A Relaxed Projection Control in the Context of Inverse Optimal Control for Discrete Nonlinear Systems

Ethan King[1] and Hien Tran[2]

Abstract— **We present a nonlinear discrete time feedback control design by searching over a family of controls and associated candidate quadratic control Lyapunov functions parameterized by the symmetric positive definite matrices using an ensemble Kalman search. Then we propose a novel relaxation of the control which can be used to tune for a particular implementation. We show that the proposed family of controls are globally exponentially stable for linear systems, and test the controls on both a linear and nonlinear example.**

I. INTRODUCTION

Designing controls for nonlinear systems remains a challenging problem. A classical approach first defines a performance cost functional of the state and control values, then constructs the control which minimizes the cost. Controls that are optimal with respect to a meaningful cost functional often have favorable properties such as robustness to measurement and model noise, and avoid counterproductive control effort [1], [2]. For linear systems, optimal controls can be constructed by the solution of the Hamilton Jacobi Bellman (HJB) equation corresponding to the cost [1]. However, for nonlinear systems the HJB equation is generally intractable, though many methods do exist for approximating the optimal control [3], [4], [5]. An alternative strategy is to approach the problem from an inverse optimal perspective, by directly designing a control for a system, then validating it as optimal with respect to a constructed cost functional.

Direct design of a control can be done by using a control Lyapunov function (CLF). For instance Freeman and Kokotovic showed that given a CLF for a continuous time system, a control can be constructed to be optimal with respect to a corresponding cost functional [6]. Construction of a CLF though, is difficult in general. For discrete time systems Ornelas-Tellez et al. propose searching over a family of controls and associated quadratic CLFs parameterized by the symmetric positive definite (SPD) matrices [7], [8].

We propose searching over a related family of controls defined by projections with respect to the norms characterized by the SPD matrices, and give a novel relaxation of the proposed controls, which can be used to tune the response for a particular implementation.

Control design following Onelas-Tellez et al. has been implemented on a variety of nonlinear systems, with the

[1]Ethan King is with the Department of Mathematics, Center for Research in Scientific Computation, North Carolina State University, Raleigh, NC 27695, USA eaking4@ncsu.edu
[2]Hien Tran is a Professor in the Department of Mathematics and Director for the Center for Research in Scientific Computation, North Carolina State University, Raleigh, NC 27695, USA tran@ncsu.edu

necessary SPD matrix found by using the speed gradient method [8], particle swarm optimization [9], [7], the extended Kalman filter [10], and more recently an ensemble Kalman filter [11].

For the synthesis of the controls considered here, we present an ensemble Kalman search procedure to find an appropriate SPD matrix. Ensemble Kalman methods use ensemble statistics and a Kalman filter update to minimize a least squares cost functional. They are derivative free with straightforward parallelizable implementations. The ensemble Kalman filter was introduced for state estimation in noisy systems by Evensen in [12] and has since been adapted for a widening array of applications within inverse problems, including parameter estimation [13], [14] and training of neural networks [15], [16].

This paper is structured as follows. In section II, we define a control for discrete time systems which can be understood as an iterative projection with respect to a SPD matrix. In section III, we present a relaxation of the control for implementation and prove global exponential stability of the relaxed control for linear systems. In section IV and V, we introduce an ensemble Kalman search procedure for identifying controls for a given system. In section VI, we test the efficacy of the presented control design strategy on both a linear and nonlinear example.

II. P PROJECTION CONTROL

Let $f : \mathbb{R}^d \to \mathbb{R}^d$ and $g : \mathbb{R}^d \to \mathbb{R}^{d \times m}$ describe a discrete dynamical system, with control inputs $u \in \mathbb{R}^m$, where from state $x_k \in \mathbb{R}^d$, using control u_k at time step k, the next state is given by

$$x_{k+1} = f(x_k) + g(x_k)u_k. \tag{1}$$

The control objective is to drive the state of the system to a given point which, without loss of generality, is taken here to be the origin.

For a symmetric positive definite (SPD) matrix $(P \in \mathbb{R}^{d \times d}_{sym})$ and $(x \in \mathbb{R}^d), (y \in \mathbb{R}^d)$ let

$$\langle x \mid y \rangle_P \doteq x^T P y \qquad ||x||_P \doteq \sqrt{x^T P x}$$

be the P inner product and norm respectively on $\mathbb{R}^d$.

An SPD matrix P is sought such that system (1) with the control

$$u_k = \text{argmin}_{u \in \mathbb{R}^m} \frac{1}{2} ||f(x_k) + g(x_k)u||_P^2 \tag{2}$$

will converge to the origin from a given set of initial states.

Using what will be assumed here, that for all $(x \in \mathbb{R}^d)$ the columns of $g(x)$ are nonzero and linearly independent,

the control (2) is well defined for all inputs x_k, and can be written explicitly. In particular $\nabla_u[\frac{1}{2}||f(x_k)+g(x_k)u||_P^2]=0$ if and only if

$$g(x_k)^T Pg(x_k)u = -g(x_k)^T Pf(x_k).$$

Note that $g(x_k)^T Pg(x_k)$ is the Gramian matrix of the columns of $g(x_k)$ with respect to the P inner product, hence by the above assumption is positive definite, therefore invertible, so the unique feedback control at state x_k may be written as

$$u_k = -(g(x_k)^T Pg(x_k))^{-1}g(x_k)^T Pf(x_k). \quad (3)$$

The control is, at each iteration k, the projection with respect to the P norm, of the future state in the absence of control, $f(x_k)$, onto the linear subspace

$$g(x_k)^{P_\perp} \doteq \{y \in \mathbb{R}^n : g(x_k)^T Py = 0\}.$$

This can be shown by the orthogonal decomposition of $f(x_k)$ with respect to the P inner product.

Supposing that there exists a P such that control (3) is globally asymptotically stable for a given system, still the control may be unfeasible for implementation as the control magnitude may be unreasonably large when $f(x_k)$ is far from the subspace $g(x_k)^{P_\perp}$.

With SPD matrices $(P \in \mathbb{R}_{sym}^{d \times d})$ and $(E \in \mathbb{R}_{sym}^{m \times m})$ a similar control to (2) can be constructed with an explicit penalty on the magnitude of the control, given by

$$u_k = \text{argmin}_{u \in \mathbb{R}^m} \left\{ \frac{1}{2}||f(x_k)+g(x_k)u||_P^2 + \frac{1}{2}||u||_E^2 \right\},$$

which has the explicit feedback form

$$u_k = -(g(x_k)^T Pg(x_k)+E)^{-1}g(x_k)^T Pf(x_k). \quad (4)$$

Controls of this form are searched over in the framework presented by Ornelas-Tellez et al.[7], [8]. Note that it includes the optimal feedback controls for linear time invariant systems with respect to an infinite horizon Linear Quadratic Regulator (LQR) functional. That is for f and g linear, with the system dynamics for some $(A \in \mathbb{R}^{d \times d})$ and $(B \in \mathbb{R}^{d \times m})$ given by

$$x_{k+1} = Ax_k + Bu_k.$$

If there exists an SPD matrix Q for an LQR cost functional of the form

$$J(\{u_k\}_{k \in \mathbb{N}}) \doteq \sum_{k=1}^{\infty} x_k^T Qx_k + u_k^T Eu_k,$$

such that P is the solution of the corresponding Algebraic Riccati equation, then (4) is the corresponding optimal LQR feedback control [1]. Therefore it is perhaps unsurprising that an effective control of this form can be found for many systems.

Rather than directly penalizing control effort as above, in the next section we consider an implicit penalty to modulate the control.

III. RELAXED P PROJECTION CONTROL

Here we study a control with an implicit penalty on the control magnitude. For a SPD matrix P and positive scaling parameters $\{\gamma_k\}_{k \in \mathbb{N}}$, we consider the control

$$u_k = \text{argmin}_{u \in \mathbb{R}^m} \frac{1}{2}(||f(x_k)+g(x_k)u||_P^2 + \frac{1}{\gamma_k}||f(x_k)+g(x_k)u-x_k||_P^2), \quad (5)$$

which may be written more intuitively as

$$u_k = \text{argmin}_{u \in \mathbb{R}^m} \frac{1}{2}(||x_{k+1}||_P^2 + \frac{1}{\gamma_k}||x_{k+1}-x_k||_P^2).$$

Using again the assumption that for all $(x \in \mathbb{R}^d)$ the columns of $g(x)$ are nonzero and linearly independent, the control (5) is well defined and can be written in an explicit feedback form. Note that

$$\nabla_u[\frac{1}{2}||f(x_k)+g(x_k)u||_P^2 + \frac{1}{2}\frac{1}{\gamma_k}||f(x_k)+g(x_k)u-x_k||_P^2] = 0$$

if and only if

$$(1+\frac{1}{\gamma_k})g(x_k)^T Pg(x_k)u = -((1+\frac{1}{\gamma_k})g(x_k)^T Pf(x_k) - \frac{1}{\gamma_k}g(x_k)^T Px_k).$$

Therefore the control is given by

$$u_k = -(g(x_k)^T Pg(x_k))^{-1}g(x_k)^T P(f(x_k) - \frac{1}{\gamma_k+1}x_k). \quad (6)$$

The sequence $\{\gamma_k\}_{k \in \mathbb{N}}$ can be selected to modulate the control when the state is far from the equilibrium, but allow for faster convergence once the state is near the origin, in particular for some $(\gamma > 0)$ and $(\alpha > 0)$ we explore the choice

$$\gamma_k = \frac{\gamma}{||x_k||_2^\alpha}, \quad (7)$$

and prove that it produces a globally exponentially stable feedback control for linear systems.

A. Global Exponential Stability for Linear Systems

Let $(A \in \mathbb{R}^{d \times d})$ and $(B \in \mathbb{R}^{d \times m})$ with controls $u \in \mathbb{R}^m$ describe the discrete dynamics

$$x_{k+1} = Ax_k + Bu_k. \quad (8)$$

Suppose that the columns of B are linearly independent, then for a given SPD matrix P, using the feedback control (3), leads to the state update recursion

$$x_{k+1} = (I - B(B^T PB)^{-1}B^T P)Ax_k.$$

Let

$$\mathbb{P}_{B^{P_\perp}} \doteq I - B(B^T PB)^{-1}B^T P$$

denote the P projection operator onto $B^{P_\perp}$.

Definition 1. *Let $P \in \mathbb{R}^{d \times d}$ an SPD matrix, P is stabilizing for (8) if with $\{\lambda_i\}_1^d$ the eigenvalues of $(\mathbb{P}_{B^{P_\perp}}A)$, for all $i \in \{1,2\ldots,d\}$, $|\lambda_i| < 1$.*

Given a stabilizing P, we can apply the modulated control (5), with sequence $\{\gamma_k\}_{k \in \mathbb{N}}$, which will give the state update recursion

$$x_{k+1} = \mathbb{P}_{B^{P_\perp}}Ax_k + \frac{1}{1+\gamma_k}B(B^T PB)^{-1}B^T Px_k, \quad (9)$$

which will converge for only mild assumptions on the sequence $\{\gamma_k\}_{k\in\mathbb{N}}$.

Theorem 2. *Let P a stabilizing SPD matrix for (8), and parameters $(\gamma > 0), (\alpha > 0)$. The feedback control (6), with $\gamma_k = \frac{\gamma}{||x_k||_2^\alpha}$ is globally exponentially stable.*

To show Theorem 2, we first establish an explicit relation for the state at each iteration in terms of the initial value when the modulated control (6) is used, from which conditions for convergence easily follow. For the following discussion we denote the range of a matrix $B \in \mathbb{R}^{d\times m}$ as

$$R(B) \doteq \{x \in \mathbb{R}^d \mid \exists z \in \mathbb{R}^m \ s.t. \ x = Bz\}.$$

Proposition 3. *Let the initial condition be $x_0 \in \mathbb{R}^d$, and let $v_0 \in B^{P\perp}$ and $z_0 \in R(B)$ give the unique P orthogonal representation, $x_0 = v_0 + z_0$. Then using feedback control (6), for all $k \geq 1$ the following holds*

$$x_k = (\mathbb{P}_{B^{P\perp}}A)^k v_0 + \sum_{m=0}^{k-1}(\prod_{i=0}^{(k-1)-m}\frac{1}{1+\gamma_i})(\mathbb{P}_{B^{P\perp}}A)^m z_0 \,. \quad (10)$$

Proof: Employing proof by induction, suppose that

$$x_k = (\mathbb{P}_{B^{P\perp}}A)^k v_0 + \sum_{m=1}^{k-1}(\prod_{i=1}^{(k-1)-m}\frac{1}{1+\gamma_i})(\mathbb{P}_{B^{P\perp}}A)^m z_0 + \prod_{i=0}^{(k-1)}\frac{1}{1+\gamma_i}z_0 \,.$$

Note

$$(\mathbb{P}_{B^{P\perp}}A)^k v_0 + \sum_{m=1}^{k-1}(\prod_{i=1}^{(k-1)-m}\frac{1}{1+\gamma_i})(\mathbb{P}_{B^{P\perp}}A)^m z_0 \in B^{P\perp} \,,$$

therefore,

$$B(B^TPB)^{-1}B^TP\left((\mathbb{P}_{B^{P\perp}}A)^k v_0 + \sum_{m=1}^{k-1}(\prod_{i=1}^{(k-1)-m}\frac{1}{1+\gamma_i})(\mathbb{P}_{B^{P\perp}}A)^m z_0\right) = 0$$

and $(z_0 \in R(B))$, so

$$B(B^TPB)^{-1}B^TPz_0 = z_0 \,.$$

Hence, it follows

$$x_{k+1} = \mathbb{P}_{B^{P\perp}}Ax_k + \frac{1}{1+\gamma_k}B(B^TPB)^{-1}B^TPx_k$$

$$= (\mathbb{P}_{B^{P\perp}}A)^{k+1}v_0 + \sum_{m=1}^{k}(\prod_{i=1}^{(k)-m}\frac{1}{1+\gamma_i})(\mathbb{P}_{B^{P\perp}}A)^m z_0 + \prod_{i=0}^{(k)}\frac{1}{1+\gamma_i}z_0 \,.$$

Let P a stabilizing matrix for (8), and $\{\lambda_i\}_{i=1}^n$ the eigenvalues of $\mathbb{P}_{B^{P\perp}}A$ with $\lambda_{max} = max\{|\lambda_i| : i \in \{1,\ldots n\}\}$. Since P is stabilizing, there exists a σ with $\lambda_{max} < \sigma < 1$ and $M > 0$ such that for all $(y_0 \in \mathbb{R}^d)$, if $\{y_k\}_{k\in\mathbb{N}}$ given by $y_1 = \mathbb{P}_{B^{P\perp}}Ay_0$ and

$$y_{k+1} = \mathbb{P}_{B^{P\perp}}Ay_k \,,$$

then for all $k \in \mathbb{N}$

$$||y_k|| < M\sigma^k ||y_0|| \,.$$

Proposition 4. *Let $x_0 \in \mathbb{R}^d$, let $v_0 \in B^{P\perp}$ and $z_0 \in R(B)$ such that $x_0 = v_0 + z_0$, then for any sequence $(\{\gamma_k\}_{k\in\mathbb{N}}$ in $]0,\infty[)$,using feedback control (6) for all $(k \geq 0)$*

$$||x_k||_2 \leq M\left(||v_0||_2 + \frac{1}{1-\sigma}||z_0||_2\right) \,.$$

Proof: From (10), and since $(\forall k \in \mathbb{N})$, $(\frac{1}{1+\gamma_k} < 1)$ it follows that

$$||x_k||_2 \leq M(\sigma)^k ||v_0||_2 + M\sum_{m=0}^{k}(\sigma)^m ||z_0||_2$$

and $0 < \sigma < 1$, hence the result.

Proposition 5. *Let sequence $(\{\gamma_k\}_{k\in\mathbb{N}}$ in $]0,\infty[)$, if there exists $\eta \in]0,1[$ such that $(\forall k \in \mathbb{N})$, $\frac{1}{1+\gamma_k} < \eta$, then the feedback control (6) is globally exponentially stable*

Proof: Let $r = max\{\sigma,\eta\}$, then using (10),

$$||x_k||_2 \leq M\sigma^k ||v_0||_2 + M\sum_{m=0}^{k-1}\eta^{(k-1)-m}\sigma^m ||z_0|| \,,$$

therefore,

$$||x_k||_2 \leq M\sigma^k ||v_0||_2 + Mkr^{k-1}||z_0||_2 \,.$$

Since $(0 < \sigma < 1)$ and $(0 < r < 1)$, then for a $1 > \kappa > max\{\sigma,r\}$, there exists an K such that for all $(x_0 \in \mathbb{R}^d)$ and $(k \in \mathbb{N})$

$$||x_k|| \leq K||x_0||\kappa^k \,.$$

Hence the control (6) is globally exponentially stable.

The proof of Theorem 2 follows immediately from Proposition 4 and Proposition 5. In particular, by Proposition 4, with $\ell = \frac{1}{1-\sigma}$ for all $(k \in \mathbb{N})$

$$\frac{1}{1+\gamma_k} \leq \frac{(M||v_0||_2 + M||z_0||_2\ell)^\alpha}{\gamma + (M||v_0||_2 + M||z_0||_2\ell)^\alpha} \,,$$

hence by Proposition 5 the control is globally exponentially stable.

IV. IDENTIFYING A STABILIZING P

The design of a control in this framework requires first computing an SPD P for which the control (3) has good convergence properties. Then an appropriate relaxation can be used to fit the control to the desired behavior for an implementation. Searching over the whole positive definite cone for an appropriate P is challenging, in particular as the dimension of the system grows. In the case where the control input is a scalar, we explore searching for stabilizing P over a selected subset of the positive definite cone.

A. P Construction for Scalar Control Input

If control input u is a scalar then at each iteration k, $g(x_k)$ is a vector in $\mathbb{R}^d$ therefore, $g(x_k)^{P\perp}$ is a hyper-plane with normal vector $Pg(x_k)$.

Suppose that there is a hyper-plane with normal vector $v \in \mathbb{R}^d$ in the neighborhood of which the system exhibits good convergence. Then it may be effective to construct a P which restricts the dynamics near the hyper-plane, that is for all $(x \in \mathbb{R}^d)$, construct P such that $Pg(x) \approx v$. We explore computing such P by constructing a Singular Value Decomposition (SVD) with the desired property.

Let $\{e_i\}_{i=1}^d$ the standard orthonormal basis and let R_v a rotation matrix that rotates e_1 to v. For example R_v constructed with two Householder reflections as,

$$R_v = (I - 2ww^T)(I - 2e_1e_1^T)$$

where $w = \frac{(-e_1-v)}{||-e_1-v||}$. Then with

$$U_v = \begin{bmatrix} | & | & | \\ v & R_v e_2 \ldots & R_v e_d \\ | & | & | \end{bmatrix}$$

let P_v be constructed with the SVD,

$$P_v = U_v \Lambda U_v^T \qquad (11)$$

where $\Lambda = diag([s, 1 \ldots, 1])$ and s is a large scaling parameter, in computations here $s = 1 \times 10^6$.

Instead of searching over the Positive Definite Cone, we search only over the set

$$\mathscr{P} = \{P_v \; : \; v \in \mathbb{R}^d, ||v|| = 1\}$$

that is over the unit normal vectors in $\mathbb{R}^d$.

V. ENSEMBLE KALMAN SEARCH

An ensemble Kalman methodology similar to that in [11] could likely be employed to find a stabilizing P, however we use a different variation with the intent to provide a more thorough search, though at greater computational expense. First we describe a general ensemble Kalman search strategy following a similar construction to [17], [18] and then detail how we implement it for a stabilizing P search. We seek only to demonstrate that a scheme of this form can be effective for identifying stabilizing P for systems, future work would be needed to evaluate appropriate construction of an algorithm for good performance.

Let $y \in \mathbb{R}^m$ be outputs or performance measures for a system and let $\hat{y} \in \mathbb{R}^m$ a desired performance or output objective, with $\mathscr{G} : \mathscr{V} \to \mathbb{R}^m : v \to y$ the map from the input space $\mathscr{V}$ to the performance measures. The problem is to find a minimizer of the difference between the observed performance and desired performance over the input space, that is, solving

$$\min_{v \in \mathscr{V}} \frac{1}{2} ||\mathscr{G}(v) - \hat{y}||^2. \qquad (12)$$

The procedure is iterative and uses an ensemble of input points at which the performance map is evaluated, then each ensemble point is updated using summary statistics of the output map over the ensemble according to the ensemble Kalman filter rule.

For iteration n, the ensemble of size J will be denoted by $\{v_n^j\}_{j=1}^J$ and the computed performance measures for each ensemble point j of generation n will be denoted by $\{y_n^j\}_{j=1}^J$, where for all $j \in \{1, 2 \ldots, J\}$

$$y_n^j = \mathscr{G}(v_n^j).$$

For iteration n the mean ensemble estimate and mean performance measure are given by

$$\bar{v}_n = \frac{1}{J} \sum_{j=1}^J v_n^j, \qquad \bar{y}_n = \frac{1}{J} \sum_{j=1}^J y_n^j.$$

Let $\Sigma_n^{v,y}$ denote the ensemble performance cross co-variance matrix, that is for $v_n^{j,i}$ element i of ensemble point j, and $y_n^{j,i}$ performance measure i of ensemble point j in iteration n, $\Sigma_n^{v,y}$ is the $(d \times m)$ matrix with entries (ℓ, i) given by

$$(\Sigma_n^{v,y})_{\ell,i} = \frac{1}{J-1} \sum_{j=1}^J (v_n^{j,\ell} - \bar{v}_n^\ell)(y_n^{j,i} - \bar{y}_n^i).$$

Let $\Sigma_n^{y,y}$ the performance measure co-variance, the $(m \times m)$ matrix with entries (ℓ, i) given by

$$(\Sigma_n^{y,y})_{\ell,i} = \frac{1}{J-1} \sum_{j=1}^J (y_n^{j,\ell} - \bar{y}_n^\ell)(y_n^{j,i} - \bar{y}_n^i).$$

The target performance objective is used in the role of the measurements of the ensemble Kalman filter. Random perturbation of the objective has been found to be needed for ensemble Kalman type schemes in order to keep the ensemble from collapsing towards the mean value [19].

Random performance measure objectives $\{\hat{y}_n^j\}_{j=1}^J$ are generated for each ensemble point j at each iteration n, by random normal perturbation of a given target objective $\hat{y}$ with co-variance R, that is

$$\hat{y}_n^j \sim \mathscr{N}(\hat{y}, R).$$

At each iteration the ensemble points are updated by the Kalman rule, for all $j \in \{1, 2 \ldots, J\}$

$$v_{n+1}^j = v_n^j + \Sigma_n^{v,y}(\Sigma_n^{y,y} + R)^{-1}(y_n^j - \hat{y}_n^j) + w_n^j, \qquad (13)$$

where $\{w_n^j\}_{j=1}^J$ are small random multi-normal perturbations with co-variance Σ_w, that is for all $j \in \{1, 2 \ldots, J\}$

$$w_n^j \sim \mathscr{N}(0, \Sigma_w).$$

A. Implementation to compute a stabilizing P

We search for stabilizing P for systems (1) with scalar control inputs over the subset of the positive definite cone $\mathscr{P}$ using an ensemble Kalman search over the unit normal vectors in $\mathbb{R}^d$.

To evaluate the controls, we use the performance map $\mathscr{G} : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^2 : (x_0, v) \to y$, which constructs matrix P_v from v as in (11) and computes the state trajectory from initial condition x_0, for a fixed number of time steps $K \in \mathbb{N}$, under feedback control (3) with $P = P_v$, then maps the final state x_K to the two performance measures

$$y^1 = ||x_K||_2 \qquad y^2 = \frac{\langle f(x_K) - x_K \mid x_K \rangle_{P_v}}{||f(x_K) - x_K||_{P_v} ||x_K||_{P_v}} \qquad (14)$$

The criterion y^1 measures how close the control has brought the state to the origin. The angle criterion y^2 measures the component of the change in the state due to the system dynamics, $f(x_K) - x_K$, which is in the subspace $g(x_{K-1})^{P_\perp}$, and directed towards the origin with respect to the P inner product. Then if the subspace is fixed, for example the hyperplane $v^{P_\perp}$, then that is the component of the change which will be preserved by the control and move the state towards the origin. Inclusion of this term was found to improve performance.

To search for a stabilizing P, the ensemble was initialized by random normal perturbation from a starting estimate v_0, where for all $j \in \{1, \dots J\}$,

$$v_0^j = v_0 + \eta_j$$

with $\eta_j \sim \mathcal{N}^d(0, \sigma_\eta^2)$.

At each search iteration we implemented the following procedure,

---

Procedure 1 Ensemble Kalman Search Update

1: generate random initial state x_0, $x_0 \sim \mathcal{N}(0, \Sigma_{x_0})$
2: **for** $j = 1, 2 \dots J$ **do**
3:   Set $v = v_n^j$, compute P_v with (11).
4:   Compute the state trajectory from x_0 under control (3) using P_v to terminal iteration K.
5:   Compute performance measures (14) for the control
6: **end for**
7: Compute the performance objectives.
8: Update the ensemble with rule (13).
9: Normalize the ensemble elements.

Note that controls are tested from only one initial state at each iteration, while a P is likely sought which will produce good convergence over a neighborhood of the origin if not globally. Control performance for ensemble points may be tested for multiple initial states at each iteration, though at significantly increased cost, for the cases we tested, using only one initial state was sufficient.

VI. TEST EXAMPLES

A. Linear Example

Consider a system of the form (8) with

$$A = \begin{bmatrix} .9974 & .0539 \\ -.1078 & 1.1591 \end{bmatrix}, \qquad B = \begin{bmatrix} .0013 \\ .0539 \end{bmatrix}.$$

The ensemble Kalman procedure was used to search for a stabilizing P of the form (11). An ensemble of 50 vectors in $\mathbb{R}^2$ was randomly initialized according to $\mathcal{N}^2(0, 0.2)$. At each search iteration solutions under the control of each ensemble point were simulated for $K = 8$ steps from a random initial state $x_0 \sim \mathcal{N}(0, 1)$. Then the performance measures (14) were computed, with the performance objectives given by

$$\hat{y}^1 \sim \mathcal{N}(0, .0002), \qquad \hat{y}^2 \sim \mathcal{N}(-.2, .01).$$

The search procedure was iterated until satisfactory performance measures were returned, after about 150 iterations the ensemble mean was,

$$v = [-.9975, -.07]^T.$$

A reference optimal LQR feedback control was computed using

$$Q = \begin{bmatrix} 350 & 0 \\ 0 & 5 \end{bmatrix}, \qquad R = \begin{bmatrix} 3.5 \end{bmatrix}.$$

A relaxed P projection control (6) using weighting parameters of the form (7), was then chosen such that the control

had roughly the same maximal magnitude, and converged in approximately the same time as the reference LQR control. A suitable choice was found to be

$$\gamma_k = \frac{.45}{||x_k||_2^3}.$$
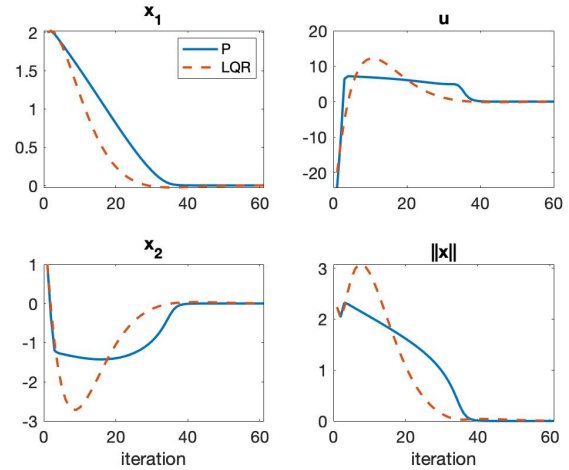
A comparison of the controls is shown in Fig. 1

Fig. 1. Comparison of LQR and relaxed P control from initial state $x_0 = [2, 1]^T$.

Note that the P control convergence rate increases as the state approaches the origin, as can be more clearly seen in Fig. 2.
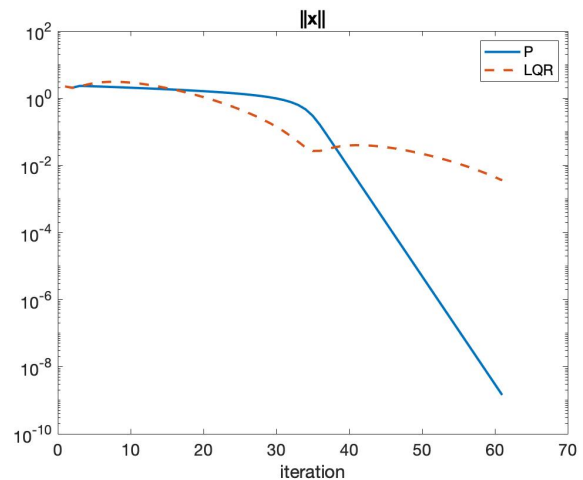

Fig. 2. Comparison of convergence of LQR and relaxed P control on Log scale.

The relaxed P projection control is slower than the LQR control to bring the system near the origin, but has a reduced overshoot, and transitions to a faster convergence rate once near the origin.

B. Nonlinear Example

Consider a system of the form (1) with

$$f(x) \doteq \begin{bmatrix} 2.2\sin(.5x_1) + .1x_2 \\ .1x_1^2 + 1.8x_2 \end{bmatrix}, \qquad g(x) \doteq \begin{bmatrix} 0 \\ 2 + .1\cos(x_2) \end{bmatrix}.$$

The Ensemble Kalman procedure was again used to search for a stabilizing P of the form (11). An ensemble of 100 vectors in $\mathbb{R}^2$ was randomly initialized according to $\mathcal{N}^2(0, 0.2)$. At each search iteration solutions under the control of each ensemble point were simulated for $K = 4$ steps from a random intial state $x_0 \sim \mathcal{N}(0,1)$. Then the performance measures (14) were computed, with the performance objectives given by

$$\hat{y}^1 \sim \mathcal{N}(0, .002), \qquad \hat{y}^2 \sim \mathcal{N}(-.2, .01).$$

The search procedure was iterated until satisfactory performance measures were returned, after about 250 iterations the ensemble mean was,

$$v = [.9950, .1003]^T.$$

A relaxed P projection control was then implemented with weighting parameters

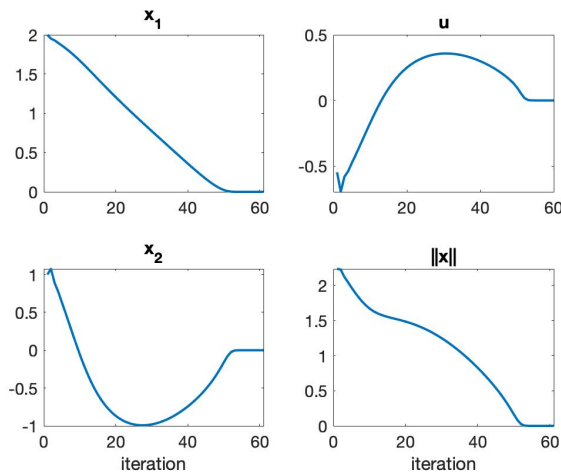$$\gamma_k = \frac{.1}{||x_k||_2^2}.$$

The result is shown in Fig. 3

Fig. 3.   Nonlinear example with relaxed P control implementation.

The parameter choice (7) for the relaxed P projection control resulted in smooth system responses for many choices of γ and α tested.

VII. CONCLUSION

The control design presented here can be effective for both linear and nonlinear systems. The advantage of the approach we present, is that an appropriate P needs to only be computed once for a system offline, then tuning of the relaxed control can be explored for a particular implementation, in contrast to the strategy presented by Ornelas-Tellez et al.

which requires supplying a quadratic cost for the control then computing an associated P, which could require iteration to fit an implementation.

While good convergence was seen in the tested examples we do not address the existence of a meaningful cost functional for which the relaxed P projection controls (5) are optimal, the structure of such a functional if it can be constructed could give good insight into how the control can be effectively employed, and is an interesting area for future work.

Analysis of the stability of the relaxed P projection control for nonlinear systems and a real time implementation for stabilization control of a double inverted pendulum is in preparation.

REFERENCES

[1] B. Anderson and J. Moore, *Optimal Control:Linear Quadratic Methods*. Englewook Cliffs, NJ: Prentice-Hall, 1990.
[2] S. Glad, "Robustness of nonlinear state feedback-a survey," *Automatica*, vol. 23, pp. 425–435, 1987.
[3] M. J. R. Sepulchre and P. V. Kokotovic, *Constructive Nonlinear Control*. Berlin: Springer-Verlag, 1997.
[4] M. Krstic and H. Deng, *Stabilization of Nonlinear Uncertain Systems*. Berlin: Springer-Verlag, 1998.
[5] S. Beeler, H. Tran, and H. Banks, "Feedback control methodologies for nonlinear systems," *Journal of Optimization Theory and Applications*, vol. 107, 02 2000.
[6] R. Freeman and P. Kokotovic, "Inverse optimality in robust stabilization," *SIAM Journal on Control and Optimization*, vol. 34, no. 4, pp. 1365–1391, 1996.
[7] F. Ornelas-Tellez, E. Sanchez, A. Loukianov, and J. Rico, "Robust inverse optimal control for discrete-time nonlinear system stabilization," *European Journal of Control*, vol. 20, 01 2013.
[8] F. Ornelas-Tellez, E. N. Sanchez, A. G. Loukianov, and E. M. Navarro-Lpez, "Speed-gradient inverse optimal control for discrete-time nonlinear systems," in *2011 50th IEEE Conference on Decision and Control and European Control Conference*, Dec 2011, pp. 290–295.
[9] R. Ruiz-Cruz, E. N. Sanchez, F. Ornelas-Tellez, A. G. Loukianov, and R. G. Harley, "Particle swarm optimization for discrete-time inverse optimal control of a doubly fed induction generator," *IEEE Transactions on Cybernetics*, vol. 43, no. 6, pp. 1698–1709, Dec 2013.
[10] M. Almobaied, I. Eksin, and M. Guzelkaya, "Inverse optimal controller based on extended kalman filter for discrete-time nonlinear systems," *Optimal Control Applications and Methods*, vol. 39, no. 1, pp. 19–34, 2018. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/oca.2331
[11] H. T. Tran and A. M. Arnold, "Ensemble kalman filtering for inverse optimal control," in *Lecture Notes in Engineering and Computer Science: Proceedings of The International MultiConference of Engineers and Computer Scientists*, Hong Kong, March 2018, pp. 526–530.
[12] G. Evensen, "Sequential data assimilation with a nonlinear quasi-geostrophic model using monte carlo methods to forecast error statistics," *Journal of Geophysical Research: Oceans*, vol. 99, no. C5, pp. 10 143–10 162, 1994.
[13] ——, "The ensemble kalman filter for combined state and parameter estimation," *IEEE Control Systems Magazine*, vol. 29, no. 3, pp. 83–104, June 2009.
[14] A. Arnold, D. Calvetti, and E. Somersalo, "Parameter estimation for stiff deterministic dynamical systems via ensemble kalman filter," *Inverse Problems*, vol. 30, no. 10, p. 105008, sep 2014.
[15] N. B. Kovachki and A. M. Stuart, "Ensemble kalman inversion: A derivative-free technique for machine learning tasks," *Inverse Problems*, 2019. [Online]. Available: http://iopscience.iop.org/10.1088/1361-6420/ab1c3a

[16] E. Haber, F. Lucka, and L. Ruthotto, "Never look back - a modified enkf method and its application to the training of neural networks without back propagation," 2018.

[17] M. A. Iglesias, K. J. H. Law, and A. M. Stuart, "Ensemble kalman methods for inverse problems," *Inverse Problems*, vol. 29, no. 4, p. 045001, mar 2013.

[18] C. Schillings and A. M. Stuart, "Analysis of the ensemble kalman filter for inverse problems," *SIAM Journal on Numerical Analysis*, vol. 55, 02 2016.

[19] G. Burgers, P. Jan van Leeuwen, and G. Evensen, "Analysis scheme in the ensemble kalman filter," *Monthly Weather Review*, vol. 126, no. 6, pp. 1719–1724, 1998.