Towards Controller Invalidation from Open- and Closed- Loop Experiments

Max Simchowitz
msimchow@berkeley.edu
Joint work with Qingqing Huang and Ross Boczar

December 15, 2016

1 Introduction

Broadly speaking, the aim of Robust Control is to design controllers which attain certain performance guarantees in the presence of systems disturbances, and plant misspecifications. For example, a cannonical problem is to design a linear controller K so that for any misspecification Δ in a constraint-set $\mathbf{\Delta}$, the corresponding inter-connection loop is stable.

In order for robust synthesis to be possible, we require our knowledge about the uncertainty - e.g., the set $\mathbf{\Delta}$ - to be pretty well specified. Indeed, for large enough families of uncertainty blocks, we may be unable to find practical controllers which stabilize all possible plants with those uncertainties. Hence, a natural net step might be to use *experimental data* about the plant to learn additional constraints on the uncertainty blocks, in order to reach a tractable robust control problem.

On one extreme, one could try to essentially learn the entire plant behavior, up to a vanishingly small uncertainty. This problem is known as System Identification, or sys-id (see, for example [1] and [5]). Unfortunately, sys-id often performs poorly in practice, because a) dynamical are complex objects that require lots of data to pin down and b) restrictive a-priori models of plant behavior (e.g., linearity) make proper learner learning difficult. On the other extreme, the "falsifiable control" literature of [2] at designing a controller by testing a sequences of candidate controllers $K_1, \dots, K_M$ in closed loop, and swapping out one controller to the next when the learner can certify that the current controller violates a certain performance specification. While this method does not rely on a-priori assumptions about plant structure, it seems too pessimistic in not taking advantage of plant modelling. Moreover, the methods do not offer any principled guidance on how to select from a a large, or possibly (uncountably) infite set of candidate controllers.

In this report, we will describe forthcoming work which tries to strike a middle ground between these two strategies. We will concern ourself with the relatively modest task of *controller invalidation*, namely, using empirical open- and closed-loop data, how can we verify that a given controller K fails to meet a certain performance objective? Ultimately, we hope that the invalidation problem will guide a principled and efficient approach to *controller synthesis*, which is concerned with *finding* a controller to achieve the desired performance guaranteed. However, since the sample complexity of the invalidation task is

still not well known, we will restrict our inquiry to this more modest problem. In particular, we would like to know the following:

1. How much open- and closed-loop data (and under what assumptions on those data) is required to invalidate a controller K?

2. What is the gap between computationally efficient invalidation procedures - such as those which rely on the S-procedure - and procedures which have access to unlimited computational power.

We emphasize that this research is a work in progress. Thus, we will focusing on clarifying our initial steps, including the problem formulation, posing the problem as a non-convex quadratic program, relaxing the program via the S-procedure, and discussing challenges in proving sample complexity bounds going forward. Before continuing, we present a brief review of some statistical and data-driven approaches to learning dynamical systems, and designing controllers.

2 Relevant Work

Our approach is motivated by the model-invalidation work of[6], which uses experimental data about unknown plant to invalidate whether or not a set of candidate uncertainty $\Delta \in \mathbf{\Delta}$ is broad enough to account for observed plant behavior. Unlike the settings we will consider, their setting assumes that the uncertains may be nonlinear causal operators (though the nomimal plant is take to be linear). The system under consideration is thus described by the following LFT:

$$\begin{bmatrix} y \\ z \end{bmatrix} = \begin{bmatrix} P_{11} & P_{12} & P_{13} \\ P_{21} & P_{22} & P_{23} \end{bmatrix} \begin{bmatrix} v \\ w \\ u \end{bmatrix} \quad \text{and} \quad v = \Delta z \tag{2.1}$$

Here, the variables represent length T-signals, so for example $z \in \mathbb{R}^{n_z T}$. We assume that the signals y and u are known (measured), that w is an unknown exogenous distribution which lies in a constraint set $\mathcal{H}_w$, and that (z, v) are unknown un-measured internal signals which lie in the constraint set $\mathcal{H}_{\mathbf{\Delta}} := \{z, v : v = \Delta z\}$. Given the observations y and u, the goal is to understand

> Do there exists disturbances $w \in \mathcal{H}_w$ and signals $(z, v) \in \mathcal{H}_\Delta$ which are consistent with the observations (y, u), in the sense that Equation 2.1 holds?

In general, this problem is computationally infeasible for abitrary sets $\mathbf{\Delta}$. One step towards tractability is to express the uncertainty in terms of the intersections quadratic constraints of the form $\mathcal{H}_w = \{w : W_i(w) \geq 0\}_{1 \leq i \leq d}$ and $\mathcal{H}_{\mathbf{\Delta}} = \{(z, v) : Q_j(z, v) \geq 0\}_{1 \leq j \leq r}$, where W_i and Q_j are *quadratic functionals*, that is, functions of the form $F(x) = x^* A x + b^* x + c$

One particular type of quadratic functional which is known to have great expressive power are IQC's, which given a signal $x(1), \ldots, X(T)$ take the form $\sum_{t=1}^{T} F_t(x) \geq 0$ for quadratic functions $F_t(x)$ (thus the "F" matrix is block-diagonal). The seminal work on using IQC's to analyze feedback systems [3] establishes that many useful notions of constraints on both signals w and pairs (z, v) that arise from plants $\Delta \in \mathbf{\Delta}$ can be formalized in terms of IQC's, including bounded signal norms, sector-inequalities for non-linearities, and $l_2 \to l_2$ plant gains (or equivalently, H_∞).

2.1 Computational Infeasibility

While describing the system in terms of quadratic contraints in general, and IQCs in particular, yields a clean problem formulation, verifying that the constraints have a non-empty intersection is, in-general, computationally intractable. The reason is that most interesting IQC's have indefinite "A"-matrices, and thus their corresponding constraint sets are non-convex (indeed, if a quadratic form F has not b and c terms, then $F(x) \geq 0$ is trivially satisfied if $A \succeq 0$).

Here, we present two IQC's of interest which express the sorts of constraints which will ultimately become important for our task. We will remark that both such constraints correspond to indefinite quadratic forms, thereby highlighting the computational difficulty of our task. The first is constraining the $l_2 \to l_2$ gain Δ by some $\gamma > 0$, that is, the constraint $\|v\|_2 \leq \gamma \|z\|_2$. This can be expressed as the quadratic constraint

$$Q_i(z, v) = \begin{bmatrix} z \\ v \end{bmatrix}^* \begin{bmatrix} \gamma^2 I & 0 \\ 0 & -I \end{bmatrix} \begin{bmatrix} z \\ v \end{bmatrix} \geq 0 \tag{2.2}$$

where we see immediately that the "A" term $\begin{bmatrix} \gamma^2 I & 0 \\ 0 & -I \end{bmatrix}$ is indefinite.

A second, more involved constraint is that the noise w is essentially quite noise. One way of expressing this is that the correlation between $w(t)$ and $w(t+k)$ decreases as the delay k increases. Assuming that $w(t)$ are scalars, we can express this condition as follows. Define the k-delayed autocorrelation

$$r_w(t) = \frac{1}{T} \sum_{t=1}^{T} w(t + k \mod T) w(t) \tag{2.3}$$

Note that if $w(t)$ are i.i.d and zero-mean, then $r(0) \geq 0$ is an unbiased estimator of $\mathbb{V}\mathfrak{d}\diagdown[w(t)]$. Hence, one definition of "white" is to say that that, for all $k \neq 0$, $|r_w(k)|$ is considerability smaller than its variance. That is,

$$|r_w(k)| \leq \gamma r_w(0) \iff \tag{2.4}$$

Letting Z denote the cyclic shift operator on length-T sequences, Equation 2.4 can be expressed as an IQC by the formula

$$\gamma r_w(0) + \sigma r_w(k) = w^*(\gamma I + \frac{\eta}{2}((Z^k)^* + Z^k))w \geq 0 \tag{2.5}$$

where k ranges from $1, \ldots, T-1$, and $\eta \in \{-1, 1\}$, yielding $2(T-1)$ quadratic constraints. Since $(Z^k)^* + Z^k)$ is indefinitely, it holds that the constraints in Equation 2.5 are indefinite.

2.2 A conservative solution: The S-procedure

By lifting to the space of tuples $x = (w, v, z)$, our problem amounts to checking that the intersection of a finite collection of quadratic functional $\{G_i(x) \geq 0\}_{0 \leq i \leq r}$ is nonempty.

A conservative approach to verifying the emptiness of the intersection, and thereby invalidating the model, is known as the S-procedure

Lemma 2.1. *Suppose there exists scalars $\tau_i \geq 0$ such that*

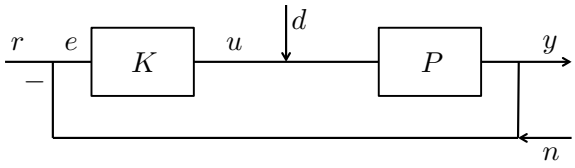
$$G_0 + \sum_{i=1}^{r} \tau_i G_i \geq 0 \tag{2.6}$$

Then, $\bigcap_{i=1}^{r} \{x : G_i(x) < 0\} \cap \{x : G_0(x) > 0\} = \emptyset$.

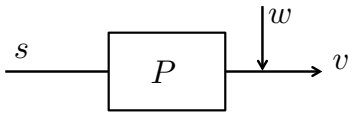
Note that the condition of Equation 2.6 amounts to checking if a linear form in τ, subject to the nonnegativity constraints $\tau_i \geq 0$ in non-negative. By clever re-parametrization, this can then by represented as semi-definite program, which can be provably solved in polynomial time.

3 Adopting the IQC framework to Controller Invalidation

In this section, we shall introduce our IQC framework for controller invalidation. In what follows, we have an unkown plant P, and our goal is to decide whether a closed loop system obtained by feeding back K into P satisfies certain performance guarantees. We will also work in the SISO setting, for the sake of simplicity, and assume all rollouts have length T. Finally, we will assume that P is given as a finite impulse response p. As a block diagram, our system is represented as follows. Our goal will be two invalidate K using N rollouts of



open-loop data, and M rollouts of closed loop data. For the open loop, we feed in an input s into the plant P, and observe $v_i = Ps_i + w_i$, where w_i is some noise. As a block diagram, Assuming that w_i satisfies lies in some constraint set $\mathcal{H}_w$, this induces the constraint that



$$v_i - p_i * s_i \in \mathcal{H}_w \tag{3.7}$$

where $*$ denotes the convolution operator. In other-words, if T_{s_i} is the Toeplitz matrix associated with s_i, then,

$$v_i - T_{s_i} p_i \in \mathcal{H}_w \tag{3.8}$$

Since v_i and s_i are known, these amount to controls of the finite-impulse response of the plant P.

Next, we perform M closed loop experiments. From the block diagram, we find that

$$y_j = p * (K(r_j - (y_j + n_j)) + d_j) \tag{3.9}$$

We will also assume that the disturbances j and d satisfy some regularity conditions, namely

$$d_j \in \mathcal{H}_d \quad \text{and} \quad n_j \in \mathcal{H}_n \tag{3.10}$$

Then, the invalidation step amount to asking whether or not the signal pairs (d_j, y_j) and (n_j, y_j) satisfy some performance guarantes. That is,

$$(d_j, y_i) \in \mathcal{H}_{(d,y)} \quad \text{and} \quad (n_j, y_j) \in \mathcal{H}_{(n,y)} \tag{3.11}$$

All in all, the following feasibility problem for all $i \in [n]$,

$$\begin{aligned}
&(d_j, y_i) \in \mathcal{H}_{(d,y)} \quad \text{and} \quad (n_j, y_j) \in \mathcal{H}_{(n,y)} \quad \text{performance constraints} \\
&y_j = p * (K(r_j - (y_j + n_j)) + d_j), \quad \text{closed-loop constraint} \\
&d_j \in \mathcal{H}_d \quad \text{and} \quad n_j \in \mathcal{H}_n \quad \text{disturbance and noise constraints} \\
&v_i - T_{s_i} p \in \mathcal{H}_w \quad \text{open-loop constraints}
\end{aligned}$$

In our case, we will choose the sets $\mathcal{H}_{(\cdot)}$ to be specified by level-sets of indefinite quadratics, as in the previous section. Moreover, the closed loop constraint is in fact a bi-linear (and thus quadratic) constraint in the unkowns p, and in n_j and d_j. Hence, we can write this constraint as the intersection of two quadratic inequality constraints, as shown below. This will leave us with a system of quadratic inequalities, which we can conservatively invalidate using the S-procedure.

3.1 Representing the Closed Loop Constraint

For $t = 1, \ldots, T$, let C_t denote the t-step convolution matrix, so that the convolution of signals $x_1 * x_2$ is given by $(x_1 * x_2)(t) = x_1^* C_t x_2$ (since $(x_1 * x_2)(t)$ is a bi-linear form in x_1 and x_2, such a C_t is guaranteed to exist). Hence, if we define the variables $c_{t,j} = C_t K(r_j - y_j)$ (note that r and y are known), we can write the losed loop constraints as

$$\begin{aligned}
y_j(t) &= p^T C_t (K(r_j - (y_j + n_j)) + d_j) \\
&= p^T C_t K(r_j - y_j) + p^T C_t (d_j - K n_j) \\
&= p^T c_{t,j} + \begin{bmatrix} p \\ d_j \\ n_j \end{bmatrix}^T \begin{bmatrix} 0 & \frac{1}{2} C_t & \frac{-1}{2} C_t K \\ \frac{1}{2} C_t & 0 & 0 \\ \frac{-1}{2} C_t K & 0 & 0 \end{bmatrix} \begin{bmatrix} p \\ d_j \\ n_j \end{bmatrix} \qquad := C_{t,j}(p, d_j, n_j)
\end{aligned}$$

where we define the function $C_{t,j}(p, d_j, n_j)$ to be the quadratic form defined above.

3.2 Constraints on Noise and Disturbance

In general, it might be impossible to invalidate an ineffective controller from infinitely many observations. Indeed, the observation noise w_j from the open loop ID is adversarially chosen, then it can greatly obfuscate our knowledge of P. Further still, if d_j and n_j are adversarially chosen, then they may chosen benignly so that the observed closed loop data has the desired performance guarantees, will on "typical" sequences of noise and disturbances, the plant's performance guarantees might falter.

Hence, for the aim of our statistical learning problem, we need to formulate some notion of what a "typical" disturbance/noise is. Then, we need to formulate a notion of a bad (i.e., falsifiable) controller based on typical sequences (indeed, a controller may actually attain the desired performance guarantees on some rare sequences but fail on average). Ultimately, our goal is to relate how "bad" a controller it is, to how frequently, and by how much, it fails performance guarantees on typical sequences of noises and disturbances.

To this end, we shall assume that all of w_i, n_j and d_j are white Gaussian noise. Moreover, we will take our performance specification to be simple disturbance/noise rejection, that is, (y_j, d_j, r_j) (resp. (y_j, n_j) must satisfy

$$Q_d(y_j, d_j, r_j) := \gamma_d^2 \|d_j\|^2 - \|y_j - r_j\|^2 \geq 0 \quad \text{and} \tag{3.12}$$

$$Q_n(y_j, n_j) := \gamma_h^2 \|n_j\|^2 - \|r_j - y_j\|^2 \geq 0 \tag{3.13}$$

Now, if we assumed that n_j and d_j were i.i.d white noise of known variances σ_n^2 and σ_d^2 , then $\|n_j\|^2$ and $\|d_j\|^2$ would concentrate tightly around $\sigma_n^2 T$ (resp $\sigma_d^2 T$). In other words, we could replace the $\gamma_d^2 \|d_j\|^2$ in the above contraints with, say $(1+\epsilon)\sigma_d^2 T$, and test approximately the same performance specification *without having to guess* d_j. In this regime, we would therefore just want to run the controller K, and see how large the error $\|r_j - y_j\|$ becomes over the course of the trials,

However, the above example only takes place under very strong modeling assumptions that a) the noise is i.i.d and b) its variance is known. When the noise is unknown, and may vary with time, then the constraints in Equation 3.12 become less vaccuous.

The next step is to then characterize: what are the appropriate contraints to place on the n_j, d_j, and w_i to enforce that they resemble "white noise"? A simple small-energy (l_2) contraint

$$\|n_j\|^2 \leq T\sigma_n^2 \quad \|d_j\|^2 \leq T\sigma_d^2 \quad \|w_j\|^2 \leq T\sigma_w^2 \tag{3.14}$$

are a start, but these constraints do not preclude highly correlated, adversarially-chosen sequences of disturbances and observation noise which may interfere with inference. A more appropriate contraint may be "average-uncorrelatedness", in the sense of of the IQC from Section 2. Namely, defining the form

$$Q_{\sigma,\eta,k}(w) = w^T \left(\frac{\sigma}{\sqrt{T}} I + \frac{\eta}{2}((Z^k)^* + Z^k)\right) w \tag{3.15}$$

where $k \in \{1, \ldots, T\}$, $\gamma > 0$ and $\eta \in \{-1, 1\}$, and Z is the $T \times T$ one-step cyclic shift operator. Thus, we might impose contraints of the form

$$Q_{\sigma_w,\eta,k}(w_i) \geq 0 \quad Q_{\sigma_n,\eta,k}(n_j) \geq 0 \quad Q_{\sigma_d,\eta,k}(d_j) \geq 0 \tag{3.16}$$

for $\eta \in \{-1, 1\}$, and $k \in \{1, \ldots, T\}$. Here, the $1/\sqrt{T}$ captures that the correlation between shifted-noise grow as the square-root of the time horizon T, whereas the unshifted auto-correlation should grow linearly in T. While this enforces that each rollout has uncorollated noise, it does not enforce the condition that successive rollouts are independent. This can be enforced as by considering the concatenations $\mathbf{w} = (w_1^T, \ldots, w_N^T)^T$, $\mathbf{d} = (d_1^T, \ldots, d_M^T)^T$, and $\mathbf{n} = (n_1^T, \ldots, n_M^T)^T$ and imposing constraints of the form

$$Q_{\sigma,\eta,k}(\mathbf{w}_i) = \mathbf{w}^T \left(\frac{\sigma}{\sqrt{NT}} I + \frac{\eta}{2}((\mathbf{Z}^k)^* + \mathbf{Z}^k)\right) \mathbf{w} \tag{3.17}$$

where $\mathbf{Z}$ is the shift operator on sequences of length NT.

3.3 Least-Squares Based Confidence Intervals for Quadratic Constraint

A more direct way to constraint P is to create a least-squares confidence interval for the plant p. Here, let $\mathbf{v} = (v_1^T, \dots, v_N^T)^T$ denote the vector in $\mathbb{R}^{TN}$ of all open loop observations, and $\mathbf{w} = (w_1^T, \dots, w_N^T)^T$ the noise vector. Further, define

$$T_{\mathbf{s}} = \begin{bmatrix} T_{s_1} \\ \dots \\ T_{s_N} \end{bmatrix} \in \mathbb{R}^{NT \times T} \tag{3.18}$$

Then,

$$\mathbf{v} = T_{\mathbf{s}} p + \mathbf{w} \tag{3.19}$$

Hence, p and the the least-squares estimator of p, $\widehat{p}$ are given by

$$p = (T_{\mathbf{s}}^T T_{\mathbf{s}})^{-1} T_{\mathbf{s}}^T (\mathbf{v} - \mathbf{w}) \quad \text{and} \quad \widehat{p} = (T_{\mathbf{s}}^T T_{\mathbf{s}})^{-1} T_{\mathbf{s}}^T (\mathbf{v}) \tag{3.20}$$

so that $p - \widehat{p} = (T_{\mathbf{s}}^T T_{\mathbf{s}})^{-1} T_{\mathbf{s}}^T (\mathbf{w})$. Under the wide noise assumption, the error is then a Gaussian vector with covariance $(T_{\mathbf{s}}^T T_{\mathbf{s}})^{-1} T_{\mathbf{s}}^T (\mathbf{w}) \{(T_{\mathbf{s}}^T T_{\mathbf{s}})^{-1} T_{\mathbf{s}}^T (\mathbf{w})\}^T$, which can be used to provide quadratic confidence intervals on p using concentration inequalities for non-isotropic Gaussians (e.g. Hanson-Wright[4])

One could also consider regularizing p to construct Lasso-based estimates, with potentially tighter confidence intervals. Indeed, constraining the FIR p to have a small l_1 makes sense in the case that P is an approximately-low order system, since then its coefficients should decay. This is another avenue we could investigate.

Finally, we remark that if the input matrix $(T_{\mathbf{s}}^T T_{\mathbf{s}})^{-1} T_{\mathbf{s}}^T (\mathbf{w})$ is sufficiently well conditioned, then as $N \to 0$, these Least-Squares and Lasso based estimators are consistent estimators of p.

4 Next Steps and Challenges

At this stage, we have yet to formalize a condition on K which we would expect to allow us to falsify it from finite data. K should fail to meet performance specifications on "typical data", with some sort of margin. It is not clear if we will impose a condition that, with high probability on white noise d_j and n_j, K fails to meet performance standards, or simply with constant probability K fails (thereby requiring many trials to observe the failure).

Further still, we will run into the issue that, as we collect more data from a generative process, we become ever more likely to observe "a-typical" sequences of disturbances which may cause K to fail, *even if K performs adequately on typical sequences.* Thus, we may need to formulate a notion of strong performance on most trials (which is often computationally hard), or work with notions of average performance over rollouts (which are considerably weaker and thus harder to falsify).

Finally, once the correct assumptions are put in place, we need to understand the sampled complexity of the S-procedure in invalidating such a model. One open question is that, while a controler K should be able to be invalidated in the limit of infinite rollouts (indeed, assuming stochastic noise, you can learn P to arbitrary high precision, and use standard robust control-theoretic tools to evaluate its performance), the over-conservative S-procedure may not even be statistically consistent!

References

[1] Moritz Hardt, Tengyu Ma, and Benjamin Recht. Gradient descent learns linear dynamical systems. *arXiv preprint arXiv:1609.05191*, 2016.

[2] Myungsoo Jun and Michael G Safonov. Automatic pid tuning: An application of unfalsified control. *Proc. IEEE CCA/CACSD*, 2:328–333, 1999.

[3] Alexandre Megretski and Anders Rantzer. System analysis via integral quadratic constraints. *IEEE Transactions on Automatic Control*, 42(6):819–830, 1997.

[4] Mark Rudelson, Roman Vershynin, et al. Hanson-wright inequality and sub-gaussian concentration. *Electron. Commun. Probab*, 18(82):1–9, 2013.

[5] Parikshit Shah, Badri Narayan Bhaskar, Gongguo Tang, and Benjamin Recht. Linear system identification via atomic norm regularization. In *2012 IEEE 51st IEEE Conference on Decision and Control (CDC)*, pages 6265–6270. IEEE, 2012.

[6] Roy Smith and Geir Dullerud. Nonlinear functional characterizations of uncertainty in model validation. *IFAC Proceedings Volumes*, 35(1):151–156, 2002.