quad (20)$$
$$\text{subject to:} \quad P \succeq A^T P A + I_2$$

Again, we emphasize that these are the LMI conditions that arise in the standard $\mathcal{H}_2$ analysis of linear systems. The results of [17] can be summarized as the following optimization problem:

$$\min_{P,\mu} \quad \frac{\mathbf{Tr}(P)}{1-\mu} \quad (21)$$
$$\text{subject to:} \quad \mu P \geq A^T P A$$
$$P \geq I_2$$

Let γ_1 and γ_2 denote the optimal values of optimization problems (20) and (21) respectively. Figure 1 shows the ratio γ_2/γ_1 as the parameter a in the A matrix approaches to 1. These results show the advantage of the method presented in the paper and suggest that the contraction analysis approach of [17] can be more conservative than Theorem 1.

Remark 2: It has been shown in [16] that the corresponding bounds for stochastic contraction analysis of nonlinear systems are optimal in the sense that they can be achieved. The methods in [17] are essentially the extensions of the results of [16] to state dependent metrics. Note, however, that our results do not contradict the optimality claims of [16] and [17], as those bounds are achievable by a scalar dynamical system [16]. For scalar systems, our bound is equal to that of [17]. The proof of the optimality results as presented in [16] does not extend to higher dimensions.

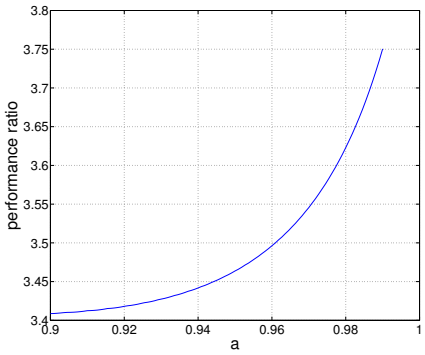

Fig. 1. The ratio γ_2/γ_1 as parameter a varies from 0.9 to 0.99.

B. A Simple Nonlinear Example

Consider discrete-time system $\tilde{\Sigma}$ defined by:

$$x_1(t+1) = \frac{a-1}{a}x_1(t) - \frac{5}{a}x_2(t) + \frac{1}{c}e^{-x_3(t)} + w(t) \quad (22)$$

$$x_2(t+1) = \frac{1}{a}x_1(t) + x_2(t) + \frac{1}{c}e^{-x_3(t)} + w(t) \quad (23)$$

$$x_3(t+1) = b(x_1(t)^2 + x_2(t)^2)^{-1} \quad (24)$$

$$y(t) = \left(x_1(t), x_2(t), e^{-x_3(t)}\right)^T \quad (25)$$

where a, b and c are positive constants. In this example we set $a = b = c = 100$. Also, $\mathbf{E}[w(t)^2] = 1$. Rewriting (24) as

$$e^{-x_3(t+1)} = e^{-b(x_1(t)^2+x_2(t)^2)^{-1}}, \quad (26)$$

we obtain a new representation of $\tilde{\Sigma}$ in the implicit form defined by (22), (23), and (26). It can be shown that Equation (8) cannot be satisfied if Theorem 1 is directly applied to (22)–(24). Reformulating Equation (24) in the exponential form (26) serves as a suitable nonlinear coordinate transformation that facilitates the analysis. In order to apply Theorem 1 to (22), (23), and (26), we compute the Jacobian matrices:

$$\dot{g}(x) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \phi_3 \end{bmatrix}, \quad \dot{f}(x) = \begin{bmatrix} \frac{a-1}{a} & -\frac{5}{a} & \phi_3 \\ \frac{1}{a} & 1 & \phi_3 \\ \phi_1 & \phi_2 & 0 \end{bmatrix}$$

where

$$\phi_3 = \phi_3(x) = -\frac{1}{c}e^{-x_3}$$

and

$$\phi_1 = \phi_1(x) = x_1\bar{\phi}(x_1, x_2), \quad \phi_2 = \phi_2(x) = x_2\bar{\phi}(x_1, x_2)$$

where

$$\bar{\phi}(x_1, x_2) = 2b\frac{e^{-b(x_1^2+x_2^2)^{-1}}}{(x_1^2 + x_2^2)^2}.$$

Note also that $\dot{h}(x) = \dot{g}(x)$. It can be shown that with $Q = I_3$, and a block diagonal matrix $P = \text{diag}(\bar{P}, p)$, where $\bar{P} \in \mathcal{S}_+^2$, and $p > 0$, a sufficient condition for (8) is given by the following matrix inequality constraint:

$$\begin{bmatrix} A^T\bar{P}A - \bar{P} + p\mu I_2 & A^T\bar{P}L \\ L^T\bar{P}A & L^T\bar{P}L - p \end{bmatrix} + I_3 \preceq 0. \quad (27)$$

where

$$\mu = \max_{x_1,x_2}\{\phi_1(x)^2 + \phi_2(x)^2\} = \frac{27e^{-3}}{2b}, \quad L = \frac{1}{c}[1 \quad 1]^T.$$

Feasibility of (27) implies that $\mathcal{G}_{P,I}(w \to y) \leq 1$ for any feasible solution P. By minimizing $\mathbf{Tr}(PW)$ subject to (27) we obtain:

$$\mathcal{E}\,\|y - \tilde{y}\|^2 \leq \inf_P \; 2\mathbf{Tr}(PW) < 1108.$$

Thus, the IMSG from w to y is less than $\sqrt{1108}/2 \approx 16.65$, that is, $\mathcal{G}_{I,I}(w \to y) < 16.65$.

IV. CONCLUSIONS

We presented new results on contraction analysis of nonlinear systems with stochastic inputs based on a notion of stochastic differential contraction which does not explicitly require an exponential rate of decay on the distance between two trajectories. As a special case, we recover the standard $\mathcal{H}_2$ analysis results for linear systems. This feature is absent in previous approaches to contraction analysis of nonlinear systems with stochastic inputs. This suggests that the presented method can in general be less conservative and provide an improvement over existing methods.

REFERENCES

[1] J. Jouffroy, "Some ancestors of contraction analysis," in *Decision and Control, 2005 and 2005 European Control Conference. CDC-ECC '05. 44th IEEE Conference on*, pp. 5450–5455, 2005.

[2] A. Pavlov, *et al.*, "Convergent dynamics, a tribute to Boris P. Demidovich," *Systems & Control Letters*, vol. 52, no. 3, pp. 257–261, 2004.

[3] D. Lewis, "Metric properties of differential equations," *American Journal of Mathematics*, vol. 71, no. 2, pp. 294–312, 1949.

[4] Z. Opial, "Sur la stabilité asymptotique des solutions dun systeme déquations différentielles," *Ann. Polinici Math*, vol. 7, pp. 259–267, 1960.

[5] P. Hartman, "On stability in the large for systems of ordinary differential equations," *Canadian Journal of Mathematics*, vol. 13, pp. 480–492, 1961.

[6] B. P. Demidovich, "Dissipativity of a nonlinear system of differential equations," *Vestnik Moscow State University, Ser. Matematika Mekhanika,* vol. Part I, no. 6, pp. 19–27 (1961), vol. Part II, no. 1, pp. 3–8, 1962.

[7] P. Hartman, *Ordinary differential equations*. Birkhauser, 1964.

[8] W. Lohmiller and J.-J. E. Slotine, "On contraction analysis for nonlinear systems," *Automatica*, vol. 34, no. 6, pp. 683–696, 1998.

[9] D. Angeli, "A Lyapunov approach to incremental stability properties," *Automatic Control, IEEE Trans. on*, vol. 47, no. 3, pp. 410–421, 2002.

[10] I. R. Manchester and J.-J. E. Slotine, "Transverse contraction criteria for existence, stability, and robustness of a limit cycle," *Systems & Control Letters*, vol. 63, pp. 32–38, 2014.

[11] M. M. Tobenkin, I. R. Manchester, and A. Megretski, "Stable nonlinear identification from noisy repeated experiments via convex optimization," *arXiv preprint arXiv:1303.4175*, 2013.

[12] M. M. Tobenkin, I. R. Manchester, J. Wang, A. Megretski, and R. Tedrake, "Convex optimization in identification of stable non-linear state space models," in *Decision and Control (CDC), 2010 49th IEEE Conference on*, pp. 7232–7237, IEEE, 2010.

[13] W. Wang and J.-J. E. Slotine, "On partial contraction analysis for coupled nonlinear oscillators," *Biological cybernetics*, vol. 92, no. 1, pp. 38–53, 2005.

[14] W. Lohmiller and J.-J. E. Slotine, "Nonlinear process control using contraction theory," *AIChE journal*, vol. 46, no. 3, pp. 588–596, 2000.

[15] K. H. Khalil, *Nonlinear Systems*. Prentice Hall, 2002.

[16] Q.-C. Pham, N. Tabareau, and J.-J. Slotine, "A contraction theory approach to stochastic incremental stability," *Automatic Control, IEEE Trans. on*, vol. 54, no. 4, pp. 816–820, 2009.

[17] Q.-C. Pham and J.-J. Slotine, "Stochastic contraction in riemannian metrics," *arXiv preprint arXiv:1304.0340*, 2013.

[18] G. E. Dullerud and F. Paganini, *A course in robust control theory*, vol. 6. Springer New York, 2000.