Stability of Model Predictive Control using Markov Chain Monte Carlo Optimisation

Elilini Siva, Paul Goulart, Jan Maciejowski and Nikolas Kantas

Abstract— **We apply stochastic Lyapunov theory to perform stability analysis of MPC controllers for nonlinear deterministic systems where the underlying optimisation algorithm is based on Markov Chain Monte Carlo (MCMC) or other stochastic methods. We provide a set of assumptions and conditions required for employing the approximate value function obtained as a stochastic Lyapunov function, thereby providing almost sure closed loop stability. We demonstrate convergence of the system state to a target set on an example, in which simulated annealing with finite time stopping is used to control a nonlinear system with non-convex constraints.**

I. INTRODUCTION

Model predictive control (MPC) solves an open loop optimal control problem at each time step [7]. The underlying optimisation problems that must be solved in MPC schemes for nonlinear systems with arbitrary constraints and objective functions defined on continuous spaces are in general non-convex. The existence of multimodalities in the objective presents further challenges in solving for global optima. In such instances, options for the optimisation method are restricted to stochastic optimisation techniques. This paper is concerned with the application of Markov Chain Monte Carlo (MCMC) methods to solving for MPC control laws for deterministic nonlinear systems.

Simulated annealing can admit an MCMC formulation for expected utility optimisation [9], [1]. This has previously been employed by [6] for nonlinear MPC with arbitrary disturbances, with the incorporation of probabilistic constraint satisfaction. Problems of feasibility and stability were not addressed there.

This paper addresses the closed-loop stability of an MPC control law implemented via the repeated application of an MCMC optimiser, specifically simulated annealing, in a receding horizon fashion. Probabilistic finite time guarantees for general simulated annealing methods have been obtained for instance by [5], in which the desired precision of the 'approximate optimum' [14] obtained is controlled by the choice of stopping temperature. The MPC control laws obtained with finite time stopping simulated annealing are suboptimal.

Conditions under which suboptimal control of deterministic systems is stabilising are examined by [8] and [12]. A modified version of the standard Lyapunov stability theorem allowing for nonuniqueness and discontinuity in the control law is presented in [12]. One condition for convergence is that the cost function reduces at each time step. This is achieved by enforcing a terminal set constraint, in which a locally asymptotically stabilizing control law exists. Key conditions on the terminal set and controller are presented in [8] to ensure that the value function is Lyapunov. The conditions ensure that it is sufficient to find a feasible solution at each timestep to guarantee stability.

The use of stochastic optimisation for obtaining MPC control laws introduces probabilistic uncertainty into the system, even when the system dynamics are deterministic. A need for establishing stability in this context therefore exists. It is not possible to meet the stability requirement in [8] and ensure reduction of the approximate value function at each time step. The standard notion of Lyapunov stability is inapplicable, so we consider stochastic Lyapunov functions, which are processes having supermartingale properties [3] in the neighbourhood of the stable point. Our contribution is the application of stochastic stability theory [13], [3] to obtain an extension of the existing stability result obtained by [8] when approximate optima are obtained with MCMC optimisation. We obtain conditions on the approximate value function to be a stochastic Lyapunov function. Specifically, we show that this is achieved by making the correct choice of stopping temperature for the simulated annealing optimisation at each time step. An MPC strategy with MCMC optimisation is outlined and its stabilising properties discussed.

Some of the notation used in this paper is outlined here. The state and control inputs are denoted x and v respectively. We use x_k to depict the *actual* measured state at time k. At the time instant k, the *prediction* of the state i steps in the future is denoted $x_{k+i|k}$. The joint process (x_k, v_k) is represented as ϕ_k. We denote the expectation of a real valued function of a random variable $Z \sim \pi$ as $\mathbb{E}_\pi[h(Z)] = \int h(z)\pi(z)\mathrm{d}z$. A function $\alpha(\cdot)$, defined on nonnegative reals, is a $\mathcal{K}$-function if it is continuous and strictly increasing with $\alpha(0) = 0$.

The paper is organised as follows. In Section II, the problem formulation and background stochastic stability theory is presented. The MPC formulation and proposed MCMC optimisation procedure are outlined in section III. In section IV, the stabilising properties of the proposed MPC strategy are proved. An illustrative example is presented in Section V and concluding remarks are made in Section VI.

II. PROBLEM FORMULATION AND BACKGROUND

We consider controlling a discrete time deterministic system whose dynamics can be described as

$$x_{k+1} = f(x_k, v_k), \tag{1}$$

where $x \in \mathbb{R}^n$ and $v \in \mathbb{R}^m$ and $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ is a nonlinear mapping of the states and inputs to the successor state.

This work is supported by the EPSRC grant EP/C014006/1 and EC project TREN/07/FP6AE/S07.71574/037180 IFLY.

E. Siva, J. Maciejowski and N. Kantas are at Cambridge University Engineering Department es403,jmm,nk234@eng.cam.ac.uk and P. Goulart is at Imperial College, London p.goulart@imperial.ac.uk

The inputs and state must satisfy the following constraints:

$$v_k \in \mathbb{U} \subset \mathbb{R}^m \tag{2}$$
$$x_k \in \mathbb{X} \subset \mathbb{R}^n. \tag{3}$$

where $\mathbb{U}$ is compact and $\mathbb{X}$ is closed. The control objective is to steer the state to the origin. Given the difficulty in obtaining a closed form solution for stabilising controllers for nonlinear systems with nonconvex constraint sets, we consider randomised sampling of control inputs. We seek to obtain control laws which give rise to almost sure asymptotic stability [3] of the system in (1) as defined below:

Definition 2.1: Asymptotic Stability

1) The origin is stable with probability one if and only if, for any $\rho > 0$, $\epsilon > 0$, there is a $\delta > 0$ such that,
$$Pr(\sup_k \|x_k\| \geq \epsilon) \leq \rho$$
for all $\|x_0\| \leq \delta$.
2) The origin is asymptotically stable with probability 1 if and only if it is stable w.p.1., and $x_n \to 0$ w.p.1 for all x_0 in some neighbourhood of the origin.

At time k, a measurement of the current state x_k is made, and control inputs v_k are sampled according to some distribution $\pi_k(v_k|x_k)$ dependent on x_k. The control law v_k is applied, giving rise to a successor state generated by (1). The joint process ϕ_k evolves according to

$$\Phi_{k+1} \sim p_{k+1}(\cdot|\Phi_k),$$

with the conditional distribution p_{k+1} given by [1]

$$p_{k+1}(\phi_{k+1}|\phi_k) = \delta(x_{k+1} - f(x_k, v_k))\pi_{k+1}(v_{k+1}|x_{k+1}). \tag{4}$$

The origin is asymptotically stable according to Definition 2.1 if the conditions of the following theorem are satisfied:

Theorem 2.1: Assume the following conditions hold:

(i.) There is a function $V : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ continuous at the origin with $V(0,0) = 0$, and a $\mathcal{K}$-function α, such that for all $x \in \mathbb{R}^n, v \in \mathbb{R}^m$,
$$V(x, v) \geq \alpha(||x||) \tag{5}$$

(ii.) There is a compact set $\mathbb{Q} \subseteq \mathbb{R}^n$ that contains the origin, and given the initial condition $x_0 \in \mathbb{Q}$, realisations of the controlled system trajectory (x_k, v_k) satisfy $x_k \in \mathbb{Q}$ with probability 1.

(iii.) There exists a $\mathcal{K}$-function $\eta(\cdot)$ such that in $\mathbb{Q}$, along the controlled system trajectory
$$\mathbb{E}_{p_{k+1}}[V(\Phi_{k+1})|\Phi_k = (x, v)] - V(x, v) \leq -\eta(\|x, v\|) \tag{6}$$

(iv.) For any $\lambda > 0$, there exists a $\delta > 0$ such that $\mathbb{E}_{\pi_0}[V(x_0, v_0)] < \lambda$ for all $||x_0|| \leq \delta$.

[1] Although the state transition from x_k to x_{k+1} is deterministic, we shall treat x_{k+1} as an extreme case of a random variable drawn from a delta function distribution as this is more convenient for our analysis.

Then, the origin is asymptotically stabilising for all $x_0 \in \mathbb{Q}$ with probability 1.

Proof:

- Stability of the origin follows from Theorem 1, [3].
- Convergence: In view of Theorem 1, [4, Ch.6, p.71], we have $\eta(\|x_k, v_k\|) \to 0$ with probability 1. Since $\eta(\cdot)$ is a class $\mathcal{K}$-function, this implies $x, v \to 0$ with probability 1.

■

In the next sections we define a cost function V and input sampling distributions π_k, using a combination of MPC and MCMC to ensure V is a stochastic Lyapunov function, and the conditions of Theorem 2.1 are satisfied.

III. MPC and MCMC Optimisation

A. MPC

We consider controlling the system in (1) via the Model Predictive Control (MPC) approach. The MPC formalism requires an explicit model of the process. The current control action is obtained via optimisation of an open-loop objective over a finite prediction horizon, which is solved in a receding horizon manner. The solution is performed online, unlike other control techniques where precomputed control laws are applied.

The prediction horizon is of length N. The state prediction $x_{k+i|k}$ is generated by the model in (7), where $v_{k+i|k}$ is the predicted open loop input:

$$x_{k+i+1|k} := f(x_{k+i|k}, v_{k+i|k}), \tag{7}$$
$$\text{for } i = 0, \ldots, N-1$$
$$x_{k|k} := x_k.$$

For the state x_k at time k, the cost is defined as:

$$V(x_k, u_k) := \sum_{j=k}^{k+N-1} L(x_{j|k}, v_{j|k}) + F(x_{k+N|k}) \tag{8}$$

with stage cost $L : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}, L(x, u) > l(\|x, u\|), (x, u) \neq (0, 0)$, where $l(\cdot)$ is a $\mathcal{K}$-function and $L(0, 0) = l(\|0, 0\|) = 0$. The terminal cost is $F : \mathbb{R}^n \to \mathbb{R}^+$. .A terminal constraint is enforced:

$$x_{k+N|k} \in X_f \subset \mathbb{X} \tag{9}$$

We require $f(\cdot), L(\cdot, \cdot)$ and $F(\cdot)$ to be continuous, $\mathbb{U}$ compact and $\mathbb{X}$ and X_f to be closed to ensure that a minimum for $V(x_k, \cdot)$ exists [8]. The finite input sequence of predicted future controls is denoted:

$$u_k := \{v_{k|k}, v_{k+1|k}, \ldots, v_{k+N-1|k}\}.$$

The feasible set of input sequences satisfying the control, state and terminal constraints is defined as $\mathbb{U}(x_k)$, and in general depends on the initial state x_k. The optimisation problem at each time instant k is to minimise $V(x_k, \cdot)$, yielding the value function

$$V^*(x_k) := \min_{u_k \in U(x_k)} V(x_k, u_k) \tag{10}$$

by obtaining the optimal sequence of inputs u_k^* from the feasible set $U(x_k)$.

$$\begin{aligned} u_k^* &= \operatorname*{argmin}_u V(x_k, u_k) \\ &=: \{v_{k|k}^*, \dots, v_{k+N-1|k}^*\}. \end{aligned}$$

The MPC control law applied to the state x_k at time k is the first input $v_{k|k}$ of the actual sequence u_k obtained by the optimisation, $v_k = v_{k|k}$, giving the closed loop system

$$x_{k+1} = f(x_k, v_k). \tag{11}$$

B. MCMC Optimisation

Here we describe the randomised sampling technique for minimising the objective function $V(x_k, \cdot)$ at each time k. A key property of the approximate optimum obtained this way is then derived which ensures the cost is a stochastic Lyapunov function. The subscript k denoting system time is dropped from the notation in this section for clarity. A suitable transformation of variables is applied to convert the original minimisation problem to that of maximisation. An example would be $g := -V + \sup_{u,x} V(x,u)$. Note that g is a function of both the state and input sequence x and u, but the optimisation variable of interest is u. At each time, we would like to sample control input sequences according to a distribution whose modes coincide with the maximisers of g, and which enables satisfaction of the latter two requirements of Theorem 2.1. We propose sampling inputs according to a distribution π proportional to the objective function raised to a positive exponent $\bar{J}$:

$$\pi(u|x) \propto g^{\bar{J}}(x,u).$$

The higher the value of $\bar{J}$, the more concentrated about the optimum values of $g(x,\cdot)$ the distribution π becomes. Sampling from this distribution is enabled by constructing a suitable Markov chain in u, with stationary distribution π. This is achieved by the following simulated annealing algorithm. We begin with a proposal distribution q, which can be chosen freely by the user but is required to be supported on the whole space. We also require a nondecreasing sequence of inverse temperatures $\{J_n\}$ with $J_n \leq \bar{J}$, known as a cooling schedule, with n denoting the chain iteration index.

Algorithm 3.1: Simulated annealing with finite time stopping
Initialise: At $n = 0$

1) Select $q_0(\cdot)$. Initialise $J = J_0$.
2) Extract a sample $U_0 \sim q_0$. Set $U = U_0$.

Repeat:

1) Draw a new input sample sequence $\tilde{U} \sim q(\cdot|U)$;
2) Record $\tilde{U}$ as a proposed new state of the Markov chain;
3) Evaluate the acceptance probability

$$\rho_J = \min\left[1, \frac{g^J(x,\tilde{U})q(U|\tilde{U})}{g^J(x,U)q(\tilde{U}|U)}\right].$$

Accept the proposal $\tilde{U}$ with probability ρ_J and set $U = \tilde{U}$. Otherwise, with probability $1-\rho_J$, leave U unchanged;

4) If $J < \bar{J}$, set J to J_{n+1}. Else keep J fixed at $\bar{J}$;

Repeat $1-4$ until convergence of the chain to the stationary distribution π is achieved. Discard the initial samples obtained with $J < \bar{J}$ and select the final sample $\hat{u}$ of those samples remaining obtained with $J = \bar{J}$.

Constraints are handled by rejecting infeasible points. Other options include appending the cost with a parameter representing a reward for constraint satisfaction. Inputs $\hat{u}$ chosen according to Algorithm 3.1 are distributed approximately according to the stationary distribution of the Markov chain in u with the 'inverse temperature' [10] capped at $\bar{J}$, given by:

$$\hat{u} \sim \pi = \frac{g^{\bar{J}}(x,u)}{\int g^{\bar{J}}(x,u)\mathrm{d}u}$$

We consider now a feature of the approximate optimum obtained by implementation of Algorithm 3.1 which is necessary for implementing the MPC scheme presented later. The closeness of the expected value of the approximate optimum to the global optimum of the cost function is determined by the choice of capping parameter $\bar{J}$. The expected value of the approximate optimum is:

$$\begin{aligned} \mathbb{E}_\pi[g(x,U)] &= \int g(x,u)\pi(u)\mathrm{d}u \\ &= \frac{\int g^{\bar{J}+1}(x,u)\mathrm{d}u}{\int g^{\bar{J}}(x,u)\mathrm{d}u}, \end{aligned}$$

since

$$\pi(u|x) = \frac{g^{\bar{J}}(x,u)}{\int g^{\bar{J}}(x,u)\mathrm{d}u}.$$

Writing $\int g^{\bar{J}}(x,u)\mathrm{d}u$ as $\|g^{\bar{J}}(x,\cdot)\|_1$, since $g(x,u) \geq 0$, the expected difference between the global and approximate optimum is given by

$$\begin{aligned} &\mathbb{E}_\pi[\sup_u g(x,u) - g(x,u)] = \\ &\qquad \sup_u g(x,u) - \frac{\|g^{\bar{J}+1}(x,\cdot)\|_1}{\|g^{\bar{J}}(x,\cdot)\|_1} \end{aligned} \tag{12}$$

The following result shows that this expected difference can be made arbitrarily small.

Lemma 3.1: For any π-measurable function $\tilde{g}$, assume that $\tilde{g} : \mathbb{R}^m \to \mathbb{R}$ is such that $\tilde{g}(u) \geq 0$ for all u and $\operatorname{ess\,sup}_u \tilde{g}(u) = \sup_u \tilde{g}(u) =: \tilde{g}^*$. For any $\epsilon > 0$, there exists a finite M such that for all $n > M$,

$$\tilde{g}^* - \frac{\|\tilde{g}^{n+1}\|_1}{\|\tilde{g}^n\|_1} \leq \epsilon.$$

Proof:

$$\|\tilde{g}^n\|_1 = \|\tilde{g}^{\frac{n+1}{2}} \tilde{g}^{\frac{n-1}{2}}\|_1$$

From Holder's inequality we have

$$\begin{aligned}
\|\tilde{g}^{\frac{n+1}{2}}\tilde{g}^{\frac{n-1}{2}}\|_1 &\leq \|\tilde{g}^{\frac{n+1}{2}}\|_2\|\tilde{g}^{\frac{n-1}{2}}\|_2 \\
\|\tilde{g}^{\frac{n+1}{2}}\tilde{g}^{\frac{n-1}{2}}\|_1^2 &\leq \|\tilde{g}^{\frac{n+1}{2}}\|_2^2\|\tilde{g}^{\frac{n-1}{2}}\|_2^2 \\
&= \|\tilde{g}^{n+1}\|_1\|\tilde{g}^{n-1}\|_1 \\
\|\tilde{g}^{n}\|_1^2 &\leq \|\tilde{g}^{n+1}\|_1\|\tilde{g}^{n-1}\|_1 \\
\frac{\|\tilde{g}^{n+1}\|_1}{\|\tilde{g}^{n}\|_1} &\geq \frac{\|\tilde{g}^{n}\|_1}{\|\tilde{g}^{n-1}\|_1} \qquad (13)
\end{aligned}$$

From [11], Chapter 6, p.71, we have

$$\lim_{n\to\infty}\frac{\|\tilde{g}^{n+1}\|_1}{\|\tilde{g}^{n}\|_1} = \sup\tilde{g} = \tilde{g}^*$$

The sequence

$$\left\{\sup_u \tilde{g}(u) - \frac{\|\tilde{g}^{n+1}(u)\|_1}{\|\tilde{g}^{n}(u)\|_1}\right\}_{n\geq 0}$$

is thus monotonically non-increasing for all positive n, and converges to 0. It follows that for any $\epsilon > 0$, there exists an M such that for all $n > M$,

$$\tilde{g}^* - \frac{\|\tilde{g}^{n+1}\|_1}{\|\tilde{g}^{n}\|_1} \leq \epsilon.$$

■

If we make the transformation $g := V_{sup} - V$, where $V_{sup} := \sup_{x,u} V$, we have

$$g^* := \sup_u g = V_{sup} - \inf_u V \qquad (14)$$

From Lemma 3.1, we have

$$\mathbb{E}_\pi[g^* - g] \leq \epsilon$$

Combining with (14) gives

$$\begin{aligned}
\mathbb{E}_\pi[V_{sup} - \inf_u V - (V_{sup} - V(x,U))] &\leq \epsilon \\
\mathbb{E}_\pi[V(x,U) - V(x,u^*)] &\leq \epsilon \qquad (15)
\end{aligned}$$

It follows that the expected difference between the global optimum of $V(x_k,\cdot)$ and the expected cost can be made arbitrarily small by suitable choice of stopping inverse temperature at each time step. This feature is required for implementation of the proposed MPC strategy detailed next, and is also necessary for the stability analysis in Section IV. As the stopping temperature need not be the same at each time step k, we will denote it $\bar{J}(k)$ henceforth.

C. MCMC with MPC

We now detail the overall receding horizon MPC scheme with the MCMC optimisation. The control input sequence $u_k(x_k)$ obtained by the simulated annealing optimisation at each time step k with state x_k must satisfy the state and input constraints. Additionally, we place a requirement on the solution quality at each time step in terms of the expected difference between the global optimum and approximate solution. The requirement is that at each time step, the expected difference between the global optimum and approximate solution is no greater than some fraction of the first stage cost at the previous timestep. This fraction is determined by the $\mathcal{K}$-function $l(\cdot)$ that bounds this stage cost from below, as defined in Section III. We present now the algorithm:

Algorithm 3.2: MPC with Simulated Annealing

1) At time $k = 0$, state x_0, find a control sequence $u_0 = \{v_{0|0}, v_{1|0}, \dots, v_{N-1|0}\}$ which satisfies (2), (3), (7) and (9). Set $v_0 = v_{0|0}$.
2) At system time k, measure state x_k, choose an inverse stopping temperature $\bar{J}(k)$ large enough which satisfies

$$\begin{aligned}
&\mathbb{E}_{\pi_k}[V(x_k,U_k) - V(x_k,u_k^*)] \leq \\
&\qquad L(x_{k-1}, v_{k-1|k-1}) - l(\|x_{k-1}, v_{k-1|k-1}\|) \qquad (16)
\end{aligned}$$

   and implement Algorithm 3.1 to obtain a control sequence
   $u_k = \{v_{k|k}, v_{k+1|k}, \dots, v_{k+N-1|k}\} = \hat{u}$ that satisfies (2), (3), (7) and (9). Set $v_k = v_{k|k}$.

From Lemma 3.1, setting the right hand side of (16) to be the ϵ term in (15), it follows that a value for $\bar{J}(k)$ exists at each timestep which satisfies the stipulated requirement on the solution quality. The bounds on solution quality obtained for simulated annealing with finite-time stopping in [5] provide a useful starting point in determining $\bar{J}(k)$. The stabilising properties of the resulting suboptimal MPC law are explored in Section IV.

Remark 3.1: Care must be taken to distinguish between the state of Markov chain generated by Algorithm 3.1 and the controlled system chain state, which evolves within a different time frame; at every time instant k, with respect to the system time reference, a Markov chain is constructed whose state evolves with iteration n, (with the system time still fixed at k) according to an inhomogeneous Markov Chain transition kernel [10] generated by Algorithm 3.1. Note we make the key assumption in the closed loop stability analysis that the extractions from the chain are generated from the stationary distribution. Criteria for diagnosing convergence to stationarity of the chain are presented in [10]. In [2], [9] it is shown that under certain conditions, including employing a logarithmic cooling schedule for incrementing $\bar{J}$, in the limit with infinite $\bar{J}$, the distribution of the Markov chain converges to a uniform distribution over the set of minimisers of $V(x,\cdot)$. In practice only a finite value of $\bar{J}$ is required, and in the next section it is shown how closed loop stability is achieved with $\bar{J}$ chosen to meet the requirements of Algorithm 3.2.

IV. Stability in MPC with Simulated Annealing

We seek to establish almost sure stability [3] of the closed loop system in (11). Before proceeding with establishing the stabilising properties of the control law that arise by implementation of Algorithm 3.2, we first require some ancillary results and assumptions. Appropriate conditions on the terminal set X_f and terminal cost $F(\cdot)$ need to be determined. We detail them next:

Assumptions 4.1:

(i) The terminal set $X_f \subset \mathbb{X}$, X_f is closed, $0 \in X_f$.

(ii) There exists a controller $u_f(x) \in \mathbb{U}, \ \forall x \in X_f$.

(iii) $f(x, u_f(x)) \in X_f$, $\forall x \in X_f (X_f$ is positively invariant under $u_f(\cdot))$.

(iv) $F(f(x, u_f(x))) + L(x, u_f(x)) - F(x) \leq 0, \forall x_k \in X_f$.

(v) There exists a $\mathcal{K}$-function $\sigma(\cdot)$, such that $\{x_0, u_0^*\}$ satisfies

$$||u_0^*|| \leq \sigma(||x_0||) \tag{17}$$

The first four conditions in Assumption 4.1 have been identified by [8] as key ingredients for closed loop stability when employing their strategy. As in [8], we require the existence of a terminal controller u_f in the terminal set X_f satisfying the properties in Assumption 4.1. The control law that is actually applied when the state is in the terminal set is obtained from the optimisation.

The third condition in Theorem 2.1 for almost sure stability concerns the initial expected approximate value function and initial state. The proof of Theorem 4.1, showing that this condition is satisfied on carrying out Algorithm 3.2, requires the following result about the initial state and initial expectation of the approximate solution:

Lemma 4.1: For any $\lambda > 0$, given that V is continuous at the origin and $V(0,0) = 0$, there exists a J and $\delta > 0$ such that

$$\mathbb{E}_{\pi_0}[V(x_0, U_0)] < \lambda - V(x_0, u_0^*) \tag{18}$$

for all $||x_0|| \leq \delta$.
Proof: For all $r \geq 0, n \geq 1,$ define $B_r := \{x \in \mathbb{R}^n : ||x|| \leq r\}$. As V is continuous at the origin, with $V(0,0) = 0$, there exists a constant $r_1 > 0$ and a $\mathcal{K}$-function $\beta(\cdot)$ such that

$$V(x, u) \leq \beta(||(x, u)||) \ \forall x \in B_{r_1}.$$

For any $\lambda > 0$, there exists a $\delta > 0$ such that $\delta < r_1$ and $\sigma(\delta) \leq r_1$.
Let $||x_0|| < \delta$. Now $||u_0^*|| < \sigma(\delta)$ and

$$V(x_0, u_0^*) \leq \beta(||x_0, u_0^*||) \leq \beta(\delta + \sigma(\delta)).$$

As $\delta \to 0, \sigma(\delta) \to 0$, and therefore $\delta + \sigma(\delta) \to 0$ and $\beta(\delta + \sigma(\delta)) \to 0$. Given $\lambda > 0$, a $\delta > 0$ satisfying $\lambda - V(x_0, u_0^*) > 0$ exists. From equation (12) and Lemma 3.1 it follows that there exists a J such that (18) holds. ∎

We are now ready to establish the convergence properties of the control law that arise by implementation of Algorithm 3.2.

Theorem 4.1: Let $\mathbb{Q}$ represent the set of states for which there exists a control sequence that satisfies (2), (3), (7) and (9) and (16). Given satisfaction of Assumptions 4.1, the MPC law arising from implementation of Algorithm 3.2 is asymptotically stabilising with region of attraction $\mathbb{Q}$.

Proof: The algorithm ensures that along trajectories of the controlled system we have $x_k \in \mathbb{Q}$, satisfying condition (ii) of Theorem 2.1 by assumption, and

$$\mathbb{E}_{\pi_{k+1}}[V(x_{k+1}, U_{k+1}) - V(x_{k+1}, u_{k+1}^*)] \leq L(x_k, v_{k|k}) - l(\|x_k, v_{k|k}\|). \tag{19}$$

The following analysis shows the equivalence of this relation and the stochastic Lyapunov condition (iii) in Theorem 2.1 . The model predictive controller $v_k = v_{k|k}$ steers x_k to x_{k+1}.

$$\begin{aligned} V(x_{k+1}, \{v_{k+1|k}, \ldots, u_f(x_{k+N|k})\}) = \\ V(x_k, \{v_{k|k}, v_{k+1|k}, \ldots v_{k+N-1|k}\}) - L(x_k, v_{k|k}) \\ - F(x_{k+N|k}) + L(x_{k+N|k}, u_f(x_{k+N|k})) \\ + F(f(x_{k+N|k}, u_f(x_{k+N|k}))) \end{aligned} \tag{20}$$

The sum of the last three terms is less than or equal to zero, from Assumption 4.1(iv), giving

$$\begin{aligned} L(x_k, v_{k|k}) \leq V(x_k, u_k) \\ - V(x_{k+1}, \{v_{k+1|k}, \ldots, u_f(x_{k+N|k})\}) \end{aligned} \tag{21}$$

Substitution of (21) into (19) yields

$$\begin{aligned} \mathbb{E}_{\pi_{k+1}}[V(x_{k+1}, U_{k+1}) - V(x_{k+1}, u_{k+1}^*)] \\ \leq V(x_k, \{v_{k|k}, v_{k+1|k}, \ldots v_{k+N-1|k}\}) \\ - V(x_{k+1}, \{v_{k+1|k}, \ldots, u_f(x_{k+N|k})\}) - l(\|x_k, v_{k|k}\|) \end{aligned} \tag{22}$$

The cost associated with $\{v_{k+1|k}, \ldots, u_f(x_{k+N|k})\}$ and state x_{k+1} is an upper bound for the cost yielded on application of the optimal sequence of inputs u_{k+1}^*.

$$V(x_{k+1}, \{v_{k+1|k}, \ldots, u_f(x_{k+N|k})\}) \geq V(x_{k+1}, u_{k+1}^*)$$

It follows that

$$\begin{aligned} \mathbb{E}_{\pi_{k+1}}[V(x_{k+1}, U_{k+1}) - V(x_{k+1}, u_{k+1}^*)] \\ \leq V(x_k, u_k) - V(x_{k+1}, u_{k+1}^*) - l(\|x_k, v_{k|k}\|) \end{aligned} \tag{23}$$

Since $V(x_{k+1}, u_{k+1}^*)$ is not a random variable,

$$\mathbb{E}_{\pi_{k+1}}[V(x_{k+1}, U_{k+1})] \leq V(x_k, u_k) - l(\|x_k, v_{k|k}\|).$$

Noting also that

$$\begin{aligned} \mathbb{E}_{\pi_{k+1}}[V(x_{k+1}, U_{k+1})] \\ = \mathbb{E}_{p_{k+1}}[V(X_{k+1}, U_{k+1}) | X_k = x_k, U_k = u_k] \end{aligned} \tag{24}$$

with p_{k+1} defined in (4), we see that condition (iii) of Theorem 2.1 is satisfied. Almost sure convergence of the state to the origin follows. From Lemma 4.1 it follows that for any $\lambda > 0$ there exists a J and $\delta > 0$ such that

$$\mathbb{E}_{\pi_0}[V(x_0, U_0)] < \lambda - V(x_0, u_0^*)$$

for all $||x_0|| \leq \delta$. Condition (iv) of Theorem 2.1 is therefore satisfied. Almost sure asymptotic stability follows from Theorem 2.1. ∎

V. Simulation Example

We now present an example to illustrate application of simulated annealing to nonlinear MPC. It is shown that an inverse stopping temperature can be determined empirically by simulation. The objective is to drive a constant speed particle in $2D$ to a target set centred at O_T whilst avoiding an obstacle of radius R centred at $O_X = [O_x \ O_y]$. The control input v_k is the change in particle heading, $\theta_k - \theta_{k-1}$. The state at time k is $x_k = [x_k \; y_k \; \theta_k]'$. The system is nonlinear in state and input and described by:

$$x_{k+1} = x_k + \begin{pmatrix} V\cos(\theta_k + u_k) \\ V\sin(\theta_k + u_k) \\ v_k \end{pmatrix}$$

where $V = 5$. The stage cost is given by:

$$L(x, u) = \left\| [x - [O_T \; 0]'] \right\|_2 \tag{25}$$

The horizon length $N = 5$, and the constraint sets for the controls and states are $\mathbb{U} = \{u \in \mathbb{R} : 0 \leq u \leq 2\pi\}$ and $\mathbb{X} = \{\mathbf{x} \in \mathbb{R}^n : (x - O_x)^2 + (y - O_y)^2 > R^2\}$.

The initial state x_0 is $[10 \; 15]'$ and the target set is centred at $O_T = [50 \; 50]'$.

The problem is solved in a receding horizon manner. We implement Algorithm 3.2 without identifying a distinct $\bar{J}(k)$ at each time k, but instead keeping $\bar{J}$ fixed for the entire sequence of optimisations. Control inputs yielding infeasible solutions are rejected. For the cooling schedule, we increment J_n by 1 every 25 iterations until $J = \bar{J}$. For the proposal distribution, we use a Gaussian random walk centred on the current value of the chain. Figure 1 depicts sample trajectories obtained with variable obstacle widths $R = \{5, 10, 15\}$ and fixed horizon length $N = 5$. It can

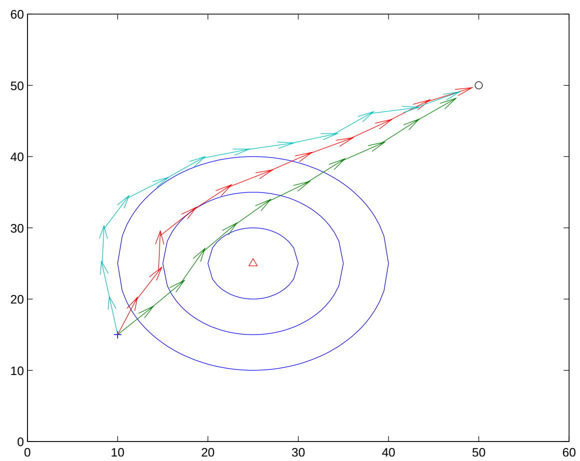

Fig. 1. Representative particle trajectories obtained with obstacles of radii 5, 10 and 15, and inverse stopping temperature $\bar{J} = 40$.

be seen that in each instance, the target set is reached. The total number of iterations per stage required for the simulations, n, were 2195, 3897 and 5755 respectively, and took 3.3, 5.7 and 8.1 seconds on a networked 2.4 GHz Dual Processor. A burn-in period of 500 iterations was used, and convergence of the chain was judged to be achieved after a certain number of states (1000) were accepted. As the obstacle radius is increased, a greater extent of resampling is required as the infeasible region increases, resulting in an increased number of iterations required. As the value of $\bar{J}$ was fixed for the entirety of each simulation, and not tuned for each optimisation at k, there is scope for a reduction in the number of iterations required.

VI. Concluding Remarks

In this paper we consider application of simulated annealing to the MPC optimisation problems that arise for nonlinear, deterministic discrete time systems. We present a suboptimal MPC scheme with almost sure stabilising properties. This is a useful extension to the deterministic stability result in [8], [12], for cases where near optimality is a solution requirement, and the optimisation method is stochastic. As in [8], we are required to enforce a terminal set constraint with similar conditions on the terminal set and terminal controller which is assumed to exist. As a monotonic cost reduction constraint cannot be imposed when a randomised optimiser is used, a condition is placed on the solution quality in terms of its expected value. Although we do not explictly state how to determine a stopping temperature at each time step to achieve this, we prove that such a value exists, and it is shown by simulation that it can be empirically determined for our example.

We are currently extending this work to establish almost sure stability of MPC schemes for stochastic nonlinear systems with an expected objective minimisation formulation.

References

[1] A. Doucet, S. J. Godsill, and C. Robert, "Marginal maximum a posteriori estimation using Markov Chain Monte Carlo," *Statisitcs and Computing*, vol. 12, pp. 77–84, 2002.

[2] Geman and Geman, "Stochastic relaxation, gibbs distribution and the bayesian restoration of images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 6, pp. 721–741, 1984.

[3] H. Kushner, "On the stability of stochastic dynamical systems," in *Proceedings of the National Academy of Sciences*, 1965.

[4] ——, *Introduction to Stochastic Control*. Holt, Rinehart and Winston, 1971.

[5] A. Lecchini-Visintini, J. Lygeros, and J. Maciejowki, "Simulated annealing: Rigorous finite-time guarantees for optimisation on continous domains," in *Advances in Neural Information Processing 20*, 2007.

[6] J. M. Maciejowki, A. Lecchini-Visintini, and J. Lygeros, "NMPC for complex stochastic systems using Markov Chain Monte Carlo," in *Assessment and Future Directions of Nonlinear Model Predictive Control*. Springer Berlin / Heidelberg, 2007.

[7] J. M. Maciejowski, *Predictive Control with Constraints*. Prentice Hall, 2001.

[8] D. Mayne, J. Rawlings, C. Rao, and P. Scokaert, "Constrained model predictive control:stability and optimality," *Automatica*, vol. 36, pp. 789–814, 2000.

[9] P. Muller, B. Sanso, and M. D. Iorio, "Optimal Bayesian design by inhomogeneous Markov Chain simulation," *Journal of the American Statistical Association*, vol. 99, pp. 788–798, 2004.

[10] C. Robert and G. Casella, *Monte Carlo Statistical Methods*. Springer, 2004.

[11] W. Rudin, *Real and Complex Analysis*. McGraw-Hill, 1987.

[12] P. Scokaert, D. Mayne, and J. Rawlings, "Suboptimal Model Predictive Control (Feasibility Implies Stability)," *IEEE Transactions on Automatic Control*, vol. 44, pp. 648–654, 1999.

[13] S.P.Meyn, "Ergodic theorems for discrete time stochastic systems using a stochastic Lyapunov function," *SIAM Journal of Control and Optimisation*, vol. 27, pp. 1409–1439, 1989.

[14] M. Vidyasagar, "Randomised algorithms for robust controller synthesis using statistical learning theory," *Automatica*, vol. 37, pp. 1515–1528, 2001.