Abstract
We describe an optimal adversarial attack formulation against autoregressive time series forecast using Linear Quadratic Regulator (LQR). In this threat model, the environment evolves according to a dynamical system; an autoregressive model observes the current environment state and predicts its future values; an attacker has the ability to modify the environment state in order to manipulate future autoregressive forecasts. The attacker’s goal is to force autoregressive forecasts into tracking a target trajectory while minimizing its attack expenditure. In the white-box setting where the attacker knows the environment and forecast models, we present the optimal attack using LQR for linear models, and Model Predictive Control (MPC) for nonlinear models. In the black-box setting, we combine system identification and MPC. Experiments demonstrate the effectiveness of our attacks.

Introduction
Adversarial learning studies vulnerability in machine learning, see e.g. (Vorobeychik and Kantarcioglu 2018; Joseph et al. 2018; Liu et al. 2017; Biggio and Roli 2017; Lowd and Meek 2005). Understanding optimal attacks that might be carried out by an adversary is important, as it prepares us to manage the damage and helps us develop defenses. Time series forecast, specifically autoregressive model, is widely deployed in practice (Hamilton 1994; Box et al. 2015; Fan and Yao 2008) but has not received the attention it deserves from adversarial learning researchers. Adversarial attack in this context means an adversary can subtly perturb a dynamical system at the current time, hence influencing the forecasts about a future time. Prior work (Alfeld, Zhu, and Barford 2016; 2017) did point out vulnerabilities in autoregressive models under very specific attack assumptions. However, it was not clear how to formulate general attacks against autoregressive models.

There are extensive studies on batch adversarial attacks against machine learning algorithms. But there is much less work on sequential attacks. We say an attack is sequential if the attacker take actions over time. There are batch attacks against support vector machine (Biggio, Nelson, and Laskov 2012; Biggio et al. 2014), deep neural networks (Goodfellow, Shlens, and Szegedy 2014; Nguyen, Yosinski, and Clune 2015), differentially-private learners (Ma, Zhu, and Hsu 2019), contextual bandits (Ma et al. 2018), recurrent neural networks (Papernot et al. 2016), online learning (Wang and Chaudhuri 2018) and reinforcement learning (Ma et al. 2019). Some of these victims are sequential during deployment, but they can be trained from batch offline data; hence they can be prone to batch attacks. In contrast, (Jun et al. 2018) and (Zhang and Zhu 2019) study sequential attacks against autoregressive model, which is closer to these two papers. Meanwhile, control theory is receiving increasing attention from the adversarial learning community (Recht 2018; Zhu 2018; Lessard, Zhang, and Zhu 2018). Our work strengthens this connection.

This paper makes three main contributions: (1) We present an attack setting where the adversary must determine the attack sequentially. This generalizes the setting of (Alfeld, Zhu, and Barford 2016; 2017), where the adversary can decide the attack after observing all environmental state values used for forecast. (2) We formulate the attacks as an optimal control problem. (3) When the attacker know the environmental dynamics and forecast model (white-box setting), we solve the optimal attacks with Linear Quadratic Regulator (LQR) for the linear case, or Model Predictive Control (MPC) and iterative LQR (iLQR) for the nonlinear case; when the attacker does not know the environmental or forecaster(black-box setting), we additionally perform system identification.

The Attack Setting
Autoregressive Review
To fix notation, we briefly review time series forecasting using autoregressive models. There are two separate entities:

1. The environment is a fixed dynamical system with scalar-valued states $x_t \in \mathbb{R}$ at time $t$. The environment has a (potentially non-linear) $q$-th order dynamics $f$ and is subject to zero-mean noise $w_t \in \mathbb{R}$ with $\mathbb{V}(w_t) = \sigma^2$. Without manipulations from the adversary, the environmental state
evolves as
\[ x_{t+1} = f(x_t, \ldots, x_{t-q+1}, u_t) \]  
(1)
for \( t = 0, 1, \ldots \). We take the convention that \( x_i = 0 \) if \( i < 0 \). We allow the dynamics \( f \) to be either linear or nonlinear.

2. The forecaster makes predictions of future environmental states, and will be the victim of the adversary attack. In this paper we mainly focus on a fixed linear \( AR(p) \) autoregressive forecaster, regardless of whether the environment dynamics \( f \) is linear or not. Even though we allow the forecast model to be nonlinear in black-box setting, we use linear function to approximate the nonlinear autoregressive model. We also allow the possibility \( p \neq q \). At time \( t \), the forecaster observes \( x_t \) and uses the \( p \) most recent observations \( x_{t-p+1}, \ldots, x_{t} \) to forecast the future values of the environmental state. A forecast is made at time \( t \) about a future time \( t' > t \), we use the notation \( y_{t'|t} \) to denote it.

Specifically, at time \( t \) the forecaster uses a standard \( AR(p) \) model to predict. It initializes by setting \( y_{t+1-i|t} = x_{t+1-i} \) for \( i = 1, \ldots, p \). It then predicts the state at time \( t+1 \) by
\[ y_{t+1|t} = \hat{\alpha}_0 + \sum_{i=1}^{p} \hat{\alpha}_i y_{t+1-i|t}, \]  
(2)
where \( \hat{\alpha}_0, \hat{\alpha}_1, \ldots, \hat{\alpha}_p \) are coefficients of the \( AR(p) \) model. We allow the \( AR(p) \) model to be a nonlinear function in the black-box setting. The \( AR(p) \) model may differ from the true environment dynamics \( f \) even if \( f \) is linear; for example, the forecaster may have only obtained an approximate model from a previous learning phase. Once the forecaster predicts \( y_{t+1|t} \), it can plug the predictive value in (2), shift time by one, and predict \( y_{t+2|t} \), and so on. Note all these predictions are made at time \( t \). In the next iteration when the true environment state evolves to \( x_{t+1} \) and is observed by the forecaster, the forecaster will make predictions \( y_{t+2|t+1}, y_{t+3|t+1} \), and so on.

The Attacker
We next introduce a third entity – an adversary (a.k.a. attacker) – who wishes to control the forecaster’s predictions for nefarious purposes. The threat model is characterized by three aspects of the adversary:

(i) Knowledge: In the white-box setting, the attacker knows everything above; in the black-box setting, neither environmental dynamics nor forecaster model are known to the attacker.

(ii) Goal: The adversary wants to force the forecaster’s predictions \( y_{t'|t} \) to be close to some given adversarial accuracy target \( \gamma_{t'|t} \) (the dagger is a mnemonic for attack), for selected pairs of \((t, t')\) of interest to the adversary. Furthermore, the adversary wants to achieve this “small attacks”. These will be made precise below.

(iii) Action: At time \( t \) the adversary can add \( u_t \in \mathbb{R} \) (the “control input”) to the noise \( w_t \). Together \( u_t \) and \( w_t \) enter the environment dynamics via:
\[ x_{t+1} = f(x_t, \ldots, x_{t-q+1}, u_t + w_t). \]  
(3)
We call this the state attack because it changes the underlying environmental states, see Figure 1.

White-Box Attack as Optimal Control
We now present an optimal control formulation for the white-box attack. Following control convention (Lee and Markus 1967; Kwakernaak and Sivan 1972), we rewrite several quantities from the previous section in matrix form, so that we can define the adversary’s state attack problem by the tuple \((F, C, \{y^i_{t'|t}\}, \{Q_{t'|t}\}, R, T)\). These quantities are defined below.

We introduce a vector-valued environment state representation (denoted by boldface) \( x_t := (x_t, x_{t-1}, \ldots, x_{t-p+1})^\top \in \mathbb{R}^p \). The first entry \( x_0 \) serves as an offset for the constant term \( \hat{\alpha}_0 \) in (2). We let \( x_0 \) be the known initial state. It is straightforward to generalize to an initial distribution over \( x_0 \), which adds an expectation in (9). We rewrite the environment dynamics under adversarial control (3) as \( x_{t+1} = F(x_t, u_t, w_t) := (f(x_t, \ldots, x_{t-p+1}, u_t + w_t), x_{t+1}, \ldots, x_{t-p+2})^\top \). If \( f \) is nonlinear, so is \( F \).

In control language, the forecasts are essentially measurements of the current state \( x_t \). In particular, we introduce a vector of \( p \) predictions made at time \( t \) about time \( t' = t + 1, \ldots, t'+q \) as \( y_{t'|t} := (1, y_{t+1|t}, y_{t+2|t}, \ldots, y_{t'+q-1|t})^\top \in \mathbb{R}^{p+1} \). (For completeness, we let \( y_{t'|t} = x_t \) when \( t' \leq t \).) The forecaster’s \( AR(p) \) forecast model specifies the (linear) measurements as follows. We introduce the \((p+1) \times (p+1)\) measurement (i.e. forecast) matrix \( C \),
\[ C = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ \hat{\alpha}_0 & \hat{\alpha}_1 & \hat{\alpha}_2 & \cdots & \hat{\alpha}_{p-1} & \hat{\alpha}_p \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{bmatrix}. \]  
(4)

The adversary usually only has an estimate of \( f \). It is likely that the forecaster’s \( AR(p) \) model has a different order \( p \) than the environment dynamics’ order \( q \). For simplicity, we will assume \( p = q \) below, but explain how to handle \( p \neq q \) in Appendix A.
There are cases that \( y_{t+1} \) does not depend on \( x_t \) but has been decided at time \( t < t \). In such cases, we can simply redefine \( y_{t+1} \) to be \( y_{t+1} \) and rewrite the prediction matrix. For simplicity, we assume \( y_{t+1} \) always depends on \( x_t \) in the rest of the paper.

We also vectorize adversarial reference target: \( y_{t+1} := (1, y_{t+1}, y_{t+2}, \ldots, y_{t+p+1})^T \in \mathbb{R}^{p+1} \). We simply let \( y_{t+1} = 0 \) when \( t' \leq t \); this is non-essential. In fact, for \( (t', t) \) pairs that are uninteresting to the adversary, the target value \( y_{t+1} \) can be undefined as they do not appear in the control cost later.

The cost of the adversary consists of two parts: (1) how closely the adversary can force the forecaster’s predictions to match the adversarial reference targets; (2) how much control \( u_t \) the adversary has to exert. The \( (p+1) \times (p+1) \) matrices \( Q_{t+1} \) define the first part, namely the cost of the attacker failing to achieve reference targets. In its simplest form, \( Q_{t+1} \) is a \((p+1) \times (p+1) \) matrix with all zero entries except for a scalar weight \( \beta_{t+1} \in [0, 1] \) at (2, 2). In this case, \( Q_{t+1} \) simply picks out the \((t' \mid t)\) element:

\[
(y_{t+1} - y_{t+1})^T Q_{t+1} (y_{t+1} - y_{t+1}) = \beta_{t+1} (y_{t+1} - y_{t+1})^2.
\]

For simplicity, we use \( \|y_{t+1} - y_{t+1}\|_{Q_{t+1}}^2 \) to denote (6). Critically, by setting the weights \( \beta_{t+1} \) the attacker can express different patterns of attack. For example:

- If \( \beta_{t+1} = 1 \) for all \( t = 0 \ldots T - 1 \) and 0 otherwise, the adversary cares about all the predictions made at all times about the final time horizon \( T \). In this case, it is plausible that the adversarial target \( y_{t+1} := y_{t} \) is a constant w.r.t. \( t \).
- If \( \beta_{t+1} = 1 \) for all \( t \) and 0 otherwise, the adversary cares about all the forecasts about “tomorrow.”
- If \( \beta_{t+1} = 1, \forall t, t' \), the adversary cares about all predictions made at all times.

Obviously, the adversary can express more complex temporal attack patterns. The adversary can also choose \( \beta_{t+1} \) value in between 0 and 1 to indicate weaker importance of certain predictions.

The \( R \) matrix defines the second part of the adversary cost, namely how much control expenditure \( u_t \) the adversary has to exert. In the simplest case, we let \( R \) be a scalar \( \lambda > 0 \):

\[
\|u_t\|_R^2 := \lambda u_t^2.
\]

We use \( T \) to denote the prediction time horizon: \( T \) is the last time index (expressed by \( t' \)) to be predicted by the forecaster. We define the adversary’s expected quadratic cost \( J(u_{0:(T-1)}) \) for action sequence \( u_0, \ldots, u_{T-1} \) by

\[
J(u_{0:(T-1)}) = \sum_{t=1}^{T-1} \sum_{t'=t+1}^{T} \|y_{t+1} - y_{t+1}\|_{Q_{t+1}}^2 + \sum_{t=0}^{T-1} \|u_t\|_R^2 | x_0 \rangle.
\]

Since the environment dynamics can be stochastic, the adversary must seek attack policies \( \phi_t : \mathbb{R}^{p+1} \mapsto \mathbb{R} \) to map the observed state \( x_t \) to an attack action:

\[
u_t = \phi_t(x_t) \quad \forall t.
\]

Given an adversarial state attack problem \((F, C, {\{y_{t+1}\}_{t}}, \{Q_{t+1}\}, R, T)\), we formulate the optimal state attack as the following optimal control problem:

\[
\min_{\phi_0, \ldots, \phi_{T-1}} J(u_{0:(T-1)})
\]

s.t. \( x_0 \) given

\[
u_t = \phi_t(x_t), \quad t = 0, 1, \ldots, T - 1
\]

\( x_{t+1} = F(x_t, u_t, w_t), \quad t = 0, 1, \ldots, T - 1
\]

\( y_{t+1} = C^T x_t, \quad \forall t' > t
\)

We next propose solutions to this control problem for linear \( F \) and nonlinear \( F \), respectively. For illustrative purpose, we focus on solving the problem when the attack target is to affect \( x_T \). This implies \( \beta_{t+1} = 0 \) when \( t' \geq t + 2 \). Under this assumption, \( J(u_{0:(T-1)}) \) has the following form:

\[
\mathbb{E}_{x_0, T-1} \sum_{t=1}^{T-1} \|C x_t - y_{t+1}\|_{Q_{t+1}}^2 + \sum_{t=0}^{T-1} \|u_t\|_R^2 | x_0 \rangle.
\]

More attack targets are studied in the experiment section.

**Solving Attacks Under Linear \( F \)**

When the environment dynamics \( f \) is linear, the scalar environment state evolves as \( x_{t+1} = \alpha_0 + \sum_{i=1}^{p} \alpha_i x_{t+i-1} + u_t + w_t \), where the coefficients \( \alpha_0, \ldots, \alpha_p \) in general can be different from the forecaster’s AR\( p \) model (2). We introduce the corresponding vector operation:

\[
x_{t+1} = F(x_t, u_t, w_t) := Ax_t + B(u_t + w_t),
\]

where \( A \) has the same structure as \( C \) in (4) except each \( \alpha \) is replaced by \( \alpha \) and \( B = \{0, 1, \ldots, 0\}^T \). The adversary’s attack problem (9) reduces to stochastic Linear Quadratic Regulator (LQR) with tracking, which is a fundamental problem in control theory (Kwakernaak and Sivan 1972). It is well known that such problems have a closed-form solution, though the specific solution for stochastic tracking is often omitted from the literature. In addition, the presence of a forecaster in our case alters the form of the solution. Therefore, for completeness we provide the solution in Algorithm 1.

The derivation is in Appendix. Once the adversarial control policies are computed, the optimal attack sequence is given by \( u_t = \phi_t(x_t), t = 0, 1, \ldots, T - 1 \). The astute reader will notice that, \( u_{T-1} = \phi_{T-1}(x_{T-1}) = 0 \). This is to be expected: \( u_{T-1} \) affects \( x_T \), but it cannot affect forecasts after the prediction time horizon \( T \), which the adversary does not care. To minimize the control expenditure, the adversary’s rational behavior is to set \( u_{T-1} = 0 \).

**Solving Attacks Under Non-Linear \( F \)**

When \( f \) is nonlinear the optimal control problem (9) in general does not have a closed-form solution. Instead, we introduce an algorithm that combines Model Predictive Control (MPC) (Garcia, Prett, and Morari 1989; Kouvaritakis...
with the actual noise instantiation \( w \).

There are two issues that make the problem hard: the ex-

For planning we simply replace the random variables

Known as "nominal cost" in (Kouvaritakis and Cannon 2015).

To address the first issue, we adopt an approximation technique

Solved, which is the case for our nonlinear

\( \{ w_{t+1} \}_{t=\tau}^{\infty} \)

As the outer loop and Iterative Linear regulator (ILQR) (Li and Todorov 2004) as the inner loop to find an approximately optimal attack. While these techniques are standard in the control community, to our knowledge our algorithm is a novel application of the techniques to adversarial learning.

The outer loop performs MPC, a common heuristic in non-

A Greedy Control Policy as the Baseline State

\( \phi^G(x_t) \in \arg \min_u g_t(x_t, u) \). When \( F \) is linear, \( \phi^G(x_t) \) can be obtained in closed-form. We show the solution in Appendix D.

When \( f \) is nonlinear, we let noise \( u_t = 0 \) and solve the following nonlinear problem using numerical solvers:

Black-Box Attack via System Identification

We now consider black-box attack setting where the environment dynamics \( f \) and forecaster's model \( C \) are no longer
The buffer produces the environmental state value $x_{t+1} = F(x_t, u_t, 0)$; then uses MPC and LQR to design an attack action.

The attacker use linear models to estimate both the environmental dynamics and forecast model. At time $t$, the forecast model is estimated over the visible forecasts: $y_{t|t}, (t, t') \in \Lambda(t - b, t)$.

We can summarize the proposed black-box attack method in Algorithm 4.

**Experiments**

We now demonstrate the effectiveness of control-based white-box attacks on time series forecast problems. We compare the optimal attacks computed by LQR (for linear $f$), MPC+ILQR (for nonlinear $f$), black-box attack, greedy attacks, and the no-attack baseline. While the attack actions were optimized under an expectation over random noise $w$ (c.f. (9)), in the

known to the attacker. Both the environment forecast models are allowed to be nonlinear. The attacker will perform system identification (Dean et al. 2017), and solve LQR as an inner loop and MPC as an outer loop.

The attacks picks an estimation model order $p$ for both the environment and forecaster. It also picks a buffer length $b$. In the first $b + p - 1$ iterations, the attacker does no attack but collects observations on the free-evolving environment and the forecasts. Then, in each subsequent iteration the attacker estimates a linear environmental model $\hat{a}$ and linear forecast model $\hat{c}$ using a rolling buffer of previous $b + p - 1$ iterations. The buffer produces $b$ data points for the attacker to use MLE to solve $p + 1$ unknowns in environment model. The attacker then uses MPC and LQR to design an attack action.

The attacker use linear models to estimate both the environmental dynamics and forecast model. At time $t$, the environmental dynamics is estimated over the environmental state value $x_{t-b-p+1}, \ldots, x_t$ and action sequences $u_{t-b}, \ldots, u_{t-1}$:

$$\min_{a_0, p} \sum_{t=b}^{t-1} \left(a_0 + \sum_{i=1}^{p} a_i x_{t+1-i} + u_t - x_{t+1}\right)^2, \quad (19)$$

we use $\hat{a}(t, b)$, a $p + 1$-dimensional column vector to denote the minimum point of (19). Due to the nature of attacking autoregressive models, $B = (0, 1, 0, \ldots, 0) \in \mathbb{R}^{(p+1) \times 1}$ is always known to the attacker.

We use $\hat{c}_0, \hat{c}_1, \ldots, \hat{c}_p$ to denote the estimation of forecast model and let

$$C(\hat{c}_0, \hat{c}_1, \ldots, \hat{c}_p) = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ \hat{c}_0 & \hat{c}_1 & \hat{c}_2 & \cdots & \hat{c}_{p-1} & \hat{c}_p \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{bmatrix}. \quad (20)$$

To denote the corresponding forecast matrix. Let $\Lambda(t_1, t_2)$ denote the set of prediction indices which are visible to the attacker: $\Lambda(t_1, t_2) = \{(t, t') \mid t_1 \leq t \leq t_2, y_{t|t'} \text{ is visible to the attacker}\}$. At time $t$, the forecast model is estimated over the visible forecasts: $y_{t|t}, (t, t') \in \Lambda(t - b, t)$.

We use $\hat{c}(t, b)$, a $p + 1$-dimensional column vector to denote the minimum point of (21). If the attacker only observes sufficient predictions for “tomorrows”, i.e. $\Lambda(t_1, t_2) = \{(t, t') \mid t_1 \leq t \leq t_2, t' = t + 1\}$, then $\hat{c}_0, \hat{c}_1, \ldots, \hat{c}_p$ is the OLS solution to (21). However, for more complex prediction pattern, (21) might involve polynomials of $\hat{c}_0, \hat{c}_1, \ldots, \hat{c}_p$.

We can summarize the proposed black-box attack method in Algorithm 4.
experiments we report the actual realized cost based on the noise instantiation that the algorithm experienced:
\[
\sum_{t=1}^{T-1} \sum_{t'=t+1}^{T} \|C^{t'-t}x_t - y_{t|t}'\|_Q^2 + \sum_{t=0}^{T-1} \|u_t\|_R^2
\]
where the noise sequence \( \{w_{t'}\}_{t'=0}^{T-2} \) is incorporated implicitly in \( x_t \), together with the actual attack sequence \( \{u_{t'}\}_{t'=1}^{T-1} \). To make the balance between attack effect (the quadratic terms involving \( Q \)) and control expenditure (the term involving \( R \)) more interpretable, we let \( R := \lambda = \beta \sum_{t=1}^{T-1} \sum_{t'=t+1}^{T} \beta |t'| / T \).

The Effect of \( Q_{t'|t} \) on Attack

In our first synthetic example we demonstrate the adversary’s ability to target different parts of the forecasts via \( Q \), the quadratic coefficient in cost function. Figure 2 illustrates three choices of attack targets \( Q \): attack “tomorrow”, “last day” and “all predictions”.

![Figure 2: LQR solution on three attack patterns Q. y-axis shows the value of x and u. Each blue forecast curve beginning at time t shows the sequence \( \{y_{t|t-1}\}_{t=1}^{T} \). The pattern of attack defined by the corresponding \( Q \) is highlighted with * on the forecast curves.](image)

Figure 3: LQR vs. Greedy attacks. The black horizontal lines in the right plot mark the vertical axis range of the middle plot.

We now show the LQR attack policy is better than the greedy attack policy. We let the environment evolves by an \( AR(3) \) model: \( x_{t+1} = f(x_t, x_{t-1}, x_{t-2}, w_t) = 0.4x_t - 0.3x_{t-1} - 0.7x_{t-2} + w_t \), \( w_t \sim N(0, 0.1^2) \). The initial values are \( x_0 = 10, x_{-1} = x_{-2} = 0 \), prediction horizon \( T = 15 \). This environment dynamic is oscillating around 0.

We let the forecaster’s model be: \( x_{t+1} = 0.41x_t - 0.29x_{t-1} - 0.68x_{t-2} \), \( Q \) is “tomorrows”. The attacker wants the forecaster to predict a sequence oscillating with smaller amplitude, \( y_{t+1|t} \) is set as following: we simulate \( x_{t+1} = f(x_t, x_{t-1}, x_{t-2}, 0) \), then, let the attack reference target be \( y_{t+1|t-1} = 0.5x_t \). We set \( \beta = 0.1 \).

For simplicity, we let all adversarial reference targets \( y_{t+1|t} \) be the constant 1. We let the environment evolve according to an \( AR(1) \) model: \( x_{t+1} = 1 + 0.5 x_t + w_t \). We let the noise \( w_t \sim N(0, 0.1^2) \) and the initial state \( x_0 = 0 \). We simulate the case where the forecaster has only an approximation of the environment dynamics, and let the forecaster’s model be \( x_{t+1} = 0.9 + 0.6x_t \) which is close to, but different from, the environment dynamics. For illustrative purpose, we set the prediction time horizon \( T = 10 \). Recall that the attacker can change the environment state by adding perturbation \( u_t \):
\[
x_{t+1} = 1 + 0.5x_t + u_t + w_t.
\]
We set \( \beta = 0.1 \).

We run LQR and compute the optimal attack sequences \( u \) under each \( Q \) scenario. They are visualized in Figure 2. Each attack is effective: the blue *’s are closer to the green target line on average, compared to where they would be in the upper-left no-attack panel. Different target selection \( Q \) will affect the optimal attack sequence.

**Comparing LQR vs. Greedy Attack Policies**

We now show the LQR attack policy is better than the greedy attack policy. We let the environment evolves by an \( AR(3) \) model: \( x_{t+1} = f(x_t, x_{t-1}, x_{t-2}, w_t) = 0.4x_t - 0.3x_{t-1} - 0.7x_{t-2} + w_t \), \( w_t \sim N(0, 0.1^2) \). The initial values are \( x_0 = 10, x_{-1} = x_{-2} = 0 \), prediction horizon \( T = 15 \). This environment dynamic is oscillating around 0.

We let the forecaster’s model be: \( x_{t+1} = 0.41x_t - 0.29x_{t-1} - 0.68x_{t-2} \), \( Q \) is “tomorrows”. The attacker wants the forecaster to predict a sequence oscillating with smaller amplitude, \( y_{t+1|t} \) is set as following: we simulate \( x_{t+1} = f(x_t, x_{t-1}, x_{t-2}, 0) \), then, let the attack reference target be \( y_{t+1|t-1} = 0.5x_t \). We set \( \beta = 0.1 \).

For simplicity, we let all adversarial reference targets \( y_{t+1|t} \) be the constant 1. We let the environment evolve according to an \( AR(1) \) model: \( x_{t+1} = 1 + 0.5 x_t + w_t \). We let the noise \( w_t \sim N(0, 0.1^2) \) and the initial state \( x_0 = 0 \). We simulate the case where the forecaster has only an approximation of the environment dynamics, and let the forecaster’s model be \( x_{t+1} = 0.9 + 0.6x_t \) which is close to, but different from, the environment dynamics. For illustrative purpose, we set the prediction time horizon \( T = 10 \). Recall that the attacker can change the environment state by adding perturbation \( u_t \):
\[
x_{t+1} = 1 + 0.5x_t + u_t + w_t.
\]
We set \( \beta = 0.1 \).

We run LQR and compute the optimal attack sequences \( u \) under each \( Q \) scenario. They are visualized in Figure 2. Each attack is effective: the blue *’s are closer to the green target line on average, compared to where they would be in the upper-left no-attack panel. Different target selection \( Q \) will affect the optimal attack sequence.

**Comparing LQR vs. Greedy Attack Policies**

We now show the LQR attack policy is better than the greedy attack policy. We let the environment evolves by an \( AR(3) \) model: \( x_{t+1} = f(x_t, x_{t-1}, x_{t-2}, w_t) = 0.4x_t - 0.3x_{t-1} - 0.7x_{t-2} + w_t \), \( w_t \sim N(0, 0.1^2) \). The initial values are \( x_0 = 10, x_{-1} = x_{-2} = 0 \), prediction horizon \( T = 15 \). This environment dynamic is oscillating around 0.

We let the forecaster’s model be: \( x_{t+1} = 0.41x_t - 0.29x_{t-1} - 0.68x_{t-2} \), \( Q \) is “tomorrows”. The attacker wants the forecaster to predict a sequence oscillating with smaller amplitude, \( y_{t+1|t} \) is set as following: we simulate \( x_{t+1} = f(x_t, x_{t-1}, x_{t-2}, 0) \), then, let the attack reference target be \( y_{t+1|t-1} = 0.5x_t \). We set \( \beta = 0.1 \).

For simplicity, we let all adversarial reference targets \( y_{t+1|t} \) be the constant 1. We let the environment evolve according to an \( AR(1) \) model: \( x_{t+1} = 1 + 0.5 x_t + w_t \). We let the noise \( w_t \sim N(0, 0.1^2) \) and the initial state \( x_0 = 0 \). We simulate the case where the forecaster has only an approximation of the environment dynamics, and let the forecaster’s model be \( x_{t+1} = 0.9 + 0.6x_t \) which is close to, but different from, the environment dynamics. For illustrative purpose, we set the prediction time horizon \( T = 10 \). Recall that the attacker can change the environment state by adding perturbation \( u_t \):
\[
x_{t+1} = 1 + 0.5x_t + u_t + w_t.
\]
We set \( \beta = 0.1 \).

We run LQR and compute the optimal attack sequences \( u \) under each \( Q \) scenario. They are visualized in Figure 2. Each attack is effective: the blue *’s are closer to the green target line on average, compared to where they would be in the upper-left no-attack panel. Different target selection \( Q \) will affect the optimal attack sequence.
The attacker can change the noise to be $w_t \sim N(0, 0.1^2)$. The dynamic is $x_{t+1} = 2x_t/(1 + 0.8x_t^2) + w_t$. We let $x_0 = 3$, $T = 50$. We simulate this dynamical system to $t = 50$ and get a sample sequence from this dynamical system. The forecaster’s $AR(1)$ model $C$ is estimated from this sequence. We introduce an attacker who can add perturbation $u_t$ to change the state value: $x_{t+1} = 2x_t/(1 + 0.8x_t^2) + u_t + w_t$, $Q$ is “tomorrow”. The attack target is set to be $y_{t+1|t} = 2$. We let $\lambda = 0.01$.

Both the attacker and the oracle do nothing but observe the $x_t$ and $y_{t|t}$ at time $t = 0, \ldots, b + p - 2$, and attack the forecast at time $t = b + p - 1, \ldots, T - 2$. For system identification, we let $b = 15$, $l = 5$, $p = 3$. For the oracle, the time step of MPC is set to be $l = 10$. Inside the MPC loop, the stopping condition of iLQR is $tol = 10^{-4}$. The maximum iteration of iLQR is set to be 1000.

We run 100 trials, the last one is shown in Figure 5. The mean actual realized cost of system identification attack and oracle MPC+iLQR attack are 4.20, 1.43 respectively. Even though the cost of System identification attack is larger than that of the oracle, it is evident from Figure 5 that the attacker can quickly force the forecasts (blue) to the attack target (green) after $t=25$; the chaotic period between $t=20$ and $t=25$ is the price to pay for system identification.

### Acknowledgment

We thank Laurent Lessard, Yuzhe Ma, and Xuezhou Zhang for helpful discussions. This work is supported in part by NSF 1836978, 1545481, 1704117, 1623605, 1561512, the MADLab AF Center of Excellence FA9550-18-1-0166, and the University of Wisconsin.

### Conclusion

In this paper we formulated adversarial attacks on autoregressive model as optimal control. This sequential attack problem differs significantly from most of the batch attack work in adversarial machine learning. In the white-box setting, we obtained closed-form LQR solutions when the environment is linear, and good MPC approximations when the environment is nonlinear. In the black-box setting, we propose a method via system identification then perform MPC. We demonstrated their effectiveness on synthetic and real data.
