1998-11-01

A Theory of Satisficing Decisions and Control

Richard L. Frost

Michael A. Goodrich
mike@cs.byu.edu

Wynn C. Stirling
wynn_stirling@byu.edu

Original Publication Citation

BYU ScholarsArchive Citation

A Theory of Satisficing Decisions and Control

Michael A. Goodrich, Wynn C. Stirling, and Richard L. Frost

Abstract—**The existence of an optimal control policy and the techniques for finding it are grounded fundamentally in a global perspective. These techniques can be of limited value when the global behavior of the system is difficult to characterize, as it may be when the system is nonlinear, when the input is constrained, or when only partial information is available regarding system dynamics or the environment. Satisficing control theory is an alternative approach that is compatible with the limited rationality associated with such systems. This theory is extended by the introduction of the notion of *strong satisficing* to provide a systematic procedure for the design of satisficing controls. The power of the satisficing approach is illustrated by applications to representative control problems.**

I. INTRODUCTION

A. Background and Solution Formulation

CONTROL problems are typically characterized by two desiderata that may be in tension: 1) the commitment to achieving the fundamental goal of the controller, such as tracking, regulation, or terminal control, and 2) the commitment to a performance criterion, such as minimum control effort or minimum time. Usually, these two commitments are combined into a single performance index to be minimized by applying techniques based on Bellman's principle of optimality or Pontryagin's minimum principle. This view of control is designed to obtain the best solution with respect to a given performance metric.

For many problems, however, an optimal solution is either intractable, is prohibitively expensive, or is difficult to justify because of possibly unwarranted assumptions, such as an oversimplified performance metric. Controller designs for nonlinear systems, in particular, are difficult to obtain via the optimality paradigm. In such cases, the engineering approach usually is either to find a modification of the problem such that the optimal solution to the modified problem is a satisfactory solution to the original problem, or to adopt an *ad hoc* solution. These approaches are, of course, problematic, and if they fail, the designer must seek an alternative paradigm compatible with the information available.

Optimality is not the only possible paradigm of rational choice. Economics has motivated the need for alternative paradigms that are commensurate with the available knowledge and capabilities. Simon [1] introduced the concept of a *satisficing* decision as one that, although perhaps not op-

timal, meets minimum requirements, or is "good enough." A satisficing search is a search for an optimal solution that terminates when a solution is found such that the cost of further searching exceeds the expected benefits of doing so [2]. An optimal solution is clearly satisficing, but the notion of a minimum standard may persist even if a best solution either does not exist or is not attainable. Furthermore, a satisficing solution is distinct from a "suboptimal" solution, since the latter concept presupposes the existence of an optimal solution and is usually obtained by simplifying the original problem. Since a satisficing solution must meet a minimum standard of performance, it is also different from *ad hoc* solutions, which are often based largely on vague notions of desirability under specific circumstances.

In this paper we introduce a concept of satisficing that also draws on the notion of cost/benefit tradeoffs, but in a way that is quite different from its use simply as a stopping rule for a search procedure based on an optimality paradigm. Section II develops this concept by first summarizing epistemic utility theory (a theory of cognitive decision making) and then adapting this theory to the practical context, resulting in a new theory of satisficing control. Section III presents a methodology for satisficing control for problems of the general nonlinear form $\dot{\mathbf{x}}(t) = \mathbf{f}[\mathbf{x}(t), \mathbf{u}(t), t]$, where $\mathbf{x}(\cdot)$ is the state vector of a dynamical system, $\mathbf{u}(\cdot)$ is a control input, and $\mathbf{f}(\cdot, \cdot, \cdot)$ is a dynamical model. Key features of this approach are that it 1) incorporates performance measures and design principles to characterize terminal and transition costs; 2) is amenable to a systematic design procedure; 3) is capable of meaningful operation in the presence of limited, or local, system models and information; 4) does not require restrictive modeling assumptions such as linearity or time-invariance; and 5) yields comparable performance when applied to problems for which optimal solutions are available. Examples are presented which demonstrate the design procedures and allow comparisons of computational requirements and performance results with optimal solutions. Section IV establishes conditions for satisficing control to be consistent with optimal control. Finally, Section V summarizes important results from the paper.

B. Related Literature

The synthesis approach employed herein uses a temporally local planning horizon to generate controls. Model predictive control (MPC), also known as moving horizon and receding horizon control, employs such a planning horizon for designing controllers that operate in nonlinear, constrained, and uncertain environments. MPC design requires the identification of a system model, and the specification of a system performance

metric defined over a finite planning horizon. The success of MPC is due, in large part, to its ability to handle uncertain nonlinear systems with state and input constraints such as those found in complex industrial processes. Successes in application are supported by theoretical advances, such as the characterization and specification of sufficient conditions for stability and observability [3]–[7]. Reference [6] provides a stability theorem for time-invariant nonlinear systems, but the theorem requires that the system satisfy a number of technical conditions that are difficult to verify. Reference [8] develops error bounds for receding horizon controllers for nonlinear systems, but only for time-invariant systems. If time-invariance does not hold, there seem to be few theoretical results regarding the performance of receding-horizon control techniques for nonlinear systems. Similar results apply for receding horizon H_∞ control [9], [10].

The approach presented in this paper represents the consequences of a decision by a cost-like attribute called *liability* and a benefit-like attribute called *accuracy*. Partitioning the consequence set into these attributes recalls the generalized potential field (GPF) approach to robot path planning and obstacle avoidance [11], [12]. In the GPF methodology, a goal is represented as an attractive potential, obstacles are represented as repulsive potentials, and the path along the negative gradient of the combined potentials is selected as a collision free path. With the application of harmonic potential fields [13], [14], the problem of a robot remaining in an attractive local minima is avoided, but the problems with forming a globally attractive potential field in the presence of moving obstacles remains. Additionally, although computationally efficient, GPF's do not consider the optimality of the resulting path [15]. A method proposed in [11] deals with the moving obstacle problem using a GPF formalism by incorporating the *view-time* concept. This concept appears to be a special case of a receding planning horizon as employed in MPC. By employing representations of cost and benefit in a MPC format, we obtain an efficient method for accommodating both the fundamental controller objective as well as run-time performance considerations.

Other mathematical developments [16]–[21] of the satisficing concept are motivated by the desire to make robust decisions in the presence of uncertainty. These developments compare a utility defined over the consequences of a decision to a decision threshold. This decision threshold depends only on nature and not on decision consequences. The approach presented herein is similar to these other developments in that controls are justified on the basis of a comparison, but, by contrast, our approach compares two utilities defined over the expected consequences of a decision (i.e., the decision threshold depends upon both control actions and the expected state of nature).

II. A Theory of Satisficing Decisions

A. Application of Epistemic Utility Theory to Control

Seeking the best solution to a control problem is analogous to the epistemological stance of seeking the truth regarding an inquiry. This epistemological goal is a very ambitious

one, and success cannot always be realized due to limitations of information and resources. In recognition of this fact, an alternative school of epistemological thought has emerged that professes a more modest goal than truth-seeking. The goal of this alternative school is *error avoidance*, and the methodology employed to achieve this goal is termed *epistemic utility theory* [22].

Epistemic utility theory employs two utilities, rather than one. The first utility is designed to characterize the truth support of the propositions being evaluated (subjective probability), and the second probability measure is designed to characterize the informational[1] value of *rejecting* them. These two utilities are developed independently. For example, the truth support of rival scientific theories may be assessed through their conformance with observations, and their informational value may be assessed in terms of their simplicity, explanatory power, or predictive power. The fundamental content of an epistemic utility-based approach is that, by endowing the two utilities with the mathematical structure of probabilities, they quantify the attributes of the propositions in comparable units, and may be compared. Those propositions whose truth-support does not outweigh their informational-value-of-rejection should be rejected. All other propositions should be retained as serious possibilities. This procedure retains all propositions that are considered "good," without focusing exclusively on a search for one that is deemed "best." In this way, serious errors are avoided since no credible and valuable propositions will be eliminated from consideration. This approach results, in general, in a weaker decision than does a truth-seeking approach, since the set of unrejected propositions may not be a singleton set.

We adapt epistemic utility theory to action by re-interpreting the notions of truth and information in a practical setting. We first extend the notion of "truth" to the more general concept of *accuracy*, meaning *conformity to a given standard*. In the epistemological context, the standard is truth; in the controller design context, the standard is the achievement of the fundamental goal of the controller. Consequently, control actions that have high accuracy support are likely to achieve the fundamental design goal of the controller (for example, set-point regulation). Next, we adapt the epistemic notion of information to a controls context by introducing the concept of *liability*, meaning *susceptibility or exposure to something undesirable*. An action has high liability exposure if, independently of its accuracy, it is costly in terms of resource consumption (for example, control effort), exposure to hazard, or other encumbrance.

Let U denote the set of possible control actions, and let Θ denote the state of nature (such as the system state, disturbances, and plant parameter values). Let $\mathcal{B}$ denote the Borel field in U. We wish to define probability functions to characterize the accuracy and liability of a set $G \in \mathcal{B}$ as parameterized by the state of nature. We require these utilities

---

[1] Information, as used in this context, is similar to the usage employed by Johnson-Laird: "The more possible states of affairs that a proposition eliminates from consideration, the more semantic information it contains." [23, p. 218]. This usage should not be confused with Shannon information, which is defined in terms of entropy.

to be transition probabilities. Let $P_A : \mathcal{B} \times \Theta \mapsto [0,1]$, be a transition probability such that, for fixed $\vartheta \in \Theta$, $P_A(\cdot; \vartheta)$ is a probability measure, and for fixed $G \in \mathcal{B}$, $P_A(G; \cdot)$ is a measurable function. $P_A(\cdot; \vartheta)$ characterizes the accuracy support of elements of $\mathcal{B}$, given that ϑ is the state of nature. Similarly, let $P_L : \mathcal{B} \times \Theta \mapsto [0,1]$, be a transition probability characterizing the liability exposure of elements of $\mathcal{B}$.

The *epistemic utility function* [22], [24]–[26] is the convex combination of accuracy support and liability exposure

$$\varepsilon(G, \vartheta) = \alpha P_A(G; \vartheta) + (1 - \alpha)(1 - P_L(G; \vartheta)) \quad (1)$$

where $\alpha \in [0, 1]$. After a positive linear transformation, a more convenient form for this utility function is

$$\varepsilon_b(G, \vartheta) = P_A(G; \vartheta) - b P_L(G; \vartheta) \quad (2)$$

where $b = (1 - \alpha)/\alpha$. The parameter, b, is termed the *index of rejectivity*[2], in the sense that large values of b imply an increased willingness to reject propositions. The condition $b \approx 0$ ($\alpha \approx 1$) corresponds to minimal concern for liability. The condition $b \approx 1$ ($\alpha \approx \frac{1}{2}$) corresponds to equal concern for accuracy and liability. As $b \to \infty$ ($\alpha \approx 0$), liability concerns dominate accuracy concerns, and eventually all propositions with nonzero liability exposure are rejected. For fixed ϑ, $P_L(G; \vartheta)$ and $P_A(G; \vartheta)$ are probability measures, but neither represents the subjective *belief* concerning the event G. They are best understood as a utility function and an inutility function that happen to possess the structure of transition probabilities. $P_A(G; \vartheta)$ is a measure of how much accuracy support is associated with G when ϑ is the state of nature, and $P_L(G; \vartheta)$ characterizes the liability exposure associated with G.

Let the probability P_Θ be a (possibly subjective) distribution of the states of nature that represent beliefs regarding elements of Θ. If the state of nature is known, then P_A and P_L may be evaluated. Generally, however, we will view the state of nature as a random variable. Let $\mathcal{F}$ be the Borel field in Θ and let $P_\Theta : \mathcal{F} \mapsto [0, 1]$ denote a probability measure such that $P_\Theta(W)$ represents belief that $W \in \mathcal{F}$ contains the actual state of nature. The expected value of the epistemic utility function (2) is, for $G \in \mathcal{B}$

$$\bar{\varepsilon}_b(G) = \int_\Theta [P_A(G; \vartheta) - b P_L(G; \vartheta)] P_\Theta(d\vartheta)$$
$$= \bar{P}_A(G) - b\bar{P}_L(G) \quad (3)$$

where

$$\bar{P}_A(G) = \int_\Theta P_A(G; \vartheta) P_\Theta(d\vartheta) \quad (4)$$

$$\bar{P}_L(G) = \int_\Theta P_L(G; \vartheta) P_\Theta(d\vartheta). \quad (5)$$

We now offer a definition of *satisficing* in terms of epistemic utility theory. Define the equivalence class of sets

$$\mathcal{C}_b = \left\{ S \in \mathcal{B} : S = \arg\max_{G \in \mathcal{B}} \bar{\varepsilon}_b(G) \right\} \quad (6)$$

[2] In deference to Levi's term, *boldness*, we use b to denote rejectivity.

that is, the family of all measurable sets that maximize expected epistemic utility. Let S_b be any member of this equivalence class. Since $\bar{P}_A$ and $\bar{P}_L$ each have unit mass on U, $S_b \neq \emptyset$ if the rejectivity, $b \leq 1$. Furthermore, b can always be chosen to ensure that $\bar{P}_A(S_b) > 0$. S_b is termed a *maximal satisficing set* for rejectivity b; if $G \subset S_b$, then G will be termed a *satisficing set*.

B. Characterizing the Maximally Satisficing Sets

Although the measures P_L and P_A are obtained independently, it is of interest to establish any natural connections between these two measures. We may perform the Lebesgue decomposition of $\bar{P}_L$ relative to $\bar{P}_A$ to obtain $\bar{P}_L = \bar{P}_{L_1} + \bar{P}_{L_2}$, where $\bar{P}_{L_1} \ll \bar{P}_A$ ($\bar{P}_{L_1}$ is absolutely continuous with respect to $\bar{P}_A$) and $\bar{P}_{L_2} \perp \bar{P}_A$ ($\bar{P}_{L_2}$ and $\bar{P}_A$ are mutually singular). There exists a pair of disjoint sets B_1 and B_2, with $U = B_1 \cup B_2$, such that $\bar{P}_A$ (and, hence, $\bar{P}_{L_1}$) is concentrated on B_1, and $\bar{P}_{L_2}$ is concentrated on B_2. Let S_b be a maximally satisficing set, and define $S^\| = S_b \cap B_1$ and $S^\perp = S_b \cap B_2$. Based on these definitions, we establish that $S^\perp$ and P_{L_2} can be neglected in characterizing S_b.

Lemma 1: $\bar{P}_A(S^\perp) = \bar{P}_{L_1}(S^\perp) = \bar{P}_{L_2}(S^\perp) = \bar{P}_{L_2}(S^\|) = 0$.

Proof: $\bar{P}_A(S^\perp) = 0$ by definition of absolute continuity. $\bar{P}_{L_1}(S^\perp) = \bar{P}_{L_2}(S^\|) = 0$ since $\bar{P}_{L_1}$ is concentrated on $S^\|$ and $\bar{P}_{L_2}$ is concentrated on $S^\perp$. Clearly $\bar{P}_A(S^\perp) - b\bar{P}_L(S^\perp)$ cannot be less than zero or $\bar{P}_A(S^\|) - b\bar{P}_L(S^\|) > \bar{P}_A(S_b) - b\bar{P}_L(S_b)$ which is a contradiction to (6). Then, since $\bar{P}_A(S^\perp) = \bar{P}_{L_1}(S^\perp) = 0$ and $\bar{P}_A(S^\perp) - b\bar{P}_{L_1}(S^\perp) - b\bar{P}_{L_2}(S^\perp) \geq 0$ it follows that $\bar{P}_{L_2}(S^\perp) = 0$. $\square$

Because the set $S^\perp$ has measure zero with respect to both the accuracy support and liability exposure measures, we can, without loss of generality, restrict attention to $S_b = S^\|$. Furthermore, given this restriction we can, again without loss of generality, restrict attention to $\bar{P}_L = \bar{P}_{L_1}$ on S_b.

By the Radon-Nikodym theorem, there exists a nonnegative measurable function, h, termed the *Radon-Nikodym derivative* of $\bar{P}_{L_1}$ with respect to $\bar{P}_A$, such that for any $G \in \mathcal{B}$,

$$\bar{P}_{L_1}(G) = \int_G h(u) \bar{P}_A(du).$$

The function h is also called the *likelihood ratio* between $\bar{P}_{L_1}$ and $\bar{P}_A$ on $\mathcal{B}$. We may write

$$\bar{\varepsilon}_b(G) = \bar{P}_A(G) - b\bar{P}_L(G) = \bar{P}_A(G) - b[\bar{P}_{L_1}(G) + \bar{P}_{L_2}(G)]$$
$$= \bar{P}_A(G) - b \int_G h(u) \bar{P}_A(du) - b\bar{P}_{L_2}(G). \quad (7)$$

In applying this theory to controller design, we will restrict our attention, in this paper, to single-input systems, and define the action, or control, space U as the interval $U = [u_{\min}, u_{\max}]$. We require that the accuracy support and liability exposure distribution functions assign all of their utility mass to this interval. Let $F_A(u; \vartheta) = P_A([u_{\min}, u]; \vartheta)$, and $F_L(u; \vartheta) = P_L([u_{\min}, u]; \vartheta)$ denote the corresponding distribution functions. We may also construct the expected accuracy support distribution, $\bar{F}_A(u) = \bar{P}_A([u_{\min}, u])$, and the expected liability exposure distribution, $\bar{F}_L(u) = \bar{P}_{L_1}([u_{\min}, u])$. We

will generally have $\Theta \subset \mathbb{R}^n$ for $n \geq 1$, and we let $F_\Theta(\vartheta)$ be the distribution function corresponding to the measure P_Θ. In this paper, we will restrict our attention to continuous distributions so that F_A, F_L and F_Θ are all differentiable, with density functions $f_A(u;\vartheta) = \frac{d}{du}F_A(u;\vartheta)$, $f_L(u;\vartheta) = \frac{d}{du}F_L(u;\vartheta)$ and $f_\Theta(\vartheta) = \frac{d}{d\vartheta}F_\Theta(\vartheta)$, respectively. Rewriting (4) and (5) in terms of density functions, the expected accuracy and expected liability exposure densities are

$$\bar{f}_A(u) = \int_\Theta f_A(u;\vartheta)f_\Theta(\vartheta)\,d\vartheta \tag{8}$$

$$\bar{f}_L(u) = \int_\Theta f_L(u;\vartheta)f_\Theta(\vartheta)\,d\vartheta. \tag{9}$$

In a deterministic environment, the state of nature is known, and the probability density function will concentrate all of the mass on the true value; that is, if ϑ_0 is the true state of nature, then $f_\Theta(\vartheta) = \delta(\vartheta - \vartheta_0)$, so $\bar{f}_A(u) = f_A(u;\vartheta_0)$ and $\bar{f}_L(u) = f_L(u;\vartheta_0)$. In a stochastic environment, f_Θ will characterize the uncertainty regarding the state of nature. This uncertainty may enter, for example, through process noise, sensor noise, or other random phenomena in the system. Throughout the remainder of this paper, we will treat only the deterministic problem, and assume that the true state of nature, ϑ_0, is known.

Theorem 2: If $\bar{f}_L$ and $\bar{f}_A$ exist, the Radon–Nikodym derivative $h(u) = \frac{\bar{f}_L(u)}{\bar{f}_A(u)}$ almost everywhere over the set B_1.

Proof: For all measurable $G \subseteq B_1$ we have $P_L = P_{L_1}$, and $\bar{P}_{L_1}(G) = \int_G \bar{f}_L(u)\,du$. By the Radon–Nikodym theorem, we also have $\bar{P}_{L_1}(G) = \int_G h(u)\bar{P}_A(du) = \int_G h(u)\bar{f}_A(u)\,du$. Since these expressions must hold for every measurable set $G \subseteq B_1$, we must have that $\bar{f}_L(u) = h(u)\bar{f}_A(u)$ almost everywhere. Furthermore, since P_{L_1} is absolutely continuous with respect to P_A, $\bar{f}_A(u) = 0$ implies $\bar{f}_L(u) = 0$, so $h(u) = \frac{\bar{f}_L(u)}{\bar{f}_A(u)}$ almost everywhere on B_1. □

In our subsequent development, we restrict attention to the case $\bar{P}_A(S_b) > 0$ (in controller design we require b to be small enough such that this always occurs). With this restriction, the equivalence class of maximally satisfying sets (6) becomes

$$\mathcal{C}_b = \left\{ S \in \mathcal{B} : S = \arg\max_{G \in \mathcal{B}} \left\{ \int_G [1 - bh(u)]\bar{f}_A(u)\,du \right\} \right.$$
$$\left. = \arg\max_{G \in \mathcal{B}} \left\{ \int_G [\bar{f}_A(u) - b\bar{f}_L(u)]\,du \right\} \right\} \tag{10}$$

and we may therefore view, without loss of practical significance, the following set as the maximally satisfying set

$$S_b = \{u : \bar{f}_A(u) - b\bar{f}_L(u) \geq 0\}. \tag{11}$$

The rejectivity, b, represents how prone the decision rule is to rejection. The larger the rejectivity, the smaller the set of unrejected elements of U, since $b_1 < b_2$ implies $S_{b_1} \subset S_{b_2}$. S_b will generally not be a singleton set, and there may even be a continuum of satisfying values. In contrast to utility maximizing decision-making procedures, this approach relaxes the requirement for a unique best decision. Instead, all decisions for which the ratio of accuracy support to liability exposure meets or exceeds the rejectivity value are admitted. Obviously, only one control can actually be implemented, but,

from a strictly satisfying point of view, one may choose any of the unrejected control decisions with confidence that the action will yield justifiable performance. Thus the designer has considerable latitude in the ultimate choice of the control to be implemented.

C. Strongly Satisfying Control

Satisfying, as defined herein, is a liberal notion of performance: broadly speaking, a control is satisfying if the good (characterized by accuracy support) outweighs the bad (characterized by liability exposure). Furthermore, the maximal satisfying set, S_b, will generally not be a singleton set, and there may be many satisfying possibilities. When determining a control law which is to govern a plant, a single input is required. Since each control u in the maximally satisfying set S_b is justifiable by the satisfying principle as an acceptable input to the plant, how can a single control be selected[3]?

Although all controls in S_b are satisfying, they are not necessarily all equal. In selecting a control, if a choice exists between two controls of equal liability exposure but differing accuracy support, it is reasonable to select the one with higher accuracy support. Similarly, if a choice exists between two controls of equal accuracy support, it is reasonable to select the one with lower liability exposure. For every $u \in U$ let

$$B_A(u) = \{v \in U : \bar{f}_L(v) < \bar{f}_L(u) \text{ and } \bar{f}_A(v) \geq \bar{f}_A(u)\}$$
$$B_L(u) = \{v \in U : \bar{f}_L(v) \leq \bar{f}_L(u) \text{ and } \bar{f}_A(v) > \bar{f}_A(u)\}$$

and define the set of actions that are *strictly better* than u (i.e., the set of actions that dominate u)

$$B(u) = B_A(u) \cup B_L(u) \tag{12}$$

that is, $B(u)$ consists of all possible actions that have lower liability exposure but not lower accuracy support than u, or have higher accuracy support but not higher liability exposure. If $B(u) = \emptyset$ then no actions can be preferred to u in both accuracy support and liability exposure and u is a (weakly) nondominated action.

The set

$$\mathcal{E} = \{u \in U : B(u) = \emptyset\} \tag{13}$$

is termed the *equilibrium* set. Whereas the satisfying set is determined by comparing a control's accuracy support against its liability exposure, the equilibrium set is determined by comparing controls against each other. Actions $u \in \mathcal{E}$ possesses an important equilibrium property: within the set, perturbations in u cannot increase accuracy support without also increasing liability exposure, nor can liability exposure be decreased without also decreasing accuracy support. The *strongly satisfying* set is defined as the intersection of the satisfying set and the equilibrium set

$$\mathcal{S}_b = \mathcal{E} \cap S_b. \tag{14}$$

---

[3] The necessity of answering this question is not restricted to satisfying controllers, but also arises for fuzzy logic controllers (FLC's) [27]–[29]. Loosely speaking, the solution for FLC's is to apply a defuzzification procedure that performs some kind of weighted average using the utility of each admissible control in the weighting.

Observe that the definition of the equilibrium set and the strongly satisficing set apply for U finite, countable, or uncountable.

Theorem 3: If U is closed, then S_b is closed and $\mathcal{S}_b \neq \emptyset$.

Proof: Let $\rho(u, v) = |\bar{f}_A(u) - \bar{f}_A(v)| + b|\bar{f}_L(u) - \bar{f}_L(v)|$. Then $\rho(u, v)$ defines a metric on the equivalence class of points with the same accuracy support and the same liability exposure. Let v be a point of closure of S_b. Then for all $\varepsilon > 0$ there is a $u_\varepsilon \in S_b$ such that $\rho(u_\varepsilon, v) < \varepsilon$. Now

$$
\begin{aligned}
\varepsilon &> \rho(u_\varepsilon, v) \\
&= |\bar{f}_A(u_\varepsilon) - \bar{f}_A(v)| + b|\bar{f}_L(u_\varepsilon) - \bar{f}_L(v)| \\
&\geq [\bar{f}_A(u_\varepsilon) - b\bar{f}_L(u_\varepsilon)] - [\bar{f}_A(v) - b\bar{f}_L(v)]
\end{aligned}
$$

implies that $\varepsilon + [\bar{f}_A(v) - b\bar{f}_L(v)] > [\bar{f}_A(u_\varepsilon) - b\bar{f}_L(u_\varepsilon)] \geq 0$. This inequality holds since $u_\varepsilon \in S_b$ implies $\bar{f}_A(u_\varepsilon) - b\bar{f}_L(u_\varepsilon) \geq 0$. This is true for any $\varepsilon > 0$, which means that $\bar{f}_A(v) - b\bar{f}_L(v) \geq 0$, implying that $v \in S_b$. Thus, S_b is closed.

We now show $\mathcal{S}_b$ is nonempty by constructing a specific element of this set. Define the *most discriminating* satisficing control

$$
u_D = \arg\sup_{z \in S_b} \{\bar{f}_A(z) - b\bar{f}_L(z)\}. \tag{15}
$$

Since $S_b \neq \emptyset$ (b is chosen to guarantee this), u_D exists, but may be an equivalence class if (15) does not have a unique solution. By construction, $\bar{f}_A(u_D) - b\bar{f}_L(u_D) \geq \bar{f}_A(u) - b\bar{f}_L(u) \; \forall u \in U$, or, equivalently,

$$
\bar{f}_A(u_D) - \bar{f}_A(u) \geq b[\bar{f}_L(u_D) - \bar{f}_L(u)] \quad \forall u \in U. \tag{16}
$$

For $\bar{f}_A(u) > \bar{f}_A(u_D)$, (16) implies $\bar{f}_L(u) > \bar{f}_L(u_D)$. For $\bar{f}_L(u) < \bar{f}_L(u_D)$, (16) implies $\bar{f}_A(u) < \bar{f}_A(u_D)$. Hence, there is no $u \in U$ such that both $f_L(u_D) < f_L(u)$ and $f_A(u_D) > f_A(u)$ which means that $u_D \in \mathcal{E}$. $\quad\square$

Thus $\mathcal{S}_b$ always contains at least one element, namely a most discriminating satisficing control. If $b = 0$, the most discriminating control is the *most accurate* satisficing control,

$$
u_A = \arg\sup_{z \in S_b} \{\bar{f}_A(z)\}. \tag{17}
$$

This limiting case represents a very aggressive stance to achieve the goal at the risk of excessive cost. A most accurate satisficing control may be considered for cases with large variations in $\bar{f}_A$ and small variations in $\bar{f}_L$. Another limiting case occurs as $b \to \infty$, resulting in a *least liable* satisficing control,

$$
u_L = \arg\inf_{z \in S_b} \{\bar{f}_L(z)\}. \tag{18}
$$

This procedure is very conservative, and reflects a willingness to compromise the fundamental goal in the interest of reducing cost. It may be appropriate when there are large variations in $\bar{f}_L$ relative to small variations in $\bar{f}_A$.

D. An Important Special Case

An important class of accuracy support density functions is the set of density functions that are concave over U. That is, for $\lambda \in [0, 1], \lambda \bar{f}_A(u_1) + (1 - \lambda)\bar{f}_A(u_2) \leq \bar{f}_A(\lambda u_1 + (1 - \lambda)u_2)$ for all $u_1 \in U, u_2 \in U$. Similarly, an important class

of liability exposure density functions is the set of density functions that are convex over U. That is, for $\lambda \in [0, 1]$, $\lambda \bar{f}_L(u_1) + (1 - \lambda)\bar{f}_L(u_2) \geq \bar{f}_L(\lambda u_1 + (1 - \lambda)u_2)$ for all $u_1 \in U, u_2 \in U$. For these classes, the following theorems are important in developing a synthesis procedure. Recall from (17) and (18) that u_A and u_L are the most discriminating controls for $b = 0$ and $b \to \infty$, respectively.

Lemma 4: For continuous accuracy support and continuous liability exposure, $\mathcal{E} \subset \{u \in U : \bar{f}_A^+(u)\bar{f}_L^+(u) \geq 0\} \cup \{u_A, u_L\}$ where

$$
\bar{f}_A^+(u) = \lim_{\lambda \downarrow 0} \frac{\bar{f}_A(u + \lambda) - \bar{f}_A(u)}{\lambda} \tag{19}
$$

denotes the right derivative of $\bar{f}_A$, and similarly define $\bar{f}_L^+$ as the right derivative of $\bar{f}_L$.

Proof: Since u_A and u_L represent the most discriminating controls for $b = 0$ and $b \to \infty$, respectively, by the proof of Theorem 3, they are in $\mathcal{E}$. For any other $u \in \mathcal{E}$, concavity of $\bar{f}_A$ and convexity of $\bar{f}_L$ imply that accuracy support and liability exposure are either both nonincreasing or both nondecreasing in the neighborhood of u. Hence, $\bar{f}_A^+(u)\bar{f}_L^+(u) \geq 0$. $\quad\square$

If, in addition to being continuous, both $\bar{f}_A$ and $\bar{f}_L$ are differentiable with derivatives denoted $\bar{f}_A'$ and $\bar{f}_L'$, respectively, the equilibrium set satisfies

$$
\mathcal{E} \subset \{u \in U : \bar{f}_A'(u)\bar{f}_L'(u) \geq 0\} \cup \{u_A, u_L\}. \tag{20}
$$

Moreover, for concave $\bar{f}_A$ and convex $\bar{f}_L$ the following lemma establishes a necessary and sufficient condition for determining the equilibrium set.

Lemma 5: For differentiable concave $\bar{f}_A$ and differentiable convex $\bar{f}_L$

$$
\mathcal{E} = \{u \in U : \bar{f}_A'(u)\bar{f}_L'(u) \geq 0\} \cup \{u_A, u_L\}.
$$

Proof: Let $u \in \{v \in U : \bar{f}_A'(v)\bar{f}_L'(v) \geq 0\} \cup \{u_A, u_L\}$. If $u = u_A$ or $u = u_L$ then $u \in \mathcal{E}$ since the most discriminating control is always in the equilibrium set. Otherwise, u must satisfy $\bar{f}_A'(u)\bar{f}_L'(u) \geq 0$ which implies that $\exists \lambda \geq 0$ such that $\bar{f}_A'(u) - \lambda \bar{f}_L'(u) = 0$. However, since the sum of two concave functions ($\bar{f}_A(u)$ and $-\lambda\bar{f}_L(u)$) is also concave then the u which satisfies $\bar{f}_A'(u) - \lambda\bar{f}_L'(u) = 0$ is a most discriminating control for $b = \lambda$. Since a most discriminating control cannot be dominated, $u \in \mathcal{E}$. Thus,

$$
\mathcal{E} \supset \{u \in U : \bar{f}_A'(u)\bar{f}_L'(u) \geq 0\} \cup \{u_A, u_L\}.
$$

This result, coupled with Lemma 4 establishes the desired result. $\quad\square$

Theorem 6: Let $U \subset \mathbb{R}$. For $\bar{f}_A$ a concave density function and $\bar{f}_L$ a convex density function over U the maximal satisficing set S_b is convex. Moreover, for concave differentiable accuracy support and convex differentiable liability exposure, the equilibrium set $\mathcal{E}$ and the strongly satisficing set $\mathcal{S}_b$ are convex.

Proof: Convexity of the maximal satisficing set is shown by establishing that, for $u_1 \in S_b$ and $u_2 \in S_b$, the point $v_\lambda = \lambda u_1 + (1 - \lambda)u_2 \in S_b$ for any $0 \leq \lambda \leq 1$. By concavity of $\bar{f}_A$ and convexity of $\bar{f}_L$, we get that

$$\bar{f}_A[\lambda u_1 + (1 - \lambda)u_2] \geq \lambda \bar{f}_A(u_1) + (1 - \lambda)\bar{f}_A(u_2)$$
$$-\bar{f}_L[\lambda u_1 + (1 - \lambda)u_2] \geq -\lambda \bar{f}_L(u_1) - (1 - \lambda)\bar{f}_L(u_2),$$

hence

$$\bar{f}_A[\lambda u_1 + (1 - \lambda)u_2] - b\bar{f}_L[\lambda u_1 + (1 - \lambda)u_2]$$
$$\geq \lambda[\bar{f}_A(u_1) - b\bar{f}_L(u_1)] + (1 - \lambda)[\bar{f}_A(u_2) - b\bar{f}_L(u_2)].$$

Now, since $u_1, u_2 \in S_b$ we know that $\bar{f}_A(u_1) - b\bar{f}_L(u_1) > 0$ and $\bar{f}_A(u_2) - b\bar{f}_L(u_2) > 0$, whence

$$\bar{f}_A(v_\lambda) - b\bar{f}_L(v_\lambda) > 0.$$

Thus S_b is convex.

To establish the convexity of $\mathcal{E}$, first note that by Lemma 4

$$\mathcal{E} \subset \{u \in U : \bar{f}'_A(u)\bar{f}'_L(u) \geq 0\} \cup \{u_A, u_L\}. \quad (21)$$

If $u_A = u_L$ then $\mathcal{E} = \{u_A\}$. Otherwise, let $u_1, u_2 \in \mathcal{E}$ with $u_1 < u_2$, and suppose $\bar{f}'_A(u_1)$ and $\bar{f}'_A(u_2)$ are of opposite sign. Since $\bar{f}_A$ is concave, this requires that $\bar{f}'_A(u_1) > 0$ and $\bar{f}'_A(u_2) < 0$. For u_1 and u_2 to be elements of $\mathcal{E}$, requires that $\bar{f}'_L(u_1) > 0$ and $\bar{f}'_L(u_2) < 0$, which is impossible since $\bar{f}_L$ is convex. Thus the directional derivatives of $\bar{f}_A$ and $\bar{f}_L$ cannot change sign in $\mathcal{E}$. Since $\bar{f}_A$ is concave, there can be at most one sign change in the derivative, so it can be concluded that $\bar{f}'_A(v_\lambda)$ has the same sign as $\bar{f}'_A(u_1)$ and $\bar{f}'_A(u_2)$. A similar argument holds for $\bar{f}'_L(v_\lambda)$, and consequently $\bar{f}'_A(v_\lambda)\bar{f}'_L(v_\lambda) \geq 0$ so, by Lemma 5, $v_\lambda \in \mathcal{E}$. Finally, since the intersection of convex sets is convex, $\mathcal{S}_b$ is convex. $\square$

This theorem means that, for concave accuracy support and convex liability exposure defined on an interval $U = [u_{\min}, u_{\max}]$, the maximally satisficing set is also an interval. Moreover, $\mathcal{E} = [\min\{u_A, u_L\}, \max\{u_A, u_L\}]$. In the next section, it will be shown that this is a useful characteristic.

The following theorem establishes an equivalence between the equilibrium set and the set of most discriminating controls. This theorem is useful because it says all elements of $\mathcal{E}$ (not just the endpoints, u_A and u_L, are maximizing elements, whence $\mathcal{S}_b$ contains only satisficing and maximizing elements.

Theorem 7: Let $U \in \mathbb{R}$, let $\bar{f}_A$ be concave density function, differentiable over the interior of U, and let $\bar{f}_L$ be a convex density function, differentiable over the interior of U. Then for every $u \in \mathcal{E}$ there exists a rejectivity value $b \in [0, \infty)$ such that u is a most discriminating satisficing control. Furthermore, if $\bar{f}_A - b\bar{f}_L$ is strictly concave for every $b \in [0, \infty)$, then for every such b there corresponds a unique most discriminating satisficing control.

Proof: Let $u \in \mathcal{E}$. Define $u_* = \min\{u_A, u_L\}$ and $u^* = \max\{u_A, u_L\}$. For these limiting cases, it has previously been established that by assigning rejectivity values of $b \to \infty$ and $b = 0$, respectively, u_L and u_A are most discriminating satisficing controls. Fix $u \in (u_*, u^*)$. The function $\bar{f}_A(v) - b\bar{f}_L(v)$ is extremized when $\bar{f}'_A(v) - b\bar{f}'_L(v) = 0$. If $\bar{f}'_L = 0$ then $\bar{f}'_A = 0$ whence, by concavity of $\bar{f}_A$ and convexity of $\bar{f}_L$, $v = u_A = u_L$ for any b. Otherwise, when $\bar{f}'_L \neq 0$,
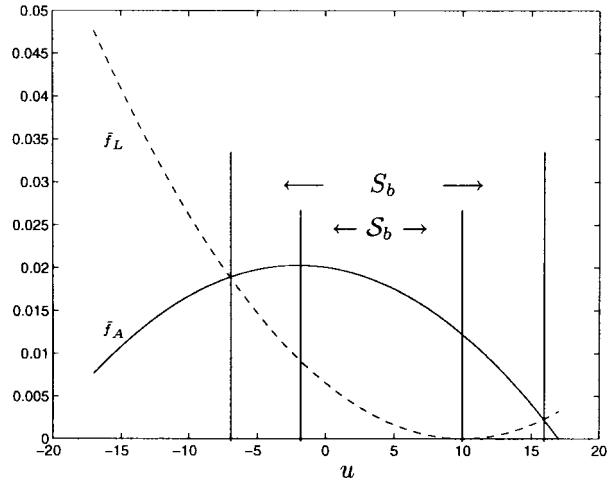

Fig. 1. Satisficing regions for a concave $\bar{f}_A$ and convex $\bar{f}_L$ plotted as densities.

evaluating $\bar{f}'_A(v) - b\bar{f}'_L(v) = 0$ at $v = u$ implies $b_u = \frac{\bar{f}'_A(u)}{\bar{f}'_L(u)}$ is the rejectivity required to render u a most discriminating satisficing control. Furthermore when $\bar{f}_A - b\bar{f}_L$ is strictly concave for all $b \in [0, \infty)$, $u = \arg\max_{v \in (u_*, u^*)}\{\bar{f}_A(v) - b\bar{f}_L(v)\}$ is the unique most discriminating satisficing control. $\square$

Fig. 1 illustrates the maximal satisficing and strongly satisficing sets for a concave differentiable $\bar{f}_L$ and convex differentiable $\bar{f}_A$, with the accuracy support and liability exposure plotted as functions of u. The satisficing and strongly satisficing sets are shown on the plot for $b = 1$. In the figure, S_b consists of those u for which $\bar{f}_A(u)$ exceeds $\bar{f}_L(u)$. The set $\mathcal{S}_b$ consists of those u for which no control with higher accuracy support exists for a given liability exposure level, and for which no control with lower liability exposure exists for a given accuracy support level. Observe that for $u \in \mathcal{S}_b$ the accuracy support and liability exposure have slopes with the same sign (see Lemma 5), whence controls in this region are in equilibrium.

III. APPLICATIONS OF SATISFICING CONTROL

Accuracy support and liability exposure are mechanisms to implement the goals and design ideals of the problem. If information is available over the full extent of the problem, then the functions may be designed from a global perspective. If a global solution is not available or is not implementable, we may still incorporate whatever information is available to design the accuracy support and liability exposure functions via the receding horizon concept. By permitting the designer to tailor the structure of these functions according to what is actually known or defensibly assumed, the problem may be cast in its natural setting. This capability frees the designer from the need to make arbitrary assumptions simply to invoke a global solution technique.

In this section, we first develop a synthesis procedure for satisficing control using receding control horizons. We then look at two problem classes: quadratic regulation of linear and nonlinear systems, and minimum time problems. The linear

quadratic regulator provides a benchmark for performance and computational comparisons; the inverted pendulum (nonlinear quadratic regulator) provides a solution for a nonlinear plant with nonstationary plant parameters using a computationally feasible solution method, and Zermelo's problem provides a benchmark for performance and computational comparisons of minimum time problems.

A. Controller Formulation

1) Receding Planning Horizons and Influence Vectors: Consider the nonlinear, time-varying, single-input, n-dimensional system

$$\mathbf{x}(t+1) = \mathbf{f}[\mathbf{x}(t), u(t), t], \quad t = 0, 1, \ldots, t_f - 1 \quad (22)$$

together with the performance index

$$J = \phi[\mathbf{x}(t_f)] + \sum_{t=0}^{t_f-1} L[\mathbf{x}(t), u(t), t] \quad (23)$$

where ϕ and L are given performance indexes (for example, quadratic forms), and where the terminal time, t_f, is unspecified. General global solutions for this nonlinear control problem are not easily obtainable, but many such problems may be addressed from a local perspective. One natural approach is to invoke a receding, or rolling, horizon control strategy [6], [8]. This approach consists of implementing a feedback controller through a series of repeated open-loop calculations based on the instantaneous state. For a discrete-time receding horizon of length d, the next d values, $\{u(t), \ldots, u(t + d - 1)\}$, are computed as functions of the current state, $\mathbf{x}(t)$. The control $u(t)$ is implemented, producing a state $\mathbf{x}(t + 1)$, the horizon is shifted forward one time unit, and the process is repeated. For example, a one-step control horizon ($d = 1$) would require the design of only $u(t)$, the control for the current time increment. For $d = 2$, both $u(t)$ and $u(t + 1)$ are required.

In this paper, our approach is to employ a receding horizon and to specify f_A and f_L as functions of the control input, u, and the state of nature, ϑ, such that f_A assigns high accuracy support to those control values that tend to achieve the goal of the system, and f_L assigns low liability exposure to those control values that tend to conform to the design principles.

Because the system model (22) may involve delays (we will assume it is causal), it is possible that $u(t + \ell)$ will have an explicit influence on elements of the state vector beyond the $t + \ell + 1$st step. Consequently, we may define the *explicit influence horizon* of $u(t + \ell)$, denoted $D_{t+\ell}$, as the maximum number of time increments for which $u(t + \ell)$ has explicit influence on any component of the state.

If $D_{t+\ell} > 1$, let $\xi_\ell(t + \ell + k)$ be the sub-state of $\mathbf{x}(t + \ell + k)$ that is an explicit function of $u(t + \ell)$. Continuing in this manner, we may generate d sequences of substates, denoted

the *influence vectors*, of the form

$$\boldsymbol{\xi}_0(u(t)) = \{\xi_0(t+1), \ldots, \xi_0(t + D_t)\}$$
$$\boldsymbol{\xi}_1(u(t+1)) = \{\xi_1(t+2), \ldots, \xi_1(t + D_{t+1})\}$$
$$\vdots \quad (24)$$
$$\boldsymbol{\xi}_{d-1}(u(t+d-1)) = \{\xi_{d-1}(t+d+1), \ldots,$$
$$\xi_{d-1}(t + d + D_{t+d})\}.$$

The collection $\{\boldsymbol{\xi}_\ell(u(t+\ell)), \ 0 \le \ell \le d - 1\}$ then represents the state elements that are explicit functions of $u(t), \ldots, u(t + d - 1)$, and using these elements we may formulate *local performance indexes*, $\Phi[\boldsymbol{\xi}_0(u(t)), \ldots, \boldsymbol{\xi}_{d-1}(u(t+d-1))]$ and $\Lambda[\boldsymbol{\xi}_0(u(t)), \ldots, \boldsymbol{\xi}_{d-1}(u(t+d-1))]$ to compute f_A and f_L. As demonstrated in the subsequent examples, these local performance indexes are formed according to the same principles that are used to define the global performance indexes, ϕ and L.

2) Synthesis for Restricted Plant Structure: Given this receding horizon framework and the notion of an influence vector, we now restrict attention to a set of fixed conditions and design a satisficing controller for the resulting structure. The use of an influence vector allows us to account for the relative order of the system (see, for example [4]) while using Euler integration. We can thus use this simple discretization method (the first restriction) producing the discrete-time dynamical expression given by

$$\mathbf{x}(t+1) = \mathbf{x}(t) + T\dot{\mathbf{x}}(t) \quad (25)$$

where T is the sampling time. The second restriction is that we consider single input nonlinear systems of the form

$$\dot{\mathbf{x}}(t) = \mathbf{f}[\mathbf{x}(t), t] + D[\mathbf{g}[\mathbf{x}(t), t]]\mathbf{h}[u(t), t] \quad (26)$$

where each $\mathbf{f}, \mathbf{g}, \mathbf{h}$ are bounded, and where $\mathbf{h}$ is differentiable with respect to u for all time (this allows us to minimize the local cost functions presented in the next section); and where D is the diagonal operator which places all of the elements of the $n \times 1$ vector $\mathbf{y}$ into the diagonal elements of the $n \times n$ matrix $D[\mathbf{y}]$.

We now construct the influence vectors for this restricted class of systems. Let the subscript i denote the ith element of the vector; for example, x_i is ith element of $\mathbf{x} = [x_1, \ldots, x_n]^T$ and $f_i(\mathbf{x}, t)$ is the ith element of the vector $\mathbf{f}(\mathbf{x}, t) = [f_1(\mathbf{x}, t), \ldots, f_n(\mathbf{x}, t)]^T$. The discretized dynamics for x_i are given, via (25)–(26), by

$$x_i(t+1) = f_i(\mathbf{x}(t), t) + g_i(\mathbf{x}(t), t)h_i(u(t), t).$$

The following procedure then determines the minimum delay δ_i required before $x_i(t+\delta_i)$ is influenced by $u(t)$. At any step, let $\mathcal{I}$ be the set of indexes for which δ_i is known.

1) Set $i = j = 1$.
2) For each i, if $g_i(u(t)) \neq 0$ set $\delta_i = j$. (These elements are immediately influenced by $u(t)$.)
3) If all δ_i are known quit. Else $j = j + 1$.
4) Identify all $f_i[\mathbf{x}(t+j-1), t+j-1]$, $i \notin \mathcal{I}$ that depend upon any $x_k \in \mathbf{x}(t+j-1)$, $k \in \mathcal{I}$. For each of these k set $\delta_i = j$. (These elements are not immediately (at time

$t+j-1$) influenced by $u(t)$, but instead are influenced at time $t+j$ by $\mathbf{x}(t+j-1)$ via $f_i[x_k(t+j-1), t+j-1]$.)

5) Repeat step 3.

This procedure first identifies elements $x_i(t+1)$ affected directly by $u(t)$. It then identifies functions f_j that are affected by $x_i(t+1)$. These functions will directly affect $x_j(t+2)$ which will then affect new f_j, and so on. Any δ_i not identified in $n-1$ steps (given that $\mathbf{x}$ is $n \times 1$) are not affected by $u(t)$ (i.e., are not controllable at time t). The influence vector is then identified as

$$\boldsymbol{\xi}_0(u(t)) = [x_1(t+\delta_1), \ldots, x_n(t+\delta_n)]^T$$

and the influence horizon $D_t = \max_{i=1,\ldots n} \delta_i$.

Identification of the influence vectors $\boldsymbol{\xi}_j$ for $j = 1, \ldots, d-1$ then proceeds analogously. However, for purposes of identifying useful local cost functions, we introduce a third restriction. Although we allow the system matrices (such as $\mathbf{f}(\cdot, t)$) to vary with time, we require that the influence horizon D_t and each δ_i are constant. We refer to such a system as a *restricted time varying* system, and this restriction allows us to use local cost functions that employ the same influence vector structure. Thus, $\boldsymbol{\xi}_k(u(t+k)) = [x_1(t+k+\delta_1), \ldots, x_n(t+k+\delta_n)]^T$ for all $k = 0, \ldots, d-1$.

3) Constructing $\bar{f}_A$ and $\bar{f}_L$ Using Local Planning Horizons: From a local perspective, we view $t + d$ as the terminal time, t_f. Let the state of nature include the current state of the system, that is, $\vartheta = \mathbf{x}(t)$. Because we associate f_A with the goal, $f_A[\mathbf{u}; \mathbf{x}(t)]$ is designed to have large values for $\mathbf{u} = [u_0, \ldots, u_{d-1}]^T$ such that $\Phi[\boldsymbol{\xi}_0(u_0), \ldots, \boldsymbol{\xi}_{d-1}(u_{d-1})]$ is small. We obtain f_A by negating and normalizing this function. Define

$$g_A[\mathbf{u}; \mathbf{x}(t)] = \sup_{\mathbf{z} \in U^d} \{\Phi[\boldsymbol{\xi}_0(z_0), \ldots, \boldsymbol{\xi}_{d-1}(z_{d-1})]\} - \Phi[\boldsymbol{\xi}_0(u_0), \ldots, \boldsymbol{\xi}_{d-1}(u_{d-1})] + \epsilon \quad (27)$$

where $\mathbf{z} = [z_0, \ldots, z_{d-1}]^T$, U^d is the d-dimensional control space $U \times \cdots \times U$, and ϵ is a small number inserted to insure that all of the values of g_A are nonzero. The probability density function f_A is obtained from g_A by normalizing

$$f_A[\mathbf{u}; \mathbf{x}(t)] = \frac{g_A[\mathbf{u}; \mathbf{x}(t)]}{\int_{U^d} g_A[\mathbf{z}; \mathbf{x}(t)] \, d\mathbf{z}}. \quad (28)$$

f_A takes its maximum at the values $\mathbf{u}$ that drive Φ closest to zero, but also assigns significant accuracy support to control values in the neighborhood of this control value.

The liability exposure density function f_L is a measure of how well the control decision $\mathbf{u}$ complies with the incremental cost functional (cost-to-go), $\Lambda[\boldsymbol{\xi}_0(u(t)), \ldots, \boldsymbol{\xi}_{d-1}(u(t+d-1))]$, from the local perspective. Consequently we define

$$g_L[\mathbf{u}; \mathbf{x}(t)] = \Lambda[\boldsymbol{\xi}_0(u_0), \ldots, \boldsymbol{\xi}_{d-1}(u_{d-1})] - \inf_{\mathbf{z} \in U} \{\Lambda[\boldsymbol{\xi}_0(z_0), \ldots, \boldsymbol{\xi}_{d-1}(z_{d-1})]\} + \epsilon. \quad (29)$$

Converting this into a probability density function by normalizing yields

$$f_L[\mathbf{u}; \mathbf{x}(t)] = \frac{g_L[\mathbf{u}; \mathbf{x}(t)]}{\int_{U^d} g_L[\mathbf{z}; \mathbf{x}(t)] \, d\mathbf{z}}. \quad (30)$$

This function represents the local cost of the decision to take action $\mathbf{u}$. Actions that result in large values of f_L will have higher liability than actions that result in small values of f_L. These valuations are made independently of the likelihood that a given choice will achieve the desired goal—the only consideration is liability exposure, which we equate with cost.

4) Independence of Accuracy and Liability: These local cost functions associate controls with high accuracy support (high f_A) with small terminal cost, and controls with high liability exposure (high f_L) with large incremental cost. When accuracy support exceeds liability exposure, the corresponding control produces a terminal state of sufficient value that the incremental effort required to reach the terminal state is justified. This explicit tradeoff between tolerance for goal achievement (such as maximum position error within a certain tolerance) and effort (such as energy consumption) is of fundamental importance in physical controller design.

From a globally optimal perspective (long planning horizon) this tradeoff in the cost function can be resolved by analyzing the global performance of the resulting controller. However, from a local perspective with unknown time-varying conditions the optimal resolution between the need to, for example, regulate the system without expending excessive fuel, cannot be computed *a priori*. Instead, by independently allocating a unit of accuracy support over all possible controls and a unit of liability exposure over the same set, the benefit/cost tradeoff can be justifiably performed without resorting to analysis of global performance. In other words, the benefit and costs of control can be independently assessed and the tradeoff between these elements can be performed using local information.

B. Linear System with Quadratic Performance

Since the solution to the optimal linear quadratic regulator is well known, it provides a convenient benchmark against which to evaluate satisficing solutions. Consider the following time-varying single-input system:

$$\mathbf{x}(t+1) = A(t)\mathbf{x}(t) + B(t)u(t), \quad t = 0, 1, \ldots, t_f - 1 \quad (31)$$

where $\mathbf{x}(t)$ is an n-dimensional state vector, $A(t)$ is an $n \times n$ matrix, $u(t)$ is a scalar input, and $B(t)$ is an n-dimensional vector. We wish to choose the control according to a quadratic performance index, so (23) becomes

$$J = \mathbf{x}^T(t_f)[P - Q(t_f)]\mathbf{x}(t_f) + \sum_{t=0}^{t_f-1} [\mathbf{x}^T(t+1)Q(t+1)\mathbf{x}(t+1) + R_u(t)u^2(t)] \quad (32)$$

where $P \geq 0$, $Q(t) \geq 0$ and $R_u(t) \geq 0$.

To demonstrate the application of satisficing control, we develop an epistemic utility-based approach to a temporally local version of this problem. Locality is invoked via a receding horizon control strategy. The most restrictive such version is a one-step control horizon, that is, $d = 1$. It is important to emphasize that, because the satisficing controller uses only temporally local information, it is insensitive to a time-invariance assumption. Thus, the matrices A, B, Q, and R may be time-varying without appreciable change to

the procedure. Furthermore, these matrices need not even be known in advance of the instantaneous horizon, $t + d$. In this latter case, it would not be possible to obtain a globally optimal solution.[4]

5) Derivation of Accuracy Support and Liability Exposure:
Let $U_m = (-u_m, u_m)$ be a finite open interval representing the range of admissible controls. We begin by specifying g_A and g_L according to (27) and (29). To calculate g_A, we observe that "regulation" occurs when all terminal substates in $\mathbf{x}$ are zero. Thus, we penalize all terminal states equally whereby we set $P - Q(t_f) = I$. For the problem addressed in this section, the explicit influence horizon is $D_t = 1$. The influence vector is $\boldsymbol{\xi}_0(u) = \mathbf{x}(t+1) = A(t)\mathbf{x}(t) + B(t)u$, where we have dropped the time dependence on u to emphasize that at each time step a new control is generated. Accuracy support is based on the local performance index

$$\Phi[\boldsymbol{\xi}_0(u)] = \mathbf{x}^T(t+1)[P - Q(t_f)]\mathbf{x}(t+1)$$
$$= [A(t)\mathbf{x}(t) + B(t)u]^T[A(t)\mathbf{x}(t) + B(t)u]. \quad (33)$$

Similarly, liability exposure is based on the local performance index

$$\Lambda[\boldsymbol{\xi}_0(u)] = \mathbf{x}^T(t+1)Q(t+1)\mathbf{x}(t+1) + R_u(t)u^2$$
$$= [A(t)\mathbf{x}(t) + B(t)u]^T Q(t+1)[A(t)\mathbf{x}(t)$$
$$+ B(t)u] + R_u(t)u^2. \quad (34)$$

Because Φ and Λ are quadratic in u, it is possible to identify the minimum of these local functions

$$u_{\mathcal{E}_A} = \arg \min_{u \in U_m} \Phi[\boldsymbol{\xi}_0(u)] \quad (35)$$

$$u_{\mathcal{E}_L} = \arg \min_{u \in U_m} \Lambda[\boldsymbol{\xi}_0(u)]. \quad (36)$$

From an implementation perspective, it is helpful to allocate all of the accuracy support mass and all of the liability exposure mass to the equilibrium set. It is easily shown that for quadratic performance indexes the boundaries of the equilibrium set are determined by the minimum values of Φ and Λ. Let

$$u_* = \min\{u_{\mathcal{E}_A}, u_{\mathcal{E}_L}\} \quad (37)$$

$$u^* = \max\{u_{\mathcal{E}_A}, u_{\mathcal{E}_L}\} \quad (38)$$

$$U = [u_*, u^*] = \mathcal{E} \quad (39)$$

and assume $u_m \gg |u_*|$ and $u_m \gg |u^*|$ so that the boundaries of U_m do not affect these values.

Since $\Lambda[\boldsymbol{\xi}_0(u)]$ is quadratic, it assumes its unique minimum at $u_{\mathcal{E}_L}$. Thus, (29) becomes

$$g_L[u; \mathbf{x}(t)] = \Lambda[\boldsymbol{\xi}_0(u)] - \Lambda[\boldsymbol{\xi}_0(u_{\mathcal{E}_L})] + \epsilon \quad (40)$$

with normalizing term

$$G_L[\mathbf{x}(t)] = \int_{u_*}^{u^*} g_L(w; \mathbf{x}(t)) \, dw. \quad (41)$$

Both (40) and (41) can be easily obtained in closed form using (34). Similarly, since $\Phi[\boldsymbol{\xi}_0(u)]$ is quadratic in u it achieves its maximum at $u_{\mathcal{E}_L}$ when restricted to $\mathcal{E}$. Hence, (27) becomes,

$$g_A[u; \mathbf{x}(t)] = \Phi[\boldsymbol{\xi}_0(u_{\mathcal{E}_L})] - \Phi[\boldsymbol{\xi}_0(u)] + \epsilon \quad (42)$$

[4]We postpone such a demonstration to the inverted pendulum example.

with normalizing term

$$G_A[\mathbf{x}(t)] = \int_{u_*}^{u^*} g_A(w; \mathbf{x}(t)) \, dw. \quad (43)$$

Again, both (42) and (43) can be easily obtained in closed form using (33).

The most accurate and least liable controls are given by

$$u_A = u_{\mathcal{E}_A} = -[B^T(t)B(t)]^{-1}B^T(t)A(t)\mathbf{x}(t) \quad (44)$$
$$u_L = u_{\mathcal{E}_L} = [R_u(t) + B^T(t)Q(t+1)B(t)]^{-1}$$
$$\times B^T(t)Q(t+1)A(t)\mathbf{x}(t). \quad (45)$$

In this development, all belief is placed on $\mathbf{x}$, whence $\bar{f}(u) = f(u; \mathbf{x})$. Because $\bar{f}_A$ is concave and $\bar{f}_L$ is convex, $\mathcal{S}_b$ is convex (see Theorem 6), which implies that all possible strongly satisficing controls may be obtained via convex combinations of u_L and u_A. For $\lambda \in [0, 1]$ define $u_\lambda = \lambda u_L + (1 - \lambda)u_A$. For $\lambda \approx 0$, the control tends to reduce the accumulated cost at the expense of large terminal error, and for $\lambda \approx 1$, the control tends to reduce the terminal error at the expense of the accumulated cost. Thus, λ is a design parameter for a synthesizing procedure. There exists a λ_D such that the most discriminating control is given by $u_D = u_{\lambda_D}$. The most discriminating control u_D can be calculated directly in a manner similar to the calculations of u_A and u_L (i.e., minimizing $\bar{f}_A(u) - b\bar{f}_L(u)$ with respect to u), yielding

$$u_D = -[B^T(t)(I + b'Q(t+1))B(t) + b'R_u(t)]^{-1}$$
$$\times B^T(t)(I + b'Q(t+1))A(t)\mathbf{x}(t) \quad (46)$$

where $b' = b\frac{G_A[\mathbf{x}(t)]}{G_L[\mathbf{x}(t)]}$. Let $P_D(t+1) = I + b'Q(t+1)$ and define the satisficing gain

$$K_D(t) = [B^T(t)P_D(t+1)B(t) + b'R_u(t)]^{-1}$$
$$\times B^T(t)P_D(t+1)A(t) \quad (47)$$

so that $u_D = -K_D(t)\mathbf{x}(t)$. The satisficing control for the linear quadratic regulator is a state-feedback control, and has a structure similar to the optimal feedback control. Because the gain K_D is a function of the state $\mathbf{x}(t)$ (via b'), however, the feedback is not linear.

6) Simulation Results and Comparison to Optimal Control:
When $A(t)$, $B(t)$, $Q(t)$ and $R_u(t)$ are defined for all time and known beforehand, the optimal solution can be obtained of the form $u(t) = -K(t)\mathbf{x}(t)$, where $K(t)$ is the Kalman gain (see [30, Theorem 6.28])

$$K(t) = \{B^T(t)[Q(t+1) + P(t+1)]B(t) + R_u(t)\}^{-1}$$
$$\times B^T(t)[Q(t+1) + P(t+1)]A(t) \quad (48)$$

and where $P(t) = A^T(t)[Q(t+1) + P(t+1)][A(t) - B(t)K(t)]$, with terminal condition $P(t_f) = P - Q(t_f)$. In the special case where $A(t)$, $B(t)$, $Q(t)$, and $R_u(t)$ are constant, the system is time-invariant and a steady-state solution exists of the form $u(t) = K_\infty \mathbf{x}(t)$, where the steady-state Kalman gain $K_\infty = (B^T S B + R)^{-1} B^T S A$, and S is obtained via the algebraic Riccati equation [31, Eqs. (2.4)–(12)].
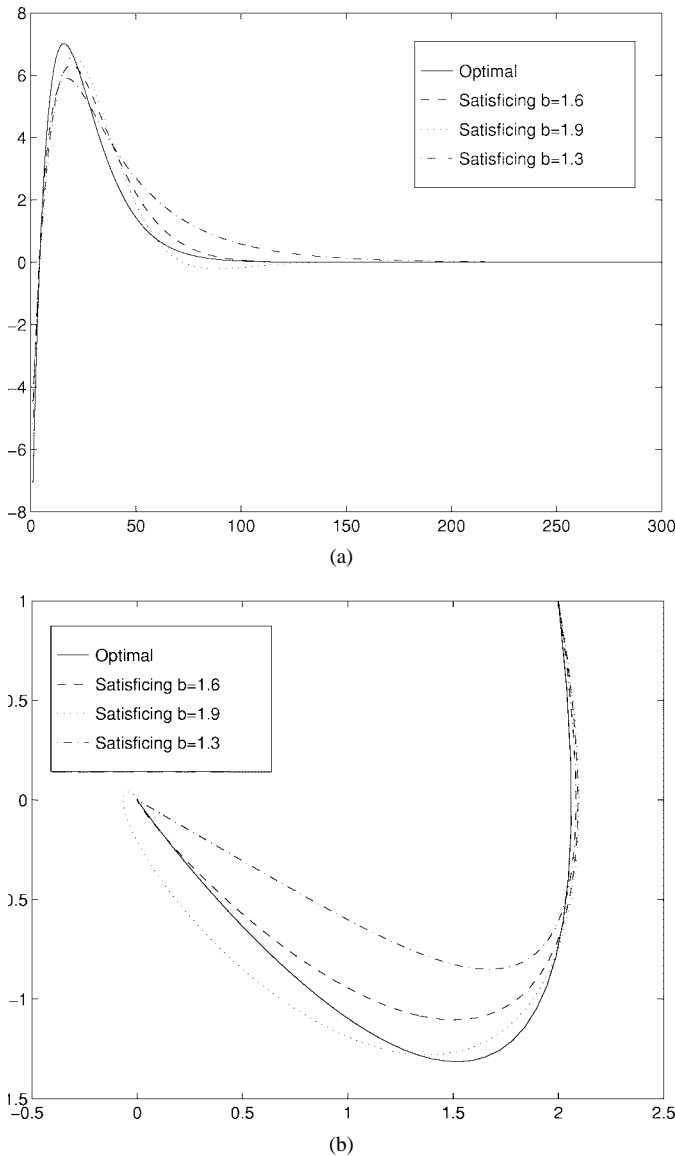
Fig. 2.   Performance results for the single-agent linear regulator problem: (a) control history and (b) phase plane.

To compare the optimal and satisficing control policies, we use an unstable second-order linear time-invariant system example taken from [32]. Let

$$A = \begin{bmatrix} 0.9974 & 0.0539 \\ -0.1078 & 1.1591 \end{bmatrix} \quad B = \begin{bmatrix} 0.0013 \\ 0.0539 \end{bmatrix}$$

$$Q = \begin{bmatrix} 0.25 & 0 \\ 0 & 0.05 \end{bmatrix} \quad R_u = [0.05] \quad P = I.$$

Since negligible performance is lost by using steady-state gains, we compare with the results using $K_\infty = [-0.5522 \ -5.9690]$. The resulting control history is plotted in Fig. 2(a), with the solid curve representing the optimal steady-state control, and the dashed curves representing the one-step satisficing control for three different values of rejectivity.

The steady-state optimal trajectory is shown in Fig. 2(b) as the solid curve, and the solution of the one-step satisficing epistemic utility-based controller is displayed with the dashed curves for three different rejectivity values. The most discrim-

inating satisficing control, u_D, was employed. For comparison purposes, we define the optimal cost function $t_f \rightarrow \infty$

$$J = \frac{1}{2}\mathbf{x}(t_f)^T[P - Q(t_f)]\mathbf{x}(t_f)$$
$$+ \frac{1}{2}\sum_{t=0}^{t_f-1} \mathbf{x}(t+1)^T Q\mathbf{x}(t+1) + u(t)^T R_u u(t). \quad (49)$$

The steady-state optimal cost for this problem is $J_{\text{ss}} = 38.3$, and the cost for the one-step controllers (denoted by J_b where subscript b indicates rejectivity) are $J_{1.3} = 42.5$, $J_{1.6} = 39.9$, and $J_{1.9} = 40.0$. Thus, performance degrades approximately 4% if a satisficing solution is employed even with a planning horizon of $d = 1$. Additionally, since the normalization constants $G_A[\mathbf{x}(t)]$ and $G_L[\mathbf{x}(t)]$ can be determined in closed form as functions of u [33], the calculation of K_D consists of one matrix inversion, seven matrix multiplications, and a small number of scalar additions and multiplications. Thus, K_D can be calculated with no substantial computational burden and thus does not significantly increase computational complexity when compared to the steady-state linear state feedback. For the linear problem, these results are perhaps not very surprising since the satisficing receding horizon solution is similar to other bounded-memory approaches (see, e.g., [6] and [8]) known to have similar properties. We will have more to say about the factors of performance and computational complexity in the other examples.

C. Nonlinear, Nonstationary System with Quadratic Performance

An important control problem is the nonlinear, nonstationary regulator problem of the form (26). Suppose we wish to regulate

$$\mathbf{x}(t+1) = \mathbf{f}[\mathbf{x}(t), u(t), t], \quad \mathbf{x}(0) = \mathbf{x}_0, \quad t = 0, 1, 2, \dots \quad (50)$$

where $\mathbf{f}$ is a nonlinear vector function, about the origin $\mathbf{x} = \mathbf{0}$ in such a way that the performance index (32) is kept small. An optimal solution would minimize this performance index but, unlike the linear regulator system, no general systematic solution has been discovered for the general nonlinear problem. A conventional approach is to linearize this system about an equilibrium point and apply linear system techniques, resulting in at least a spatially local solution whose functionality is problematic. Techniques, such as feedback linearization, adaptive control, and gain scheduling are based in large part on such assumptions.

In this section, we apply a satisficing receding horizon control to a problem that has proven to be surprisingly difficult: The control of an inverted pendulum in a vertical plane with full circular freedom by applying a lateral force to the cart to which the pendulum is attached, while simultaneously regulating the position of the cart.

Consider the apparatus illustrated in Fig. 3. The problem is to bring the pendulum from vertically downward to vertically upward by applying a force to the cart. This problem is prototypical of many nonlinear control problems, and thus
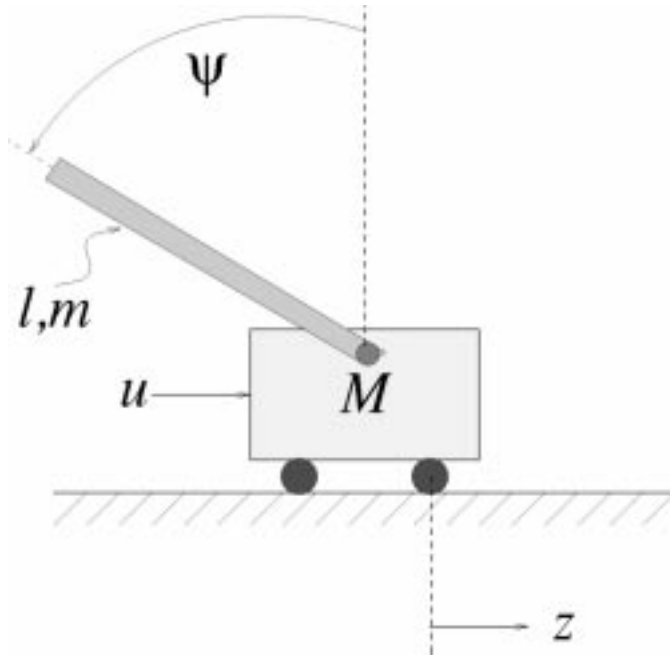
Fig. 3. Inverted pendulum on a cart.

makes a good test case for satisficing control. Traditional controllers linearize the dynamics model of the pendulum in a small region within, say $10°$ of the vertical. More recently, a fuzzy controller trained by a genetic algorithm has been shown to balance the pendulum 90% of the time if the pendulum is given a random initial position within $80°$ of the vertical and a random initial velocity less that $80°$/s [34]. In this section, we will design an epistemic utility-based controller with the control horizon $d = 1$ which will control the pendulum given any set of initial conditions, while simultaneously positioning the cart at a desired point. The only restriction made is that the initial pendulum velocity be small enough so that the sample interval is much less than the rotational period of the pendulum. To render the problem nonstationary, we will further assume that the mass of the pendulum, m, is a random walk whose future d values are known.

Let $\mathbf{x} = [\psi, z, \dot{\psi}, \dot{z}]^T$ denote the state of the cart/pendulum system. The continuous-time dynamical equation for this problem is

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})u \tag{51}$$

where $\mathbf{f} = [f_1, f_2, f_3, f_4]^T$ and $\mathbf{g} = [g_1, g_2, g_3, g_4]^T$ are given by

$$f_1(\mathbf{x}) = \dot{\psi} \tag{52}$$

$$f_2(\mathbf{x}) = \dot{z} \tag{53}$$

$$f_3(\mathbf{x}) = \frac{(M+m)g\sin\psi - ml\cos\psi\sin\psi\dot{\psi}^2}{l(M+m\sin^2\psi)} \tag{54}$$

$$f_4(\mathbf{x}) = \frac{ml\sin\psi\dot{\psi}^2 - mg\cos\psi\sin\psi}{M+m\sin^2\psi}. \tag{55}$$

$$g_1(\mathbf{x}) = 0 \tag{56}$$

$$g_2(\mathbf{x}) = 0 \tag{57}$$

$$g_3(\mathbf{x}) = \frac{-\cos\psi}{l(M+m\sin^2\psi)} \tag{58}$$

$$g_4(\mathbf{x}) = \frac{1}{M+m\sin^2\psi}. \tag{59}$$

where

M mass of the cart;
l length of the pendulum;
m mass of the pendulum;
ψ angle from vertical (measured counterclockwise);
z horizontal position of the cart;
u control input, is a lateral force applied to the cart.

1) Derivation of Accuracy Support and Liability Exposure: Using Euler integration with sample time T, the equivalent discrete-time dynamical expression is

$$\mathbf{x}(t+1) = \mathbf{x}(t) + T\{\mathbf{f}[\mathbf{x}(t)] + \mathbf{g}[\mathbf{x}(t)]u(t)\} \tag{60}$$

for $t = 0, 1, \ldots$. With this discretization procedure, $\psi(t+1)$ and $z(t+1)$ are not explicit functions of $u(t)$, but $\psi(t+2)$ and $z(t+2)$ are explicit functions of $u(t)$. Thus, $D_t = 2$, and we may identify the components of the influence vector $\boldsymbol{\xi}_0(u) = \{\xi_0(t+1), \xi_0(t+2)\}$ as

$$\xi_0(t+1) = \begin{bmatrix} \dot{\psi}(t+1) \\ \dot{z}(t+1) \end{bmatrix}$$
$$= \begin{bmatrix} \dot{\psi}(t) + T\{f_3[\mathbf{x}(t)] + g_3[\mathbf{x}(t)]u(t)\} \\ \dot{z}(t) + T\{f_4[\mathbf{x}(t)] + g_4[\mathbf{x}(t)]u(t)\} \end{bmatrix} \tag{61}$$

and

$$\xi_0(t+2) = \begin{bmatrix} \psi(t+2) \\ z(t+2) \end{bmatrix}$$
$$= \begin{bmatrix} \psi(t) + 2Tf_1[\mathbf{x}(t)] + T^2\{f_3[\mathbf{x}(t)] + g_3[\mathbf{x}(t)]u(t)\} \\ z(t) + 2Tf_2[\mathbf{x}(t)] + T^2\{f_4[\mathbf{x}(t)] + g_4[\mathbf{x}(t)]u(t)\} \end{bmatrix}. \tag{62}$$

We adopt the same quadratic performance indexes (33)–(34) that were used for the previous linear quadratic regulator example. Since the inverted pendulum is a regulator problem, the goal of the system is the same as for that problem; namely, bring the system to rest at the desired point. Let $Q(t) = \begin{bmatrix} Q_p(t) & 0 \\ 0 & Q_v(t) \end{bmatrix}$, where Q_p and Q_v are 2×2 matrices with $Q_p(t_f) = Q_v(t_f) = 0$, and let $P - Q(t_f) = I$. The local performance indexes then become

$$\Phi[\boldsymbol{\xi}_0(u)] = \xi_0^T(t+2)\xi_0(t+2) + \xi_0^T(t+1)\xi_0(t+1) \tag{63}$$

and

$$\Lambda[\boldsymbol{\xi}_0(u)] = \xi_0^T(t+2)Q_p(t+2)\xi_0(t+2) + \xi_0^T(t+1)$$
$$\times Q_v(t+1)\xi_0(t+1) + R_u(t)u^2 \tag{64}$$

in accordance with (27) through (30).

From these equations, $\bar{f}_A$ and $\bar{f}_L$ may be calculated in accordance with (27), (29), (28), and (30). Thus, problems with linear dynamics and nonlinear dynamics are both treated exactly the same; neither linearity nor time-invariance are exploited. The resulting most discriminating controller is, after some calculations,

$$u_D = -[G^T[\mathbf{x}(t)](I + b'Q(t+1))G[\mathbf{x}(t)] + b'R_u(t)]^{-1}$$
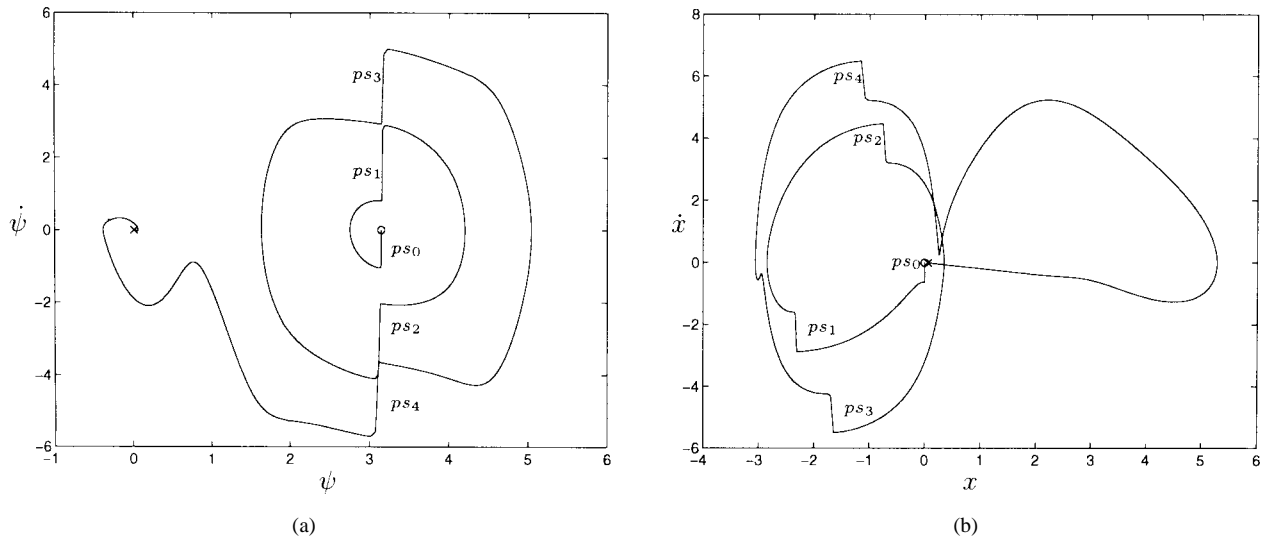$$\times G^T[\mathbf{x}(t)](I + b'Q(t+1))F[\mathbf{x}(t)] \tag{65}$$

Fig. 4. Phase planes for the inverted pendulum on a cart ($m = $ constant): (a) rotational phase plane (in radians and radians per second) and (b) translational phase plane (in meters and meters per second).

where

$$G[\mathbf{x}(t)] = [T^2 g_3[\mathbf{x}(t)] \quad T^2 g_4[\mathbf{x}(t)] \quad T g_3[\mathbf{x}(t)] \quad T g_4[\mathbf{x}(t)]]^T$$

$$F[\mathbf{x}(t)] = \begin{bmatrix} T^2 f_3[\mathbf{x}(t)] + 2T\dot{\psi}(t) + \psi(t) \\ T^2 f_4[\mathbf{x};(t)] + 2T\dot{x}(t) + x(t) \\ T f_3[\mathbf{x}(t)] + \dot{\psi}(t) \\ T f_4[\mathbf{x}(t)] + \dot{x}(t) \end{bmatrix}.$$

The following values were used in the simulation: $Q(t) = \begin{bmatrix} 30 & 1 & 0 & 0 \\ 1 & 0.3 & 0 & 0 \\ 0 & 0 & 0.2 & 0.2 \\ 0 & 0 & 0.2 & 0.2 \end{bmatrix}$, $R_u(t) = 10^{-4}$, $M = 0.455$ kg, $m = 0.21$ kg, $l = 0.61$ m, $T = 0.01$ s, and $U = [-1000, 1000]$. The off-diagonal terms in $Q(t)$ reflect coupling between ψ and z necessitated by one control input but two degrees of freedom in the system, namely the rotational and translational components.

2) Simulation Results and Discussion of Optimal Solutions: As a baseline, we first present results for the time-invariant case, that is, $m(t) = m$, a constant; we then present results for $m(t)$ a random walk. Fig. 4(a) and (b) illustrate the rotational (pendulum) and translational (cart) phase planes for the constant mass case. The "o" symbol represents the initial conditions (the cart at the origin with the pendulum in the vertical down position) and the "×" symbol represents the terminal conditions (the cart at the origin with the pendulum balanced in the vertical up position).

Fig. 5 provides the control time history for this problem. The system achieves its desired objective of balancing the pendulum at the origin by swinging the pendulum back and forth while the cart oscillates around the origin. As the cart oscillates, the pendulum gathers momentum. In the translational and rotational phase planes, this motion is manifest as growing spirals. When the amplitude increases sufficiently, the oscillation ceases and the pendulum then converges to the vertical upright position. Finally, the cart returns slowly to the origin.
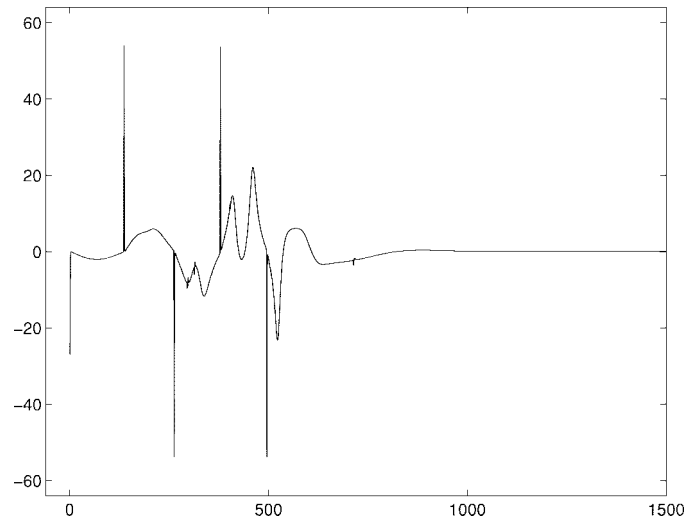

Fig. 5. Control inputs for the inverted pendulum on a cart (time measured in 0.01 s increments).

An interesting feature of this controller is that jumps occur in both translational and angular velocity when the pendulum swings through $\psi = -\pi$. This phenomenon is a consequence of the coupling between translational and rotational position in the liability exposure density function, which causes a polarity switch, denoted by ps_i in the figure, to occur between u_L and u_A. On one side of the vertical, $u_L < u_A$, and on the opposite side, $u_L > u_A$. This polarity switch creates a large change in u_D, and the resulting change in rotational rate then acts to restore the control to near its value before the polarity switch. Thus, the phenomenon appears as an impulsive control input when the pendulum goes through the vertical down position, as illustrated in Fig. 5.

Fig. 6(a) and (b) illustrate the rotational and translational phase planes for the random-mass case. Here, $m(t + 1) = m(t) + v(t)$, with $v(t)$ an uncorrelated process with each time-sample drawn from a uniform distribution over the interval $(-0.01, 0.01)$. Note that, although the trajectories differ
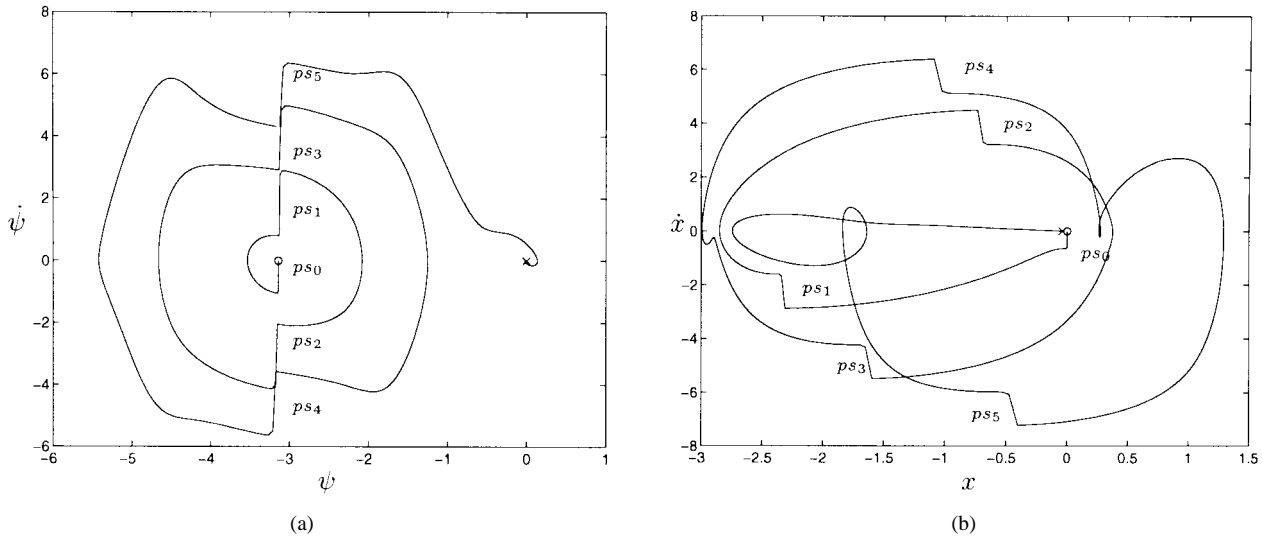
Fig. 6. Phase planes for the inverted pendulum on a cart (m = random): (a) rotational phase plane (in radians and radians per second) and (b) translational phase plane (in meters and meters per second).

quantitatively from those of the constant-mass case, the qualitative behavior is very similar. The fact that the mass is random has no effect on either the controller structure or design procedure.

It is possible to construct a solution to the discretized nonlinear problem using dynamic programming. However, such solutions can only be constructed when $m(t)$ is known for all time. In the absence of such prior information, either global robust control methods or receding horizon methods would be required, but no systematic results of comparable generality regarding the inverted pendulum are known to the authors. In summary, the satisficing approach provides a solution to this problem which could, in theory, be improved using iterative optimal control methods (such as [35]). Since closed form solutions for the controller can be obtained, computational complexity is not a significant issue.

D. Nonlinear System With Minimum-Time Performance

In this section we apply epistemic utility-based control theory to Zermelo's problem. Zermelo's problem is a minimum time problem with nonlinear dynamics for which an optimal solution is known [31], [36]. We first present the satisficing solution to this problem and compare the performance to the optimal solution.

Zermelo's problem involves a ship that must travel through a region of strong currents to reach an island, placed at the origin of a Cartesian coordinate system, in minimum time. The current vector, $[c_x, c_y]$, in Cartesian coordinates, is given as $c_x = 0$, $c_y = Vy/h$, as shown in Fig. 7, where V is the magnitude of the ship's velocity relative to the water, and h is a fixed constant. The ship's heading angle, $u = \psi$, is defined relative to the positive x axis, and is the control variable for this problem. We define $U = [0, 2\pi]$. The discretized equations of motion are

$$x(t+1) = x(t) + TV[\cos\psi(t) + y(t)/h]$$
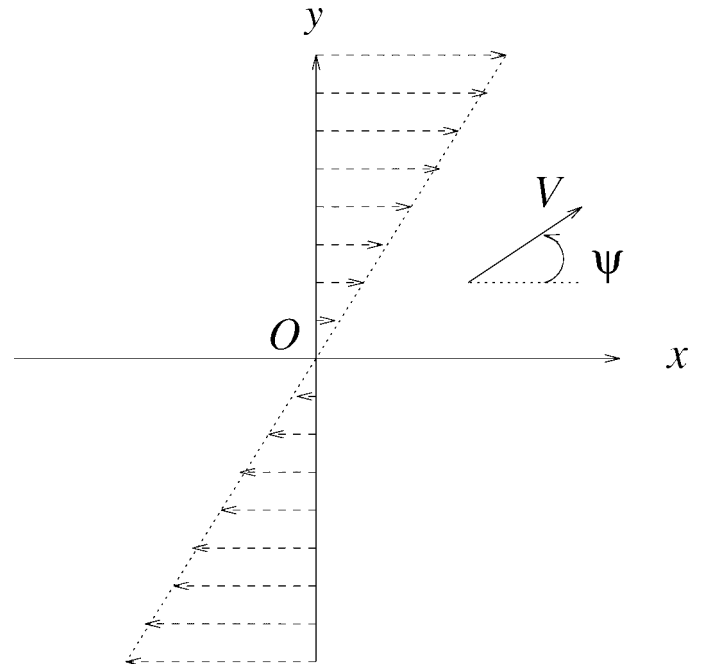$$y(t+1) = y(t) + TV\sin\psi(t)$$

Fig. 7. Geometry for Zermelo's problem.

which can easily be put in the discretized form of (26) by defining $\mathbf{x}(t) = [x(t), y(t)]^T$, $\mathbf{f} = V[y(t)/y, 0]^T$, $\mathbf{g} = \mathbf{1}$, and $\mathbf{h} = V[\cos\psi(t), \sin\psi(t)]^T$. The explicit influence horizon for this problem is $D_t = 1$, since the control, $\psi(t)$, explicitly affects both $x(t+1)$ and $y(t+1)$. For a d-length control horizon, the influence vectors are $\boldsymbol{\xi}_\ell(u(t+\ell)) = \mathbf{x}(t+\ell+1)$, $0 \le \ell \le d-1$. Specifically, for $d = 1$,

$$\boldsymbol{\xi}_0(\psi) = \mathbf{x}(t+1) = \begin{bmatrix} x(t) + TV[\cos\psi(t) + y(t)/h] \\ y(t) + TV\sin\psi(t) \end{bmatrix}.$$

For minimum-time problems with constrained final state, the performance index (23) is

$$J = \mathbf{x}^T(t_f)\mathbf{x}(t_f) + \sum_{t=0}^{t_f-1} 1 = \mathbf{x}^T(t_f)\mathbf{x}(t_f) + t_f \qquad (66)$$

where $\mathbf{x}(t) = [x(t),\ y(t)]^T$. We wish to find a minimum-time satisficing control. From the local point of view, the essence of minimizing time is to maximize velocity. Consequently, we must include a measure of velocity, or position change, in the performance index. Adding and subtracting velocity related components to (66) yields

$$J = \mathbf{x}^T(t_f)\mathbf{x}(t_f) + t_f \pm \sum_{j=1}^{d}[\mathbf{x}(t_f - d + j)$$
$$- \mathbf{x}(t_f - d + j - 1)]^T[\mathbf{x}(t_f - d + j)$$
$$- \mathbf{x}(t_f - d + j - 1)] \tag{67}$$

which may be rearranged as

$$J = \mathbf{x}^T(t_f)\mathbf{x}(t_f) + \sum_{j=1}^{d}[\mathbf{x}(t_f - d + j) - \mathbf{x}(t_f - d + j - 1)]^T$$
$$\times [\mathbf{x}(t_f - d + j) - \mathbf{x}(t_f - d + j - 1)] + t_f$$
$$- \sum_{j=1}^{d}[\mathbf{x}(t_f - d + j) - \mathbf{x}(t_f - d + j - 1)]^T$$
$$\times [\mathbf{x}(t_f - d + j) - \mathbf{x}(t_f - d + j - 1)]. \tag{68}$$

1) Accuracy Support and Liability Exposure Construction: From a local perspective, $t + d$ is associated with the terminal time, t_f. Thus, a local terminal performance index may be formed from the first two terms on the right-hand side of (68) to obtain

$$\Phi[\boldsymbol{\xi}_0(\psi_0), \ldots, \boldsymbol{\xi}_{d-1}(\psi_{d-1})]$$
$$= \mathbf{x}^T(t + d)\mathbf{x}(t + d) + \sum_{j=1}^{d}[\mathbf{x}(t + j) - \mathbf{x}(t + j - 1)]^T$$
$$\times [\mathbf{x}(t + j) - \mathbf{x}(t + j - 1)]. \tag{69}$$

For the case $d = 1$ we drop t dependence on $\psi(t)$ whence substitution of (69) into (27) yields

$$g_A[\psi; \mathbf{x}(t)] = \sup_{z \in U}\{2\mathbf{x}^T(t + 1)[\mathbf{x}(t + 1) - \mathbf{x}(t)]\}$$
$$- 2\mathbf{x}^T(t + 1)[\mathbf{x}(t + 1) - \mathbf{x}(t)] + \epsilon \tag{70}$$
$$= \sup_{z \in U}\left\{2\begin{bmatrix} x(t) + TV[\cos z + y(t)/h] \\ y(t) + TV\sin z \end{bmatrix}^T\right.$$
$$\times \left.\begin{bmatrix} TV[\cos z + y(t)/h] \\ TV\sin z \end{bmatrix}\right\}$$
$$- \begin{bmatrix} x(t) + TV[\cos\psi + y(t)/h] \\ y(t) + TV\sin\psi \end{bmatrix}^T$$
$$\times \begin{bmatrix} TV[\cos\psi + y(t)/h] \\ TV\sin\psi \end{bmatrix} + \epsilon. \tag{71}$$

The essence of (70) is the inner product $\mathbf{x}^T(t+1)[\mathbf{x}(t) - \mathbf{x}(t+1)]$. Its effect on accuracy support is best understood by considering a few simple scenarios. Suppose that the present position, $\mathbf{x}(t)$, is positive (to the right of the origin). If a control, $\psi(t)$, produces a next state, $\mathbf{x}(t+1)$, that is more positive than the present position, then the inner product is negative and the corresponding accuracy support, after shifting, is small. If, however, the next state is positive and to

the left of the present state, then the inner product is positive and the corresponding accuracy support, after shifting, is high. Similar results for other regions of state space indicate that this inner product yields a large accuracy support for controls that move the system quickly toward the origin, but low accuracy support for controls that move the system in the wrong direction or that move the system slowly. This inner product is the local basis for constructing the accuracy support density function.

To construct the liability exposure density function, we use the second two terms from (68), and again associate t_f with $t + d$, to form

$$\Lambda[\boldsymbol{\xi}_0(\psi_0), \ldots, \boldsymbol{\xi}_{d-1}(\psi_{d-1})]$$
$$= t + d - \sum_{j=1}^{d}[\mathbf{x}(t + j) - \mathbf{x}(t + j - 1)]^T$$
$$\times [\mathbf{x}(t + j) - \mathbf{x}(t + j - 1)]. \tag{72}$$

For the case $d = 1$, substituting (72) into (29) yields, after some manipulation,

$$g_L[\psi; \mathbf{x}(t)] = \sup_{z \in U}\{[\mathbf{x}(t + 1) - \mathbf{x}(t)]^T[\mathbf{x}(t + 1) - \mathbf{x}(t)]\}$$
$$- [\mathbf{x}(t + 1) - \mathbf{x}(t)]^T[\mathbf{x}(t + 1) - \mathbf{x}(t)] + \epsilon \tag{73}$$
$$= \sup_{z \in U}\left\{\begin{bmatrix} TV[\cos z + y(t)/h] \\ TV\sin z \end{bmatrix}^T\right.$$
$$\times \left.\begin{bmatrix} TV[\cos z + y(t)/h] \\ TV\sin z \end{bmatrix}\right\}$$
$$- \begin{bmatrix} TV[\cos\psi + y(t)/h] \\ TV\sin\psi \end{bmatrix}^T$$
$$\times \begin{bmatrix} TV[\cos\psi + y(t)/h] \\ TV\sin\psi \end{bmatrix} + \epsilon. \tag{74}$$

The liability exposure density function constructed from this g_L places large liability exposure on low speeds and small liability exposure on large speeds, indicating that wherever the agent travels, it should go there speedily; that is, the velocity term in (73) should be maximized. This is similar to the optimal bang-bang controllers that originate from minimum time problems with bounded controls.

2) Results and Comparison to Optimal Control: For these dynamics, the accuracy support and liability exposure are formed from (71) and (74) and the resulting boat trajectories are plotted for $V = 4.0$, $h = 1.0$, $T = 0.01$, $x(0) = 4.9h$, and $y(0) = 1.66h$ in Fig. 8. Trajectories are given for the optimal solution, a one-step satisficing solution, and a two-step satisficing solution, where the two-step satisficing solution is given for three different values of rejectivity. The corresponding times are given in Table I.

The one-step satisficing solution does not recognize that by crossing over the $y = 0$ axis the boat can gain speed, since the liability exposure for the one-step solution is uniform at $y(t) = 0$. The two-step satisficing solution, by contrast, takes advantage of the $y = 0$ crossing and decreases the amount of time required. High rejectivity ($b = 0.9$) causes the boat to overshoot the island and backtrack. Lower rejectivity ($b = 0.7$) eliminates the overshoot and decreases time, but
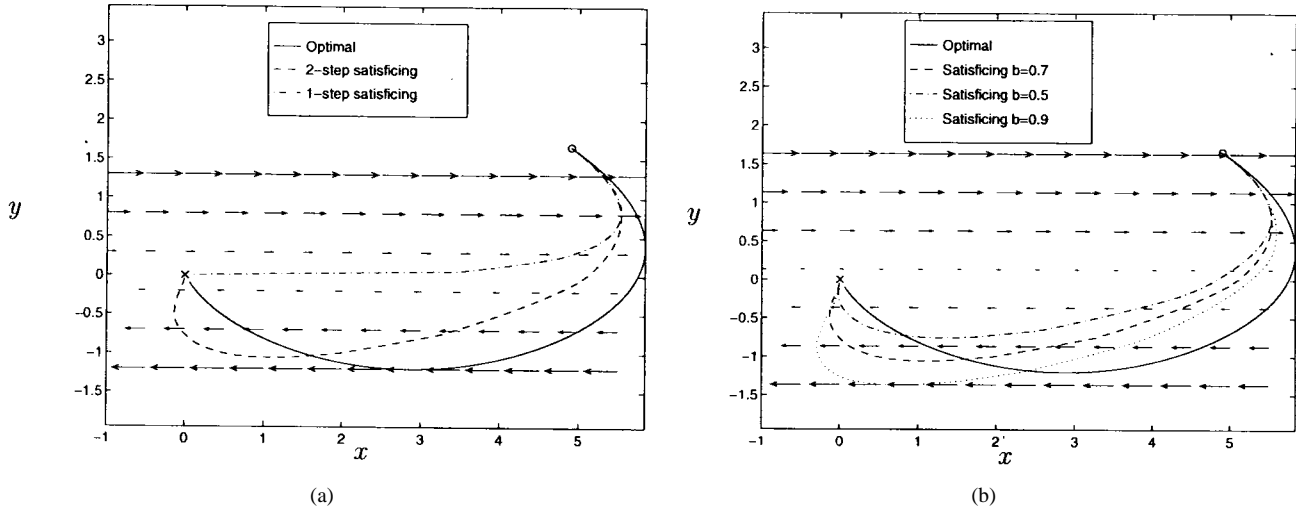
Fig. 8.   Performance results for the single-agent Zermelo's problem: (a) comparison of optimal, one-step satisficing, and two-step satisficing, and (b) comparison of optimal with two-step satisficing for varying index of rejectivity.

TABLE I
ZERMELO'S PROBLEM RESULTS

Control Method	Δt
Optimal	1.40
1-step satisficing ($b = 0.7$)	1.98
2-step satisficing ($b = 0.5$)	1.66
2-step satisficing ($b = 0.7$)	1.64
2-step satisficing ($b = 0.9$)	1.80

decreasing the rejectivity further ($b = 0.5$) increases time since currents are not fully exploited. Although changing rejectivity does affect performance slightly, the satisficing solutions are not overly sensitive to rejectivity. Thus, performance of the two-step horizon satisficing solution is not significantly worse (approximately 17% for $b = 0.7$) than the optimal control. The computational complexity of the satisficing solution is higher because no closed form solution is employed (calculation of the equilibrium set is obtained using numerical derivatives, and a search is made through the discretized control space for the most discriminating control). In the presence of nonlinear currents, however, no optimal solution of the continuous time problem is easily obtained, but a satisficing solution is of the form developed above is easily generated [33].

IV. CONSISTENCY OF RECEDING HORIZON SATISFICING CONTROL

The use of a receding horizon makes it possible to develop a tractable satisficing controller using the principles of epistemic utility theory. As the length of the control horizon, d, is increased, more of the future state values are taken into consideration for the calculation of the current control. It is therefore reasonable to expect that performance will improve with increasing d. The following theorem establishes the stronger result that in the limit as d approaches t_f, the quadratic regulator satisficing control will actually be equivalent to the optimal control.

Theorem 8 (Consistency of Quadratic Regulator): For the deterministic quadratic regulator problem (50) and (32), if the control horizon spans the full extent of the problem, that is, for $d = t_f$, then the most discriminating satisficing control is identical to the optimal control.

Proof: We will prove this result for the scalar control case only. The most discriminating control, denoted $\mathbf{u}_D = [u_0, \ldots, u_{t_f-1}]^T$, is $\mathbf{u}_D = \arg\sup_{\mathbf{z} \in \mathbb{R}^{t_f}} \{\bar{f}_A(\mathbf{z}) - b\bar{f}_L(\mathbf{z})\}$. From (28) and (30)

$$\arg\sup_{\mathbf{z} \in \mathbb{R}^{t_f}} \{\bar{f}_A(\mathbf{z}) - b\bar{f}_L(\mathbf{z})\} = \arg\sup_{\mathbf{z} \in \mathbb{R}^{t_f}} \{g_A(\mathbf{z}) - b'g_L(\mathbf{z})\} \tag{75}$$

where $b' = b\dfrac{\int_{U^{t_f}} g_A(\mathbf{z}; t)\, d\mathbf{z}}{\int_{U^{t_f}} g_L(\mathbf{z}; t)\, d\mathbf{z}}$. Also, from (27) and (29),

$$\arg\sup_{\mathbf{z} \in \mathbb{R}^{t_f}} \{g_A(\mathbf{z}) - b'g_L(\mathbf{z})\}$$
$$= \arg\inf_{\mathbf{z} \in \mathbb{R}^{t_f}} \{\Phi[\boldsymbol{\xi}_0(z_0), \ldots, \boldsymbol{\xi}_{t_f-1}(z_{t_f-1})] + b'\Lambda[\boldsymbol{\xi}_0(z_0), \ldots, \boldsymbol{\xi}_{t_f-1}(z_{t_f-1})]\}. \tag{76}$$

For the quadratic regulator problem with $d = t_f$, we take $\Phi[\boldsymbol{\xi}_0(u_0), \ldots, \boldsymbol{\xi}_{t_f-}(u_{t_f-1})] = \mathbf{x}^T(t_f)P\mathbf{x}(t_f)$ and $\Lambda[\boldsymbol{\xi}_0(u_0), \ldots, \boldsymbol{\xi}_{t_f-}(u_{t_f-1})] = \sum_{t=0}^{t_f-1} \mathbf{x}^T(t+1)Q(t+1)\mathbf{x}(t+1) + R_u(t)z_t^2$. Thus,

$$\mathbf{u}_D = \arg\inf_{\mathbf{z} \in \mathbb{R}^{t_f}} \left\{ \mathbf{x}^T(t_f)P\mathbf{x}(t_f) + b' \sum_{t=0}^{t_f-1} \mathbf{x}^T(t+1) \right.$$
$$\left. \times Q(t+1)\mathbf{x}(t+1) + R_u(t)u^2(t) \right\}. \tag{77}$$

For $b' = 1$, this is exactly the optimal quadratic regulator solution. $\quad\square$

The following theorem establishes the result that in the limit as d approaches t_f, the minimum-time satisficing control is equivalent to the optimal control. The proof is similar to the proof of Theorem 8.

Theorem 9 (Consistency of Minimum Time Controller): For the deterministic minimum-time (50) and (66), if the control horizon spans the full extent of the problem, that is, for $d = t_f$, then the most discriminating satisficing control is identical to the optimal minimum-time control.

V. CONCLUSION

We present a theory of satisficing control that builds upon the philosophy of avoiding error rather than seeking truth. The mature theoretical foundations of this philosophy are translated into a controller design procedure for single-agent, time-varying, nonlinear systems. Solutions are generated for two well-known optimal control problems (LQR and Zermelo's problem) and performance and computational requirements are compared. Using these examples, we demonstrate that satisficing controllers do not require significantly more computations and produce behavior which compares favorably to optimal solutions. When computation of optimal solutions is infeasible (such as the uncertain time-varying inverted pendulum problem), we demonstrate the ability to generate computable solutions in the presence of time-varying plants.

Unlike traditional receding horizon control methods, we independently assess controller performance on the basis of terminal conditions (such as regulation) and transition costs (such as fuel consumption). This independent assessment makes explicit the asymmetry between the fundamental goal of the controller and undesirable performance characteristics. The tradeoff that appears in many controller designs between these objectives is made explicit, and only controls for which accuracy support (benefit with respect to fundamental goal achievement) exceeds liability exposure (cost with respect to undesirable characteristics are permitted). This independent assessment facilitates the design of a receding horizon controller for the inverted pendulum problem with full circular freedom. Unlike conventional generalized potential field approaches with attractive goals and repulsive obstacles, satisficing controllers consider explicit performance objectives, and are consistent with optimal solutions in the sense that, when the planning horizon is sufficient, performance is optimal.

REFERENCES

[1] H. A. Simon, "A behavioral model of rational choice," *Quart. J. Econom.*, vol. 59, pp. 99–118, 1955.

[2] H. A. Simon, *The Sciences of the Artificial*, 3rd ed. Cambridge, MA: MIT Press, 1996.

[3] H. Michalska and D. Q. Mayne, "Moving horizon observers and observer-based control," *IEEE Trans. Automat. Contr.*, vol. 40, pp. 995–1006, June 1995.

[4] P. B. Sistu and B. W. Bequette, "Nonlinear model-predictive control: Closed-loop stability analysis," *AIChE J.*, vol. 42, pp. 3388–3402, Dec. 1996.

[5] J. Richalet, "Industrial applications of model based predictive control," *Automatica*, vol. 29, pp. 1251–1274, 1993.

[6] D. Q. Mayne and H. Michalska, "Receding horizon control of nonlinear systems," *IEEE Trans. Automat. Contr.*, vol. 35, pp. 814–824, 1990.

[7] P. O. M. Scokaert, J. B. Rawlings, and E. S. Meadows, "Discrete-time stability with perturbations: Application to model predictive control," *Automatica*, vol. 33, pp. 463–470, Mar. 1997.

[8] O. Hernández-Lerma and J. B. Lasserre, "Error bounds for rolling horizon policies in discrete-time Markov control processes," *IEEE Trans. Automat. Contr.*, vol. 35, pp. 1118–1124, 1990.

[9] S. Lall and K. Glover, "A game theoretic approach to moving horizon control," *Advances in Model-Based Predictive Control*. Oxford, U.K.: Oxford Univ. Press, 1994.

[10] S. Lall and K. Glover, "Robust performance and adaptation using receding horizon H_∞ control of time varying systems," in *Proc. Amer. Contr. Conf. 1995*, pp. 2384–2388.

[11] Y. S. Nam, B. H. Lee, and N. Y. Ko, "A view-time based potential field method for moving obstacle avoidance," in *SICE*, pp. 1463–1468, 1996.

[12] J. Guldner and V. I. Utkin, "Sliding mode control for an obstacle avoidance strategy based on an harmonic potential field," in *Proc. 32nd Conf. Decision Control*, San Antonio, TX, Dec. 1993, pp. 424–429.

[13] C. L. Connolly and R. A. Grupen, "On the applications of harmonic functions to robotics," *J. Robot. Syst.*, vol. 10, no. 7, pp. 931–946, 1993.

[14] J.-O. Kim and P. K. Khosla, "Real-time obstacle avoidance using harmonic potential functions," *IEEE Trans. Robot. Automat.*, vol. 8, pp. 338–349, June 1992.

[15] S. Sundar and Z. Shiller, "Optimal obstacle avoidance based on the Hamilton-Jacobi-Bellman equation," *IEEE Trans. Robot. Automat.*, vol. 13, pp. 305–310, Apr. 1997.

[16] M. D. Mesarovic, "Systems theoretic approach to formal theory of problem solving," in *Theoretical Approaches to Non-Numerical Problem Solving*, R. Banerji and M. D. Mesarovic, Eds. Berlin, Germany: Springer, 1970, pp. 161–178.

[17] M. D. Mesarovic and Y. Takahara, "On a qualitative theory of satisfactory control," *Inf. Sci.*, vol. 4, pp. 291–313, Oct. 1972.

[18] T. Matsuda and S. Takatsu, "Characterization of satisficing decision criterion," *Inf. Sci.*, vol. 17, pp. 131–151, 1979.

[19] ——, "Algebraic properties of satisficing decision criterion," *Inf. Sci.*, vol. 17, pp. 221–237 17.

[20] S. Takatsu, "Decomposition of satisficing decision problems," *Inf. Sci.*, vol. 22, no. 2, pp. 139–148, 1980.

[21] ——, "Latent satisficing decision criterion," *Inf. Sci.*, vol. 25, no. 2, pp. 145–152, 1981.

[22] I. Levi, *The Enterprise of Knowledge*. Cambridge, MA: MIT Press, 1980.

[23] P. N. Johnson-Laird, *The Computer and the Mind: An Introduction to Cognitive Science*. Cambridge, MA: Harvard Univ. Press, 1988.

[24] W. C. Stirling and D. R. Morrell, "Convex Bayes decision theory," *IEEE Trans. Syst., Man, Cybern.*, vol. 21, pp. 173–183, Jan./Feb. 1991.

[25] W. C. Stirling, "Coordinated intelligent control via epistemic utility theory," *IEEE Contr. Syst.*, vol. 13, pp. 21–29, Oct. 1993.

[26] I. Levi, *Decisions and Revisions*. London, U.K.: Cambridge Univ. Press, 1984.

[27] J. M. Mendel, "Fuzzy logic systems for engineering: A tutorial," *Proc. IEEE*, vol. 83, pp. 345–377, Mar. 1995.

[28] E. Cox, "Fuzzy fundamentals," *IEEE Spectrum*, pp. 58–61, Oct. 1992.

[29] M. A. Goodrich, W. C. Stirling, and R. L. Frost, "A satisficing fuzzy logic controller," in *Proc. FUZZ-IEEE'96 Int. Conf. Fuzzy Systems*, New Orleans, LA, Sept. 8–11, 1996.

[30] H. Kwakernaak, *Linear Optimal Control Systems*. New York: Wiley, 1972.

[31] F. L. Lewis, *Optimal Control*. New York: Wiley, 1976.

[32] D. Kirk, *Optimal Control Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1970.

[33] M. A. Goodrich, "A theory of satisficing control," Ph.D. dissertation, Brigham Young Univ., Provo, UT, 1996.

[34] J. Kim, Y. Moon, and B. P. Zeigler, "Designing fuzzy net controllers using genetic algorithms," *IEEE Contr. Syst.*, vol. 15, no. 3, pp. 66–72, 1995.

[35] R. Beard, G. Saridis, and J. Wen, "Improving the performance of stabilizing controls for nonlinear systems," *IEEE Contr. Syst.*, vol. 16, pp. 27–35, Oct. 1996.

[36] A. E. Bryson and Y. Ho, *Applied Optimal Control*. London, U.K.: Blaisdell, 1969.

Michael A. Goodrich received the B.S., M.S., and Ph.D. degrees (cum laude) in electrical engineering from Brigham Young University, Provo, UT, in 1992, 1995, and 1996, respectively.

He is currently doing post-doctoral work at Nissan Cambridge Basic Research, Nissan Research and Development, Inc., Cambridge, MA, in intelligent systems modeling and control. His research interests include modeling and controlling intelligent systems, decision theory, multiple-agent coordination, estimation theory, and fuzzy logic.

Wynn C. Stirling received the Honors B.A. degree (magna cum laude) in mathematics and the M.S. degree in electrical engineering from the University of Utah, Salt Lake City, in 1969 and 1971, respectively, and the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA, in 1983.

From 1972 to 1975, he was with Rockwell International Corporation, Anaheim, CA, and from 1975 to 1984, he was with ESL, Inc., Sunnyvale, CA. Since 1984, he has been with Brigham Young University, Provo, UT, where he is a Professor in the Department of Electrical and Computer Engineering. His research interests include decision theory, control theory, estimation theory, and stochastic processes.

Dr. Stirling is a member of Phi Beta Kappa.

Richard L. Frost received the B.S. degree in physics in 1975 (magna cum laude), the M.S.E.E. degree in 1977, and the Ph.D. degree in 1979 in electrical engineering, all from the University of Utah, Salt Lake City.

He was with the Lincoln Laboratory, Massachusetts Institute of Technology, Cambridge, from 1979 to 1981, on the faculty of the University of Utah from 1981 to 1984, and with the Communication Systems Division of Sperry (now L3 Communications) from 1984 to 1987, when he joined the faculty at Brigham Young University, Provo, UT. He is presently an Associate Professor in the Department of Electrical and Computer Engineering, with principal research interests in quantization and source coding and in intelligent control.