Nonasymptotic Analysis of Robust Control from Coarse-grained Identification

Stephen Tu, Ross Boczar, Andy Packard, Benjamin Recht

July 10, 2017

Abstract

This work explores the trade-off between the number of samples required to accurately build models of dynamical systems and the degradation of performance in various control objectives due to a coarse approximation. In particular, we show that simple models can be easily fit from input/output data and are sufficient for achieving various control objectives. We derive bounds on the number of noisy input/output samples from a stable linear time-invariant system that are sufficient to guarantee that the corresponding finite impulse response approximation is close to the true system in the $H_\infty$-norm. We demonstrate that these demands are lower than those derived in prior art which aimed to accurately identify dynamical models. We also explore how different physical input constraints, such as power constraints, affect the sample complexity. Finally, we show how our analysis cleanly fits within the established framework of robust control, demonstrating how a controller designed for an approximate system provably meets performance objectives on the true system.

1 Introduction

Most control design relies on establishing a model of the system to be controlled. For simple physical systems, a model with reasonable fidelity can typically be constructed from knowledge of the physics at hand. However, for complex, uncertain systems, building models from first principles becomes quickly intractable and one usually resorts to fitting models from empirical input/output data. This approach naturally raises an important question: How well must we know a system in order to control it?

In this work, we attempt to answer this question by striking a balance between system identification and robust control. We aim to identify coarse estimates of the true underlying model while coupling our estimation with precise probabilistic bounds on the inaccuracy of our estimates. With such a coarse model in hand, we can use standard robust synthesis tools that take into account the derived bounds on the model uncertainty.

More precisely, given an unknown stable discrete-time plant $G$, we bound the error accrued by fitting a finite impulse response (FIR) approximation to $G$ from noisy output measurements. These bounds balance the sample complexity of estimating an unknown FIR filter against the capability of such a filter to approximate the behavior of $G$. Indeed, we show that notably short FIR filters provide a sufficient approximation to stable systems in order to ensure robust performance for a variety of control design tasks. In particular, we demonstrate considerable savings in experimental measurements as compared to other non-asymptotic schemes that aim to precisely identify $G$. 
In the process of fitting a FIR filter, a natural question arises as to what inputs should we use to excite the unknown system. Of course, due to actuator limitations and other physical constraints, we are not free to choose any arbitrary input. Hence, we model the choice of inputs as an experiment design question, where the practitioner specifies a bounded input set and asks for the best \( p \) inputs to use to minimize FIR identification error. We propose a new optimal experiment design procedure for solving this problem, and relate it to the well studied \( A \)-optimal experiment design objective from the statistics literature [17]. This connection is used to study practical cases of input constraints. Specifically, we prove that when the inputs are \( \ell_2 \)-power constrained, then impulse responses are the optimal choice of inputs. However, we show that is not the case when the inputs are \( \ell_\infty \)-constrained. For \( \ell_\infty \) constraints, we construct a deterministic set of inputs which is within a factor of 2 to the optimal solution. Combining these designs with our probabilistic bounds, we show that for estimating a length \( r \) FIR filter \( G_r \), as long as \( p \geq 2r \), the residual \( \mathcal{H}_\infty \) error \( \|G_r - \tilde{G}_r\|_\infty \) on the estimate \( \tilde{G}_r \) satisfies \( \tilde{O}(1/\sqrt{p}) \) with high probability. This is a substantial improvement over the \( \tilde{O}(\sqrt{r}/p) \) scaling which we show occurs in the \( \ell_2 \)-constrained case.

Experimentally, we show that \( \mathcal{H}_\infty \) loop-shaping controller design on the estimated FIR model, using probabilistic bounds, can be used to synthesize controllers with both stability and performance guarantees on the closed loop with the true plant. We also demonstrate that our probabilistic bounds can be estimated directly from data using Monte Carlo techniques.

2 Related Work

Sample complexity guarantees in the system identification literature often require strong assumptions which are difficult to verify. Most analyses are asymptotic and are based on the idea of “persistence of excitation” or “mixing” [21, 12]. There has been some progress in estimating the sample complexity of dynamical system identification using machine learning tools [21, 4], but such results typically yield pessimistic sample complexity bounds that are exponential in the degree of the linear system or other relevant quantities.

Two recent results provide polynomial sample complexity for identifying linear dynamical systems. Shah et al. [20] show that if certain frequency domain measurements are obtained from a linear dynamical system, then the system can be approximately identified by solving a second-order cone programming problem. The degree of the estimated IIR system scales as \( (1 - \rho(A))^{-2} \) where \( \rho(A) \) denotes the stability radius. Similarly, Hardt et al. [8] show that one can estimate an IIR system from time domain observations with a number of measurements polynomial in \( (1 - \rho(A))^{-2} \). In this work, we show that a considerably smaller FIR-approximation with degree \( \tilde{O}((1 - \rho(A))^{-1}) \) suffices to complete many control design tasks.

2.1 Robust Control

Classical robust control literature focuses much of its effort on designing a controller while taking into account fixed bounds on the uncertainty in the model. There are numerous algorithms for controller synthesis under various uncertainty specifications, such as coprime factor uncertainty [13] or state-space uncertainty [15]. However, there are only a few branches of the robust control literature that couple identification to control design, and the identification procedure best suited for a particular control synthesis scheme is usually not specified.
2.2 $H_\infty$ identification and gain estimation

Most related to our work is the literature on $H_\infty$ identification. In this literature, noisy input/output data from an unknown stable linear time-invariant (LTI) plant is collected in either the frequency or time domain; the goal is often to estimate a model with low $H_\infty$ error. For frequency domain algorithms, see e.g. [9, 10], and for time domain algorithms, see e.g. [6]. A comprehensive review of this line of work is given by Chen and Gu [5].

The main difference between the $H_\infty$ identification literature and our work is that we assume a probabilistic noise model instead of worst-case (adversarial), and we assume that our identification algorithm is allowed to pick its inputs to the plant $G$. As we will see, these simplifying assumptions lead to simple algorithms, straightforward analysis, and finite-time sample complexity guarantees.

Another related line of work is the use of the power method [22, 18] for estimating the $H_\infty$-norm of an unknown SISO plant. The key insight in this line of work is that in the SISO case, a time-reversal trick can be applied to effectively query the system $G^* \circ G$. This approach is appealing, since the power method is known to converge exponentially quickly to the leading eigenvector. However, the leading factor in the convergence rate is the ratio of $\lambda_1/\lambda_2$, and hence providing a finite-time guarantee of this method would require a non-asymptotic analysis of the rate of convergence of the second singular value of finite sections of a Toeplitz operator.

2.3 Norms of random polynomials

A significant portion of our analysis relies on bounding the norms of random trigonometric polynomials $Q(z) = \sum_{k=0}^{r-1} \varepsilon_k z^k$. The study of the supremum norm of random finite degree polynomials was first initiated by Salem and Zygmund [19], who studied the setting where the coefficients are drawn from a symmetric Bernoulli distribution supported on $\{\pm 1\}$. Later, Kahane [11] proved that when the coefficients are distributed as an isotropic Gaussian, then with probability at least $1 - \delta$, $\|Q\|_\infty \leq O(\sqrt{r \log(r/\delta)})$. More recently, Meckes [14] extended this result to hold for independent sub-Gaussian random variables by employing standard tools from probability in Banach spaces.

In Section 3, we extend these results to the case when the coefficients follow a non-isotropic Gaussian distribution. This is important because it allows us to reduce the overall error of our estimate by using non-isotropic covariance matrices from experiment design.

3 System Identification of Finite Impulse Responses

In this section, we suppose we are given query access to $G$ via the form

$$f(u; r) = \text{Toep}_r(G)u + \xi, \quad \xi \sim N(0, \sigma^2 I_r).$$

Above, $\text{Toep}_r(G)$ is the restriction of the operator $u \mapsto g * u$ to the first $r$ timesteps, where $r$ is fixed. We assume that $u \in U$ for some compact set $U$. Therefore, the ratio of some measure of the size of $U$ to $\sigma$ serves as the signal-to-noise (SNR) ratio for our setting.

Fix a set of $p$ non-zero inputs $u_1, ..., u_p \in U$. Given a realization of $\{f(u_k; r)\}_{k=1}^p$, we can estimate the first $r$ coefficients $\{g_k\}_{k=0}^{r-1}$ of $G$ via ordinary least-squares (OLS). Calling the vector $Y := (f(u_1; r), ..., f(u_p; r)) \in \mathbb{R}^p$, it is straightforward to show that the least squares estimator
\( \tilde{g}_{0:r-1} \) is given by

\[
\tilde{g}_{0:r-1} := \begin{bmatrix}
\tilde{g}_0 \\
\tilde{g}_1 \\
\vdots \\
\tilde{g}_{r-1}
\end{bmatrix} = (Z^T Z)^{-1} Z^T Y, \quad Z := \begin{bmatrix}
\text{Toep}(u_1) \\
\vdots \\
\text{Toep}(u_p)
\end{bmatrix} \in \mathbb{R}^{p \times r}.
\]

Above, the operator \( \text{Toep}(u) \) takes a vector \( u \in \mathbb{R}^r \) as input and outputs an \( r \times r \) lower-triangular Toeplitz matrix with the first column equal to \( u \). From \( \tilde{g}_{0:r-1} \), we form the estimated finite impulse response \( \tilde{G}_r \) as \( \tilde{G}_r(z) := \sum_{k=0}^{r-1} \tilde{g}_k z^{-k} \). The Gaussian output noise assumption means that the error vector \( \varepsilon := \tilde{g}_{0:r-1} - g_{0:r-1} \) is distributed \( \varepsilon \sim N(0, (Z^T Z)^{-1}) \), and hence \( \tilde{G}_r - G_r \) is equal in distribution to the random polynomial \( Q(z) = \sum_{k=0}^{r-1} \varepsilon_k z^{-k} \).

In this section, we first provide a characterization of the behavior of the random quantity \( \|Q\|_\infty \) as a function of the covariance \( (Z^T Z)^{-1} \) and the polynomial degree \( r \). Next, we study the problem of experiment design to choose the best inputs \( u_1, \ldots, u_p \) to minimize the error \( \|Q\|_\infty \).

### 3.1 A concentration result for the error polynomial

We first address the behavior of the error \( \|Q\|_\infty \). Our main tool is a discretization result from Bhaskar et al. [1]:

**Lemma 3.1** (Bhaskar et al.). Let \( Q(z) := \sum_{k=0}^{r-1} \varepsilon_k z^{-k} \), where \( \varepsilon_k \in \mathbb{C} \). For any \( N \geq 4\pi r \),

\[
\|Q\|_\infty \leq \left( 1 + \frac{4\pi r}{N} \right) \max_{k=0,\ldots,N-1} |Q(e^{jk2\pi/N})|.
\]

Lemma 3.1 immediately reduces controlling the \( H_\infty \)-norm of a finite-degree polynomial to controlling the maxima of a finite set of points on the torus. Hence, controlling the expected value of \( \|Q\|_\infty \) and showing concentration is straightforward. Before the state the result, we define for \( z \in \mathbb{C} \),

\[
\varphi(z) := (1, z, z^2, \ldots, z^{r-1}) \in \mathbb{C}^r, \quad \varphi_1(z) := \text{Re}\{\varphi(z)\}, \quad \varphi_2(z) := \text{Im}\{\varphi(z)\}.
\]

**Lemma 3.2.** Let \( \varepsilon \sim N(0, \Sigma) \) and put \( Q(z) = \sum_{k=0}^{r-1} \varepsilon_k z^{-k} \). Define for \( \ell = 1, 2 \),

\[
\sigma^2_\ell := \sup_{z \in \mathbb{T}} \varphi_\ell(z)^T \Sigma \varphi_\ell(z).
\]

We have that

\[
\mathbb{E}\|Q\|_\infty \leq 4\sqrt{2} \sigma \sqrt{\log(8\pi r)}, \quad (3.2)
\]

where \( \sigma := \max(\sigma_1, \sigma_2) \). Furthermore, with probability at least \( 1 - \delta \), we have

\[
\|Q\|_\infty \leq 4\sqrt{2} \sigma \left( \sqrt{\log(8\pi r)} + \sqrt{\log(2/\delta)} \right), \quad (3.3)
\]

**Proof.** Set \( N = 4\pi r \) and invoke Lemma 3.1 to conclude that

\[
\|Q\|_\infty \leq 2 \max_{k=0,\ldots,4\pi r-1} |Q(e^{jk2\pi r})| \leq 2 \max_{k=0,\ldots,4\pi r-1} |\langle \varphi(e^{jk2\pi r}) , \varepsilon \rangle| \leq 2 \max_{k=0,\ldots,4\pi r-1} |\langle \varphi_1(e^{jk2\pi r}) , \varepsilon \rangle| + 2 \max_{k=0,\ldots,4\pi r-1} |\langle \varphi_2(e^{jk2\pi r}) , \varepsilon \rangle|.
\]
We first prove (3.2) by bounding \( \mathbb{E} \max_{k=0,\ldots,4\pi r-1} |\langle \varphi_\ell(e^{j k/2r}), \varepsilon \rangle| \) for \( \ell = 1,2 \). For a fixed \( k \), we have that \( \langle \varphi_\ell(e^{j k/2r}), \varepsilon \rangle \sim N(0, \varphi_\ell(e^{j k/2r})^T \Sigma \varphi_\ell(e^{j k/2r})) \). By standard results for expected maxima of Gaussian random variables, we have

\[
\mathbb{E} \max_{k=0,\ldots,4\pi r-1} |\langle \varphi_\ell(e^{j k/2r}), \varepsilon \rangle| \leq \max_{k=0,\ldots,4\pi r-1} \sqrt{\varphi_\ell(e^{j k/2r})^T \Sigma \varphi_\ell(e^{j k/2r})} \sqrt{2 \log(8 \pi r)} \\
\leq \sigma \sqrt{2 \log(8 \pi r)} .
\]

This yields (3.2).

For (3.3), using standard concentration results for suprema of Gaussian processes (see e.g. [3]), we have that with probability at least \( 1 - \delta \),

\[
\max_{k=0,\ldots,4\pi r-1} |\langle \varphi_\ell(e^{j k/2r}), \varepsilon \rangle| \leq \mathbb{E} \max_{k=0,\ldots,4\pi r-1} |\langle \varphi_\ell(e^{j k/2r}), \varepsilon \rangle| + \sigma \sqrt{2 \log(1/\delta)} \\
\leq \sigma \sqrt{2 \log(8 \pi r)} + \sigma \sqrt{2 \log(1/\delta)} .
\]

The claim (3.3) now follows from a union bound.

Note that when \( \Sigma = I \), \( \sigma^2 \leq r \) which recovers the known results from [11] up to constants. Furthermore, when \( \Sigma \) is diagonal, \( \sigma^2 \leq \text{Tr}(\Sigma) \). We will exploit this result in the sequel.

### 3.2 Experiment Design

We now consider the problem of choosing a set of inputs \( u \in U \) in order to minimize the expected error of the residual polynomial. Fixing the number of inputs \( p \) and input constraint set \( U \), the optimal experiment design problem is

\[
\text{minimize } \mathbb{E}_{z_1,\ldots,z_p \in U} \|Q\|_\infty \quad \text{s.t. } \Sigma = \sum_{i=1}^{p} \text{Toep}(u_i)^T \text{Toep}(u_i) .
\]

In (3.4) and the sequel, if the covariance matrix \( \Sigma \) is not invertible then we assign the function value \(+\infty\). Problem (3.4) is difficult to solve as written because the expected value does not have a form which is easy to work with computationally. The following approximate design problem provides a good approximation of (3.4). Let \( \{z_1,\ldots,z_m\} \subseteq \mathbb{T} \) denote a grid of points on \( \mathbb{T} \). Consider the problem

\[
\text{minimize } \max_{u_1,\ldots,u_p \in U} \varphi_\ell(z_k)^T \Sigma^{-1} \varphi_\ell(z_k) \quad \text{s.t. } \Sigma = \sum_{i=1}^{p} \text{Toep}(u_i)^T \text{Toep}(u_i) .
\]

The objective (3.5) minimizes the maximum pointwise variance of \( Q(z) \) over all points on the grid \( \{z_1,\ldots,z_m\} \). If the grid is uniformly spaced and \( m \geq 4\pi r \), then by Lemma 3.1 we can interpret (3.5) as minimizing an upper bound to the objective function in (3.4), since

\[
\mathbb{E}_{z \sim N(0,\Sigma^{-1})} \|Q\|_\infty \leq (1 + 4\pi r/m) \mathbb{E} \max_{1 \leq k \leq m} |\langle \varphi(z_k), \varepsilon \rangle| \\
\leq (1 + 4\pi r/m) \sqrt{2 \log(2m)} \sum_{\ell=1}^{2} \max_{1 \leq k \leq m} \sqrt{\varphi_\ell(z_k)^T \Sigma^{-1} \varphi_\ell(z_k)} .
\]

(3.6)
However, (3.5) is non-convex in the \( u_i \). A convex version of the problem can be written by choosing \( p_0 \) inputs \( u_1, \ldots, u_{p_0} \in U \) and solving the semidefinite program (SDP)

\[
\min_{\lambda \in \mathbb{R}^{p_0}} \max_{1 \leq k \leq m} \varphi_\ell(z_k)^T \Sigma^{-1} \varphi_\ell(z_k) \quad \text{s.t.} \quad \Sigma = \sum_{i=1}^{p_0} \lambda_i \text{Toep}(u_i)^T \text{Toep}(u_i), \quad \lambda^T 1 = 1, \quad \lambda \geq 0 . \tag{3.7}
\]

(3.7) is a convex program and can be solved with any off-the-shelf solver such as MOSEK or cvxopt.

We now study two special cases of \( U \) to show how input constraints can affect design. We first observe that when \( \Sigma \) is diagonal, continuing the estimates from (3.6), we have the following upper bound which holds since \( \|\varphi_\ell(z)\|_\infty \leq 1 \),

\[
E_{z \sim N(0,\Sigma^{-1})} \|Q\|_\infty \leq (1 + 4\pi r/m) 2\sqrt{2} \log(2m) \text{Tr}(\Sigma^{-1}) . \tag{3.8}
\]

Even though (3.8) only holds when \( \Sigma \) is diagonal, it motivates us to consider the standard A-optimal design problem

\[
\min_{u_1, \ldots, u_p \in U} \text{Tr}(\Sigma^{-1}) \quad \text{s.t.} \quad \Sigma = \sum_{i=1}^p \text{Toep}(u_i)^T \text{Toep}(u_i) . \tag{3.9}
\]

An advantage of (3.9) versus (3.5) is that the reduced complexity of the objective function allows us to make statements about optimality for special cases of \( U \). The analogous SDP relaxation of (3.9), similar to (3.7), is also more efficient to implement in practice.

Let \( F_i^* \) denote the optimal value of (3.9) with \( U = B_i(1) \) for \( i \in [1, \infty] \). It is not hard to show that \( F_i^* \) is finite and the value is attained (and hence \( \Sigma \) at the optimum is invertible). We will now study the cases \( U = B_2(1) \) and \( U = B_\infty(1) \). For \( B_2(1) \), we will show that setting all inputs to the impulse response \( u_i = e_1 \) is the optimal solution. On the other hand, for \( B_\infty(1) \), we will show that impulse responses are quite sub-optimal.

### 3.2.1 A-optimal design for \( \ell_2 \)-ball

**Lemma 3.3.** We have that \( F_2^* = \frac{r}{p} \) and the value is attained by setting \( u_1 = \ldots = u_p = e_1 \).

**Proof.** Using Schur complements, we rewrite (3.9) as

\[
\min_{u_1, \ldots, u_p \in \mathbb{R}^r : \|u_i\|_2 = 1} \sum_{i=1}^r t_i : M_i := \begin{bmatrix} t_i & e_i^T \\ e_i & \sum_{k=1}^p \text{Toep}(u_k)^T \text{Toep}(u_k) \end{bmatrix} \succcurlyeq 0 , \ i = 1, \ldots, r .
\]

Let \( u_1, \ldots, u_p \) be an optimal solution to this problem. Since \( \Sigma = \Sigma(u_1, \ldots, u_p) \) is invertible, all of its diagonal entries must be positive, and hence \( t_i > 0 \) for all \( 1 \leq i \leq r \).

Next, since \( M_i \succcurlyeq 0 \), all of its principle sub-minors are positive semi-definite as well. Examining \( M_1 \), we have that

\[
0 \prec \begin{bmatrix} t_1 & 1 \\ 1 & \Sigma_{11} \end{bmatrix} = \begin{bmatrix} t_1 & 1 \\ 1 & \sum_{k=1}^p \|u_k\|_2 \end{bmatrix} = \begin{bmatrix} t_1 & 1 \\ 1 & p \end{bmatrix} \iff t_1 \geq p^{-1} .
\]

6
We now repeat this argument for $M_2$. We know that
\[
0 \preceq \begin{bmatrix} t_2 & 0 & 1 \\ 0 & \Sigma_{11} & \Sigma_{12} \\ 1 & \Sigma_{12} & \Sigma_{22} \end{bmatrix} \iff \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12} & \Sigma_{22} \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ 0 & t_2^{-1} \end{bmatrix} \succeq 0 \\
\iff \Sigma_{22} \geq t_2^{-1}.
\]

But we know that $\Sigma_{22} \leq p$, and hence we conclude that $t_2 \geq p^{-1}$. It is clear that this argument can be repeated for $t_3, \ldots, t_r$. In the general case, we have that
\[
\Sigma_{1:k,1:k} - \begin{bmatrix} 0_{k-1 \times k-1} & 0_{k-1 \times 1} \\ 0_{1 \times k-1} & t_k^{-1} \end{bmatrix} \succeq 0 \iff \Sigma_{kk} \geq t_k^{-1}.
\]

However, $\Sigma_{kk} \leq p$, so $t_k \geq p^{-1}$. This proves a lower bound on the objective value as $r/p$. This lower bound is achieved by setting $u_1 = \ldots = u_p = e_1$.

\textbf{3.2.2 $A$-optimal design for $\ell_\infty$-ball}

We first prove a lower bound on the optimal objective value $F_\infty^\ast$. To do this, we use the following linear algebra fact.

\textbf{Lemma 3.4.} Let $A$ be an $n \times n$ positive definite matrix. We have that
\[
\text{Tr}(A^{-1}) \geq \sum_{i=1}^n A_{ii}^{-1}.
\]

\textbf{Proof.} This proof is due to Mateusz Wasilewski [23]. By the Schur-Horn theorem, we know that the eigenvalues of $A$ majorize the diagonal of $A$, i.e.
\[
\sum_{i=1}^k A_{ii} \leq \sum_{i=1}^k \lambda_i(A), \ k = 1, \ldots, n, \ \sum_{i=1}^n A_{ii} = \sum_{i=1}^n \lambda_i(A).
\]

The function $x \mapsto 1/x$ is convex for $x > 0$. This allows us to apply Karamata’s inequality, from which the claim immediately follows.

Lemma 3.4 immediately yields the following lower bound on $F_\infty^\ast$.

\textbf{Lemma 3.5.} Let $H_r$ denote the $n$-th Harmonic number. We have that
\[
F_\infty^\ast \geq \frac{H_r}{p}.
\]

\textbf{Proof.} Let $u_1, \ldots, u_p \in B_\infty(1)$ be such that (3.9) takes on a finite value. By Proposition 3.4,
\[
\text{Tr}(\Sigma^{-1}) \geq \sum_{i=1}^r \Sigma_{ii}^{-1} = \sum_{i=1}^p \left( \sum_{k=1}^{r-i+1} \sum_{\ell=1}^{r-i} (u_k)^2 \right)^{-1} \geq \frac{1}{p} \sum_{i=1}^r \frac{1}{r - i + 1} = \frac{H_r}{p}.
\]
Comparing Lemma 3.5 with Lemma 3.3 suggests that a system with $\ell_2$-power constraints can require substantially more measurements to identify than a system with $\ell_\infty$ constraints. Specifically, Lemma 3.2 states that to satisfy $E\|Q\|_\infty \leq \varepsilon$, the former requires $\tilde{O}(\sqrt{r}/\varepsilon^2)$ measurements whereas the latter requires only $\tilde{O}(1/\varepsilon^2)$ measurements. We will see in Section 4 that, ignoring log factors, $r \approx 1/1-\rho$ is a reasonable upper bound on $r$, and hence for systems which are marginally stable the gap between the two can be large.

We now study two input sequences which either achieve or come within a constant factor of achieving the lower bound $H_r/p$. Both our constructions are deterministic. We believe, however, that random $\pm 1$ inputs should achieve a similar type of guarantee, and we leave this to future work.

**Hadamard construction.** Our first construction is deterministic. Let $r = p = 2^n$. We will show that the optimal input vectors for $A$-design are $r$ orthogonal vectors in $\{-1, +1\}^r$. We give a construction for these vectors in the following proposition; this construction is the standard one for constructing a Hadamard matrix.

**Proposition 3.6.** For $n = 0, 1, 2, ...$, there exists $2^n$ vectors in $\{-1, +1\}^n$ that are orthogonal with respect to the standard $\ell_2$ inner product on $\mathbb{R}^n$.

**Proof.** We will induct on $n = 0, 1, 2, ...$, for which the base case $n = 0$ holds with $u_0 = 1$. Assume we have $2^n$ orthogonal vectors, in $\{-1, +1\}^n$, denoted $\{u_k\}$. Then, the $2^{n+1}$ vectors

$$\{\tilde{u}_k\} := \bigcup_{k=0}^{2^n-1} \left\{ \begin{bmatrix} u_k \\ u_k \end{bmatrix}, \begin{bmatrix} u_k \\ -u_k \end{bmatrix} \right\},$$

which reside in $\{-1, +1\}^{2^n+1}$, are also orthogonal.

**Lemma 3.7.** The constructed orthogonal vectors $\{u_k\}_{k=0}^{2^n-1}$ specified in Proposition 3.6 satisfy

$$M := \sum_{k=0}^{2^n-1} \text{Toep}(u_k)^T \text{Toep}(u_k) = 2^n \text{diag}(2^n, 2^n-1, 2^n-2, \ldots, 1).$$

**Proof.** This follows from straightforward manipulations shown in Appendix A.

Combining Lemma 3.5 and Lemma 3.7 implies that the construction from Proposition 3.6 is optimal for (3.9).

**Sinusoidal construction.** In the general case, when $r$ and $p$ are not equal to the same power of 2, solving for the optimal design is not straightforward. However, we give a deterministic construction that relaxes the power of 2 assumption while remaining optimal up to a constant factor of 2. Our design assumes only that $p$ is even and that $p \geq 2r$.

Let $p = 2n$ with $n \geq r$. Define $\theta_i = \frac{2\pi i}{n}$ for $i = 0, 1, ..., n-1$ and $z_i = e^{j\theta_i}$. Consider the $p$ inputs

$$\text{Re}\{\varphi(z_0)\}, \text{Im}\{\varphi(z_0)\}, \text{Re}\{\varphi(z_1)\}, \text{Im}\{\varphi(z_1)\}, \ldots, \text{Re}\{\varphi(z_{n-1})\}, \text{Im}\{\varphi(z_{n-1})\},$$

where $\varphi$ is defined as in (3.1). Clearly these $p$ inputs are contained in $B_\infty(1)$. We next show that this particular choice of inputs induces a nearly optimal design matrix.
Lemma 3.8. Consider the $p = 2n$ inputs $u_1, ..., u_p$ defined in (3.10). We have

$$M := \sum_{k=1}^{p} \text{Toep}(u_k)^T \text{Toep}(u_k) = \frac{p}{2} \text{diag}(r, r-1, r-2, ..., 1).$$

Proof. We first observe that for any complex matrix $X$,

$$\text{Re}\{X^*X\} = \text{Re}\{X\}^T \text{Re}\{X\} + \text{Im}\{X\}^T \text{Im}\{X\}.$$

Furthermore, $\text{Re}\{\text{Toep}(u)\} = \text{Toep}(\text{Re}\{u\})$ and similarly $\text{Im}\{\text{Toep}(u)\} = \text{Toep}(\text{Im}\{u\})$. Hence,

$$\text{Re}\{\text{Toep}(\varphi(z_i))^* \text{Toep}(\varphi(z_i))\} = \text{Toep}(\text{Re}\{\varphi(z_i)\})^T \text{Toep}(\text{Re}\{\varphi(z_i)\}) + \text{Toep}(\text{Im}\{\varphi(z_i)\})^T \text{Toep}(\text{Im}\{\varphi(z_i)\}).$$

Therefore, we have the identity

$$M = \text{Re}\left\{ \sum_{i=0}^{n-1} \text{Toep}(\varphi(z_i))^* \text{Toep}(\varphi(z_i)) \right\} := \text{Re}\left\{ \sum_{i=0}^{n-1} M_{zi} \right\}.$$

Let us now compute the entries of $(M_z)_{k,\ell}^{-1}$ for $z = e^{j\theta}$. Since $M_z$ is Hermitian, we have $(M_z)_{k,\ell} = (M_z)_{\ell,k}$. Hence we only need to consider the case when $k \leq \ell$. A quick calculation shows that

$$(M_z)_{k,\ell} = (r - \ell)e^{-j\theta(\ell-k)}.$$

Hence, $(M_z)_{kk} = (r - k)$ for $k = 0, ..., r - 1$. On the other hand, when $k < \ell$,

$$\sum_{i=0}^{n-1} (M_{zi})_{k,\ell} = (r - \ell) \sum_{i=0}^{n-1} e^{-j\theta_i(\ell-k)} = (r - \ell) \sum_{i=0}^{n-1} e^{-\frac{2\pi i}{n}(\ell-k)} = (r - \ell) \frac{1 - e^{-2\pi j(\ell-k)}}{1 - e^{-\frac{2\pi}{n}(\ell-k)}} = 0.$$

Above, the last equality holds since we have assumed that $n \geq r$, and hence $\ell - k < n$. Therefore, we have shown that

$$M = n \text{diag}(r, r-1, r-2, ..., 1) = \frac{p}{2} \text{diag}(r, r-1, r-2, ..., 1).$$

4 Finite Truncation Error Analysis for Stable Systems

In Section 3, we presented both probabilistic guarantees and experiment design for identification of FIR systems of length $r$, which were independent of any system specific properties of $G$. In this section, we analyze how system behavior affects the necessary truncation length needed to reach a desired approximation error tolerance.
In order to provide guarantees, we require that the underlying system $G$ is stable with stability radius $\rho \in (0,1)$. A standard fact states that stability is equivalent to the existence of a constant $C > 0$ such that the tail decay on the coefficients of the Laurent expansion $G = \sum_{k=0}^{\infty} a_k z^{-k}$ satisfies the following condition

$$|a_k| \leq C \rho^{k-1}, \quad k \geq 1. \tag{4.1}$$

Under this assumption, a simple calculation reveals that as long as

$$r \geq \frac{1}{1 - \rho} \log \left( \frac{C}{\varepsilon (1 - \rho)} \right) + 1,$$

then we have that the approximation error $\|G - G_r\|_\infty$ satisfies $\|G - G_r\|_\infty \leq \varepsilon$.

We next give a system-theoretic characterization of the quantity $C$. Intuitively, if a system has long transient behavior, then we expect the constant $C$ in (4.1) to be large, since in order to obtain a small approximation error one needs to capture the transient behavior. Our next lemma formalizes this intuition.

**Lemma 4.1.** Let $G(z) = \sum_{k=0}^{\infty} a_k z^{-k}$ be a stable SISO LTI system with stability radius $\rho \in (0,1)$. Fix any $\gamma$ satisfying $\rho < \gamma < 1$. Then for all $k \geq 1$,

$$|a_k| \leq \frac{8}{\pi} \|G(\gamma z)\|_* \gamma^k \leq \frac{8}{\pi} d \|G(\gamma z)\|_\infty \gamma^k,$$

where $d$ denotes the McMillan degree of $G$ and $\|G(\gamma z)\|_*$ is the Hankel nuclear norm of the system $G(\gamma z)$.

**Proof.** We first prove the second inequality. Letting $\sigma_1, \ldots, \sigma_{d_\gamma}$ denote the Hankel singular values of $G_\gamma := G(\gamma z)$ in non-increasing order, we immediately have $\|G_\gamma\|_* = \sum_{i=1}^{d_\gamma} \sigma_i \leq d_\gamma \sigma_1$, where $d_\gamma$ denotes the McMillan degree of $G_\gamma$. To see that $d_\gamma \leq d$, note that if $(A, B, C, D)$ is a minimum realization of $G$, then $(\gamma^{-1} A, \gamma^{-1} B, C, D)$ is a realization of $G_\gamma$. Next, recall that the inequality $\sigma_1 \leq \|G_\gamma\|_\infty$ holds. This is a consequence of the definition of the Hankel operator $\Gamma_{G_\gamma}$ as $\Gamma_{G_\gamma} = P_+ G_\gamma P_-$, where $G_\gamma$ on the RHS is understood as an $\ell_2(\mathbb{Z}) \to \ell_2(\mathbb{Z})$ operator and where $P_+$ (resp. $P_-$) orthogonally projects $\ell_2(\mathbb{Z})$ onto $\ell_2(\{0,1,\ldots\})$ (resp. $\ell_2(\{\ldots,-2,-1\})$). Since projection operators have operator norm bounded by one, we conclude that $\sigma_1 = \|\Gamma_{G_\gamma}\| \leq \|P_+\| \|G_\gamma\| \|P_-\| \leq \|G_\gamma\|$, which yields the second inequality.

Now we prove the first inequality. Let $G_h(z) = \sum_{k=1}^{\infty} a_k z^{-k}$, and let $\mathbb{D} = \{ z \in \mathbb{C} : |z| < 1 \}$. By construction, the system $G_h(\gamma z)$ has all poles contained in $\mathbb{D}$. Hence, by a result in Hankel operator theory (see e.g. [16, Theorem 7.7] and [20, Theorem 3.1]), $G_h(\gamma z)$ has a Coifman-Rochberg expansion

$$G_h(\gamma z) = \sum_{i=1}^{\infty} \lambda_i \frac{1 - |\omega_i|^2}{z - \omega_i}, \tag{4.2}$$

with $(\lambda_i)_{i=1}^{\infty} \in \ell_1(\mathbb{N})$, $\{\omega_i\}_{i=1}^{\infty} \subseteq \mathbb{D}$, and $\sum_{i=1}^{\infty} |\lambda_i| \leq \frac{8}{\pi} \|G_h(\gamma z)\|_*$. We express $G_h(z)$ in this expansion by writing

$$G_h(z) = \sum_{i=1}^{\infty} \lambda_i \frac{1 - |\omega_i|^2}{\gamma^{-1} z - \omega_i}. \tag{4.3}$$
Using the expansion (4.3),

\[ G_h(z) = \sum_{i=1}^{\infty} \lambda_i \gamma z^{-1} \frac{1 - |\omega_i|^2}{1 - \gamma |\omega_i|^2} z^{-1} \sum_{k=0}^{\infty} \gamma^k \omega_i^k z^{-k} \]

\[ = \sum_{i=1}^{\infty} \sum_{k=0}^{\infty} \lambda_i (1 - |\omega_i|^2) \gamma^{k+1} \omega_i^k z^{-(k+1)} \]

\[ = \sum_{k=0}^{\infty} \sum_{i=1}^{\infty} \lambda_i (1 - |\omega_i|^2) \gamma^{k+1} \omega_i^k z^{-(k+1)} \]

\[ = \sum_{k=1}^{\infty} \sum_{i=1}^{\infty} \lambda_i (1 - |\omega_i|^2) \gamma^{k-1} \omega_i^k z^{-k} . \]

Above, the penultimate equality follows since the double summation is absolutely integrable, and hence we can switch the order of summation by Fubini’s theorem. Hence, for each \( k \geq 1 \),

\[ |a_k| = \left| \sum_{i=1}^{\infty} \lambda_i (1 - |\omega_i|^2) \gamma^k \omega_i^{k-1} \right| \leq \gamma^k \sum_{i=1}^{\infty} |\lambda_i| \leq \frac{8}{\pi} \gamma^k \|G_h(\gamma z)\|_\ast . \]

Some remarks are in order regarding the proof of Lemma 4.1. The reader might be wondering why the Coifman-Rochberg expansion of \( G_h(z) \) is not considered directly, and why one needs to introduce the proxy system \( G_h(\gamma z) \). The reason for this is because the Coifman-Rochberg expansion of \( G_h(z) \) only has the weaker property that the coefficients \( \{w_i\} \subseteq \mathbb{D} \), and does not have the stronger property that the \( w_i \)'s satisfy \( |w_i| \leq \rho \). Hence, for a bound of the form in Lemma 4.1, the expansion of \( G_h(z) \) is insufficient. Therefore, the proxy system \( G_h(\gamma z) \) is used to force the expansion to encode a useful quantitative bound. We note that a similar technique is used in [2].

5 Robust Controller Design

In Section 3, we described how to obtain a FIR system \( G_{\text{fir}} \) with a probabilistic guarantee that \( G = G_{\text{fir}} + \Delta \), where \( \Delta \) is an LTI system satisfying \( \|\Delta\|_\infty \leq \varepsilon \). This description of \( G \) naturally lends itself to many robust control synthesis methods. In this section, we detail the application of one particular method based on \( H_\infty \) loop-shaping to a particular unknown plant.

Suppose that \( G \) is itself an FIR described with \( z \)-transform

\[ G(z) = |w_0| + \sum_{k=1}^{149} |w_k| \rho^{k-1} z^{-k} , \quad \rho = 0.95 , \]  

(5.1)

where \( w_k \sim N(0,1) \) are independent Gaussians. In this section, we will detail the design of a reference tracking controller for \( G \) using probabilistic guarantees.

5.1 Computing bounds

While the non-asymptotic bounds of Section 3 and Section 4 give us upper bounds on the error of noisy FIR approximation, the constant factors in the bounds are not optimal. Hence, strictly
relying on the bounds will cause oversampling by a constant factor of (say) 10 or more. For real
systems, this is extremely undesirable– using the sharpest bound possible is of great practical inter-
est. Fortunately, we can do this via simple Monte–Carlo simulations, which we detail in Section B
the appendix. For now, we describe the results of these simulations.

Our first Monte–Carlo simulation establishes that $G$ satisfies the tail decay specified in (4.1)
with $C = 3.9703$ and $\rho = 0.95$. If we truncate $G$ with $r = 75$, we see that our worst-case bound on
$\|G - G_r\|_\infty$ is $\|G - G_r\|_\infty \leq C\frac{\rho^{r-1}}{1-\rho} = 3.9703 \times \frac{0.95^{74}}{1-0.95} = 1.7840$. In general, assuming we have no
other information about $G$ other than the bounds on $C$ and $\rho$, this is the sharpest approximation
error bound possible, since for any system with real-valued, all non-negative Fourier coefficients,
the $H_\infty$-norm is simply the sum of the coefficients.

However, if we further assume we know the structure of $G$ as in this case where we know
the form of (5.1), but not the values of $w_k$, we can further sharper our approximation bound.
Specifically, we know that $E_{\text{approx}} := \|G - G_r\|_\infty = \sum_{k=75}^{149} |w_k| \rho^{k-1}$, and hence we can perform
another Monte–Carlo simulation to estimate the tail probability of this random variable. The result
of our simulation is that $\mathbb{P}(E_{\text{approx}} \leq 0.46) \geq 0.99$. This is a substantial improvement over the
previous bound of $E_{\text{approx}} \leq 1.7840$ which only uses the information contained in the tail decay.

Furthermore, we can use the same trick to sharpest the estimates from Lemma 3.2. We perform
our final Monte–Carlo simulation, this time on the random variable $E_{\text{noise}} := \frac{\sigma}{\sqrt{N}} \|\sum_{k=0}^{74} \xi_k z^{-k}\|_\infty$, with $\sigma^2 = 1$ and $\xi_k \sim N(0,1)$. Note that this corresponds to choosing the inputs to the system as
impulse responses, which we recall from Section 3.2.1 is optimal under $\ell_2$-power constraints. Doing
this simulation, we obtain that $\mathbb{P}(E_{\text{noise}} \leq 3.5954) \geq 0.99$.

5.2 Controller design

![Diagram](attachment:closed-loop.png)

Figure 1: Closed-loop experimental setup. The goal is to design the controller $K$. $G_{\text{fir}}$ is estimated from
noisy output data, and $\|\Delta\|_\infty$ is bounded via Monte–Carlo simulations.

Our goal is to design a controller $K$ in the setup described in Figure 1, under the assumption
that $\|\Delta\|_\infty \leq E_{\text{approx}} + E_{\text{noise}} \leq 4.0554$. This assumption comes from the calculations in Section 5.1.
We note that $\|\Delta\|_\infty / \|G\|_\infty$ fluctuates between 10-20%, so $G_{\text{fir}}$ is a relatively coarse description of
$G$. We use standard loop-shaping performance goals (see e.g. [7]). Let $T_{r\to e}$ and $T_{n\to e}$ denote the
transfer functions from $r \to e$ and $n \to e$, respectively. At low frequencies, we would like $|T_{r\to e}|$
to have small gain, and at high frequencies we would like $|T_{r\to e}| \leq 2$. Similarly, we would like
$|T_{n\to e}| \leq 2$ at low frequencies and $|T_{n\to e}|$ small at high frequencies. Of course, we would like these
goals to be achieved, in addition to closed loop stability, for all $G = G_{\text{fir}} + \Delta$. 

12
We proceed in two steps. We first design a controller with the nominal $G_{\text{fir}}$ using $\mathcal{H}_\infty$ loop-shaping synthesis (mixsyn in MATLAB). We choose weights to encourage our performance goals on $T_{r\rightarrow e}$ and $T_{n\rightarrow e}$ to be met. Next, we check that our performance goal is met, in addition to robust stability. To make the computation easier, we check the performance goals separately. First, it is well known (see e.g. [7]) that the goal on $T_{r\rightarrow e}$ is met (in addition to robust stability) if the following holds

$$\|W_1S| + \gamma|KS||_{\infty} < 1,$$

where $S = \frac{1}{1+KG_{\text{fir}}}$ and $\gamma = 4.0554$. Specifically, under (5.2), the closed loop with $K$ in feedback with $G$ is stable and achieves the performance guarantee $|T_{r\rightarrow e}(z)| \leq \frac{1}{|W_1(z)|}$ for every frequency $z \in \mathcal{T}$. On the other hand, to the best of our knowledge no simple expression for the performance goal on $T_{n\rightarrow e}$ exists, so we resort to a standard structured singular value (SSV) calculation [15].

We generate our controller $K$ via the following MATLAB commands

```matlab
w_c = 0.07; % Cross-over freq
W1 = makeweight(5000, w_c, .5, 1); % Low-freq disturbance rejection
W2 = 1.5*fir_error_bound; % Robust stability
W3 = makeweight(.5, 3 * w_c, 5000, 1); % High-freq noise insensitivity
P = augw(G_fir, W1, W2, W3);
K = hinfsyn(P);
```

In Figure 2, we plot the open loop gain $L = G_{\text{fir}}K$, sensitivity function $S = 1/(1 + L)$, and complementary sensitivity function $T = 1 - S$. Here, we see that the cross-over frequency $\omega_c \approx 0.1$. Next, in Figure 3, we plot the $\mu$ values for both the reference tracking objective $T_{r\rightarrow e}$ and the noise insensitivity objective $T_{n\rightarrow e}$, and check that both curves lies below 1 for all frequencies. Recall that this means that $G$ in feedback with $K$ is not only exponentially stable, but also satisfies both performance guarantees. Finally, in Figure 4, we plot the output $y$ as a function of a noisy square wave input $u$, to show the desired reference tracking behavior, on both the closed loop simulation
Figure 4: Reference tracking behavior of the closed loop with the model $G_{\text{fir}}$ and the actual plant $G$.

Figure 5: Reference tracking behavior as the FIR truncation length is varied from $r = 10, 30, 50, 70$.

(with $G_{\text{fir}}$), and the actual closed loop behavior (with $G$). This shows that, while the model $G_{\text{fir}}$ was a coarse grained description of $G$ with up to 20% relative error, it was faithful enough to allow for a robust controller design.

5.3 Varying truncation length

We next study the effect of truncation length $r$ on controller design. In Figure 5, we assume the same setup and performance goals as the previous section, but vary the truncation length $r \in \{10, 30, 50, 70\}$. We also include the result of a controller design which has full knowledge of the true system $G$, which we label as Opt. We see that for $r = 10$, the resulting controller unsurprisingly has undesirable overshoot behavior. However, as $r$ increases the resulting controller mimics the behavior of Opt quite closely. This plot shows that, at least for reference tracking behavior, a fairly low-fidelity model suffices.

6 Conclusions

This paper explored the use of a coarse-grained FIR model estimated from noisy output data for control design. We showed that sharp bounds on the $H_{\infty}$ error between the true unknown plant and the estimated FIR filter can be derived using tools from concentration of measure, and the constant factors on these bounds can be further refined via Monte–Carlo simulation techniques.

There are many possible future extensions of our work. We highlight a few ideas below.

MIMO Systems. While our approach can be generalized to the MIMO case by estimating filters for each input/output pair separately, we believe that when the MIMO transfer matrix has special structure (e.g. low rank), it should be possible to couple the estimation procedure to reduce the $n^2$ factor increase in sample complexity. This is motivated by the vast literature on compressed sensing, where sparse models embedded in a much larger ambient dimension can be uniquely recovered with at most a logarithmic factor more samples than the degree of the intrinsic sparsity.
**Experiment Design.** On the experiment design side, characterizing the optimal input for all $\ell_p$ balls, in order to understand how the geometry of the input set affects the error, is an interesting future direction. Our analysis for $\ell_2$ and $\ell_\infty$ revealed drastically different optimal input designs, and it would be of theoretical interest to investigate if the intermediate $\ell_p$ balls interpolate between the two extreme designs. On the more practical side, understanding the role of randomness in input design would be beneficial to practitioners of system identification.

**Nonlinear Systems.** Finally, an extension of these techniques to nonlinear systems is another exciting direction. One possible idea is to treat a nonlinear system’s Jacobian linearization as the target unknown system, and fit a FIR using our techniques by exciting the nonlinear system locally. One would expect that the controller designed on the FIR would be valid in a neighborhood, and upon exiting the neighborhood, the process would repeat itself. The challenge here remains to estimate online the regime for which a controller is valid.

**Acknowledgements**

We thank Kevin Jamieson for insightful discussions regarding Monte–Carlo simulation, Anders Rantzer for pointing out references related to this work, and Vikas Sindhwani for providing ideas around linear control techniques for nonlinear systems. RB is supported by the Department of Defense NDSEG Scholarship. AP gratefully acknowledges generous support from the FANUC Corporation as well as the National Science Foundation under grant ECCS-1405413. BR is generously supported by NSF award CCF-1359814, ONR awards N00014-14-1-0024 and N00014-17-1-2191, the DARPA Fundamental Limits of Learning (Fun LoL) Program, a Sloan Research Fellowship, and a Google Faculty Award.

**References**


Appendix

A Proof of Lemma 3.7

We will induct on \( n \), for which the base cases \( n = 1 \) holds. Denote the \( i \)-right shift operator applied to \( u_k \) by \( u_k[i] \). Now, assume the property holds for \( n \). This implies that

\[
M_{ij} = \sum_{k=0}^{2^n-1} u_k[i] \cdot u_k[j] = 0
\]

\[
M_{ii} = \sum_{k=0}^{2^n-1} u_k[i] \cdot u_k[i] = 2^n(2^n - i).
\]

Now, construct \( \{\tilde{u}_k\} \) as before, and note that for \( l = 0, \ldots, 2^n - 1 \),

\[
\tilde{u}_k[l] = \begin{cases} 
0_{i \times 1}, & i \leq 2^n \\
\begin{bmatrix} u_l[i] \\ u_l[i-2^n] \end{bmatrix}, & i > 2^n
\end{cases}
\]

Thus, for \( i \leq 2^n \), when \( i \neq j \),

\[
\tilde{M}_{ij} = \sum_{k=0}^{2^n+1} \tilde{u}_k[i] \cdot \tilde{u}_k[j]
\]

\[
= \sum_{l=0}^{2^n-1} \tilde{u}_2l[i] \cdot \tilde{u}_2l[j] + \tilde{u}_2l+1[i] \cdot \tilde{u}_2l+1[j]
\]

\[
= \sum_{l=0}^{2^n-1} u_l \cdot u_l[i] \cdot u_l[j] - u_l \cdot u_l + u_l[i] \cdot u_l[j] = 0.
\]

Furthermore,

\[
\tilde{M}_{ii} = \sum_{k=0}^{2^n+1} \tilde{u}_k[i] \cdot \tilde{u}_k[i]
\]

\[
= \sum_{l=0}^{2^n-1} u_{2l-1} \cdot \tilde{u}_{2l-1}[i] \cdot \tilde{u}_{2l-1}[i] + \tilde{u}_{2l}[i] \cdot \tilde{u}_{2l}[i]
\]

\[
= \sum_{l=0}^{2^n-1} u_l \cdot u_l[i] \cdot u_l[i] + u_l \cdot u_l + u_l[i] \cdot u_l[i]
\]

\[
= 2(2^n(2^n - i)) + 2(2^n(2^n)) = 2^{n+1}(2^n - i).
\]

A similar calculation holds for \( i > 2^n \). Thus, by induction, the property holds for all \( r = 2^n \).
B Details for Monte–Carlo simulations

In all of our simulations we are faced with the following problem which we describe in some generality. Let $X$ be a random variable distributed according to the law $\mathbb{P}$. We assume we have access to iid samples from $\mathbb{P}$. Our goal is to estimate an upper bound on $\mathbb{P}(X \geq t)$ for a fixed $t \in \mathbb{R}$.

If the law $\mathbb{P}$ admits a density $f(\cdot)$ with respect to the Lebesgue measure, a possible solution could be to solve this problem exactly by numerically integrating

$$\int_{X(\xi) \geq t} X(\xi) f(\xi) \, d\xi.$$  

However, numerical integration does not scale favorably with dimension. For our experiments, $\xi$ is 75-dimensional, which is prohibitive for numerical integration.

An alternative approach to numerical integration is to rely on concentration of measure. Let $X_1, \ldots, X_N$ be iid copies of $X$, and let $\mathbb{P}^N$ denote the product measure $\mathbb{P}^N = \bigotimes_{k=1}^N \mathbb{P}$. Using a Chernoff bound and defining $F_t := \mathbb{P}(X \geq t)$, we have

$$\mathbb{P}^N \left( \frac{1}{N} \sum_{k=1}^N 1\{X_k \geq t\} \leq F_t - \varepsilon \right) \leq e^{-N \cdot D(F_t - \varepsilon, F_t)} ,  \tag{.1}$$

where $D(p, q) = p \log(p/q) + (1 - p) \log((1 - p)/(1 - q))$ is the KL-divergence between two Bernoulli distributions. Given a $\delta \in (0, 1)$, define the random variable $Q$ as the solution to the implicit equation

$$N \cdot D \left( \frac{1}{N} \sum_{k=1}^N 1\{X_k \geq t\}, Q \right) = \log(1/\delta) . \tag{.2}$$

Note that, from a realization of $X_1, \ldots, X_N$, the realization of $Q$ from (.2) can be solved for by numerical root finding. Plugging the definition of $Q$ back into the Chernoff inequality (.1), we conclude that there exists an event $\mathcal{E}$ (in the product $\sigma$-algebra) such that on $\mathcal{E}$ the inequality $F_t \leq Q$ holds, and furthermore $\mathbb{P}^N(\mathcal{E}) \geq 1 - \delta$. This is the methodology which we use to generate all our bounds, with $\delta = 10^{-4}$. Hence, the statements of the form “$F_t \leq \gamma$” in Section 5.1 should be understood as operating under the assumption that our implementation of the simulation chose a particular realization which is contained in the simulator event $\mathcal{E}$ described previously.