

An Information Theoretic View of Stochastic Resonance

Venkat Anantharam

EECS Department
University of California Berkeley
Berkeley, CA 94720, U.S.A.
ananth@eecs.berkeley.edu

Vivek S. Borkar

School of Technology and Computer Science
Tata Institute of Fundamental Research
Homi Bhabha Rd., Mumbai 400005, INDIA
borkar@tcs.tifr.res.in

Abstract—We are motivated by the widely studied phenomenon called *stochastic resonance*, namely that in several sensing systems, both natural and engineered, the introduction of noise can enhance the ability of the system to perceive signals in the environment. We adopt an information theoretic viewpoint, evaluating the quality of sensing via the mutual information rate between the environmental signal and the observations. Viewing what would be considered noise in stochastic resonance as an open loop *control* and using Markov decision theory techniques, we discuss the problem of optimal choice of this control in order to maximize this mutual information rate. We determine the corresponding dynamic programming recursion: it involves the conditional law of certain conditional laws associated to the dynamics. We prove that the optimal control may be chosen as a *deterministic* function of this law of laws.

I. INTRODUCTION

To get an intuitive understanding of the phenomenon of stochastic resonance it is useful to revisit the original example of Benzi et al. [1]. Consider the Langevin equation

$$dx = [x(a - x^2)]dt + \epsilon dW .$$

This is a stochastic differential equation (SDE), driven by the Wiener process W . Assume that $a > 0$. The deterministic differential equation corresponding to $\epsilon = 0$ has stable solutions at $\pm\sqrt{a}$ and an unstable solution at 0. The SDE itself has two basins of attraction of equal depth centered at $\pm\sqrt{a}$.

Following [1], consider the Langevin equation with small periodic forcing (i.e. A is small)

$$dx = [x(a - x^2) + A \cos(\Omega t)]dt + \epsilon dW .$$

Heuristically, this SDE can be thought of as having two periodically varying basins of attraction around roughly $\pm\sqrt{a}$, with the depths of the basins also periodically varying. Now, if the variance of the driving noise is too small relative to the period of the forcing, the effect of the forcing will not be apparent in the solution to the SDE. If the variance of the driving noise is too high, then again the effect of the forcing will not be apparent in the solution to the SDE, because the basins of attraction themselves are washed out. However, for intermediate ranges of the driving noise variance, an interesting phenomenon occurs. Since it is easier for the noise to drive the state out of a more shallow well than out of a deeper well, the solution of the SDE tends to spend more time

in the deeper basin of attraction, and so the periodic forcing makes itself apparent in this solution. For the mathematical details that formalize this heuristic, see [1] and [5].

The possibility of using noise to enhance the ability of sensing systems to detect signals in the environment is also apparent in more simple contexts. Consider, for instance, a scalar binary hypothesis testing problem with real valued observations in a Gaussian noise environment, where the threshold based decision rule is fixed *a priori*. The signal may be 0 or -1 , and the decision rule implements the indicator map $1(x > \theta)$, where the threshold $\theta > 0$ is fixed and x denotes the real observation. Assuming that the inherent ambient Gaussian noise itself has zero mean and small variance, the strategy of adding an additional zero mean Gaussian noise to the signal before thresholding exhibits a stochastic resonance effect. If the added noise variance is very small there the probability of exceeding the threshold is too small to allow one to discriminate between the hypotheses, while if the added noise variance is too large the probability of exceeding the threshold is roughly the same, roughly $\frac{1}{2}$, for both hypotheses. However, for intermediate values of the added noise variance the signal manifests itself better in the observation. Related examples are discussed in e.g. [7], [9].

The literature on stochastic resonance is enormous and fascinating, for a survey see [6]. It has been proposed as an explanation for the ice ages (the external signal being the change in the heat flux from the sun due to the periodic change in eccentricity of the orbit of the earth) [2] and is believed to underly the uncanny ability of certain animals, e.g. crayfish, to detect extremely subdued signals in their environment [10]. In engineered systems, such as sensor arrays, a question that presents itself naturally is how best to exploit this phenomenon to enhance the ability of sensing systems to perceive a signal in the environment. Here we are motivated to think of what serves as noise or added randomization in the above examples as a *control*. In this paper we address this engineering question in a framework combining information theoretic and control theoretic ideas. Specifically, we model the engineered sensing system as a Markovian system driven by a signal in the environment, which is desired to be detected, and also driven by an additional input which is for us to choose. We study the open loop control problem of how best to choose the

additional input to maximize a mutual information rate between the signal in the environment and an observation process derived from the Markovian sensing system. The details of our problem formulation and the statement and sketches of the proofs of our results are in the subsequent sections.

II. NOTATION AND PROBLEM SETUP

(X_n) , called the *state* process, (Y_n) , called the *output* process, (S_n) , called the *signal*, and (U_n) , called the *control* process take values in finite sets \mathcal{X} , \mathcal{Y} , \mathcal{S} , and \mathcal{U} resp.

We use mnemonic notation, writing, e.g., $p(x_n)$ for $P(X_n = x_n)$ etc. The symbol $p(\cdot)$ is reserved for probabilities; kernels that are fixed throughout the discussion have their own special notation.

The system dynamics is the following :

$$p(x_{n+1}, y_{n+1}, s_{n+1}, u_{n+1} \mid x_0^n, y_0^n, s_0^n, u_0^n) =$$

$$\tau(x_{n+1}, y_{n+1} \mid x_n, s_n, u_n) \sigma(s_{n+1}) \alpha_{n+1}(u_{n+1} \mid u_0^n).$$

$(\alpha_n(u_n \mid u_0^{n-1}))$ is called the *control strategy*. An initial distribution $\iota(x_0, y_0)$ is given. Setting $p(x_0, y_0, s_0, u_0) = \iota(x_0, y_0) \sigma(s_0) \alpha_0(u_0)$ then completely defines the model.

We assume, without loss of generality, that $\min_i \sigma(i) \geq \Delta > 0$. We shall also make the assumption:

(†) $\forall(x', y'), \tau(x', y' \mid x, s, u)$ is either $0 \forall x, s, u$, or $\geq \delta \forall x, s, u$, for a given $\delta > 0$.

We assume that the associated control-independent and signal-independent directed graph on $\mathcal{X} \times \mathcal{Y}$ is aperiodic (this will be relaxed later). Set $\hat{\delta} = \delta \Delta$.

Here are some specific points to note about the model. The signal is i.i.d. with marginal distribution $\sigma(s)$. The control is randomized, but is open loop in that the controller does not observe anything about the dynamics. The kernel $\tau(x', y' \mid x, s, u)$ defines the dynamics.

The output process is thought of as being seen by an agent who wishes to learn the signal. The control process is thought of as used by another agent who wants to influence the dynamics in order to aid the observer to learn the signal. The controller has no access to the output process seen by the observer. The observer has no access to the control process used by the controller. More elaborate formulations, for instance when the controller also has access to its own observations, can be envisioned, but appear to be much more difficult to analyze.

We consider two kinds of information theoretic measures of how well the observer can learn the signal. The simpler one is to maximize the long run average of a marginal mutual information :

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n I(S_{k-1} \wedge Y_k).$$

The more complicated one is to maximize the long run mutual information rate :

$$\liminf_{n \rightarrow \infty} \frac{1}{n} I(S_0^n \wedge Y_0^{n+1})$$

We first discuss the simpler objective function and then the more complicated one.

III. MARGINAL INFORMATION RATE MAXIMIZATION

Let us call the problem at hand problem I. We would like to reformulate this problem in the language of dynamic programming with partial observations with an expected long term average reward criterion. We first observe that

$$I(S_{n-1} \wedge Y_n) \leq I(S_{n-1} \wedge Y_n \mid U_0^{n-1}).$$

where we have used the independence of S_{n-1} and U_0^{n-1} . We may thus consider an alternative control problem which we call problem II. Here the objective is to maximize the long average :

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n I(S_{k-1} \wedge Y_k \mid U_0^{k-1}).$$

Note that the maximum achievable in problem II is at least as big as the maximum achievable in problem I.

Let Π denote the set of probability distributions on $\mathcal{X} \times \mathcal{Y} \times \mathcal{S}$ and Π_0 the subset thereof of measures of the form $\nu(x, y, s) = \eta(x, y \mid s) \sigma(s)$. For $n \geq 1$, let π_n denote the conditional distribution of (X_n, Y_n, S_{n-1}) given U_0^{n-1} , i.e.

$$\pi_n(x_n, y_n, s_{n-1}) = p(x_n, y_n, s_{n-1} \mid U_0^{n-1}).$$

This is a Π_0 -valued random variable measurable with respect to U_0^{n-1} . Let $\pi_0(x_0, y_0, s_{-1}) = \iota(x_0, y_0) \sigma(s_{-1})$. Note that, for $n \geq 1$, we can write

$$I(S_{n-1} \wedge Y_n \mid U_0^{n-1}) = E[G(\pi_n)]$$

for the numerical function $G : \Pi \mapsto \mathbf{R}_+$ where $G(\pi)$ gives the mutual information between the last two of the triple of random variables having the joint distribution π . Further, we have an update rule that gives π_{n+1} from π_n and U_n for all $n \geq 0$:

$$\pi_{n+1}(x_{n+1}, y_{n+1}, s_n) =$$

$$\sum_{x_n, y_n, s_{n-1}} \tau(x_{n+1}, y_{n+1} \mid x_n, s_n, U_n) \sigma(s_n) \pi_n(x_n, y_n, s_{n-1}).$$

The derivation of this is suppressed due to space constraints. We can inductively prove that

$$\pi_n(x, y, s) \geq \hat{\delta} > 0 \forall x, y, s, \forall n. \quad (1)$$

We now define a third problem, which we call problem III. The state space for this problem is Π_0 and the action space is \mathcal{U} . The dynamics for this problem are given by the evolution equation for (π_n) . We take $\pi_0(x_0, y_0, s_{-1}) = \iota(x_0, y_0) \sigma(s_{-1})$ as the initial state in problem III. The objective in problem III is to maximize the expected long run average reward

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n E[G(\pi_k)].$$

This maximization will be taken to be over all controls $\{U_n\}$ such that for each n , U_n depends on $\pi_m, m \leq n$, and possibly some additional independent randomization.

Comparing problems II and III, we see that every strategy in problem II maps one to one to a strategy in problem III

and vice versa, and that the objectives map to each other one to one: That II reduces to III is already seen above. In turn, if one starts with III, then the recursion for $\{\pi_n\}$ shows that $\sigma(\pi_m, m \leq n) \subset \sigma(U_m, m < n)$ for all n , thus one is equivalently prescribing $\{\alpha_n\}$.

From the theory of average cost dynamic programming (see section 5 below) we know that there is an optimal control for problem III where the control U_n is chosen deterministically as a function of π_n^1 . Since the initial condition is deterministic (π_0) it follows that the optimal control strategy is deterministic. But then the objective functions of problem I and problem II are identical under this control strategy. Thus one can achieve the optimal long term reward achievable in problem II also in problem I, by implementing the deterministic control strategy (i.e. choose the control deterministically as a function of the conditional law) that was optimal for problem III. Since we can do no better in problem I than in problem II, this deterministic control strategy is optimal for problem I.

IV. ENTROPY RATE MAXIMIZATION

We next consider the more complicated objective of maximizing the long run mutual information rate :

$$\liminf_{n \rightarrow \infty} \frac{1}{n} I(S_0^n \wedge Y_0^{n+1}) .$$

The system dynamics are as before. We call this problem Ib.

Observe that

$$I(S_0^n \wedge (Y_0^{n+1}, U_0^{n+1})) \geq I(S_0^n \wedge Y_0^{n+1}) .$$

Let problem IIb mean the problem of maximizing :

$$\liminf_{n \rightarrow \infty} \frac{1}{n} I(S_0^n \wedge (Y_0^{n+1}, U_0^{n+1})) ,$$

with the system dynamics as before.

Note that we may write

$$I(S_0^n \wedge (Y_0^{n+1}, U_0^{n+1})) =$$

$$\sum_{k=0}^n [I(S_0^k \wedge (Y_0^{k+1}, U_0^{k+1})) - I(S_0^{k-1} \wedge (Y_0^k, U_0^k))] ,$$

where the term corresponding to $k = 0$ is interpreted as $I(S_0 \wedge (Y_0^1, U_0^1))$. Thus the objective is a running sum of terms. The terms of the running sum can be manipulated into the form (the details are suppressed)

$$I(S_0^n \wedge (Y_0^{n+1}, U_0^{n+1})) - I(S_0^{n-1} \wedge (Y_0^n, U_0^n)) =$$

$$H(S_n) + H(Y_{n+1} | Y_0^n, U_0^n) - H(Y_{n+1}, S_n | Y_0^n, U_0^n, S_0^{n-1}) .$$

Problem IIb is thus the problem of maximizing

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n [H(Y_{n+1} | Y_0^n, U_0^n) - H(Y_{n+1}, S_n | Y_0^n, U_0^n, S_0^{n-1})] .$$

¹We call these ‘stationary strategies’. More generally, we may consider ‘randomized stationary strategies’ φ where φ is a deterministic function that maps π_n to a probability distribution on \mathcal{U} , the idea being that the actual control U_n gets chosen according to this (conditional) law.

Consider a $\Pi_0 \times \Pi_0$ -valued process $((\phi_n, \psi_n))$, with ϕ_n for $n \geq 1$ and ψ_n for $n \geq 2$ defined by

$$\begin{aligned} \phi_n(x_n, y_n, s_{n-1}) &= p(x_n, y_n, s_{n-1} | Y_0^{n-1}, U_0^{n-1}) \\ \psi_n(x_n, y_n, s_{n-1}) &= p(x_n, y_n, s_{n-1} | Y_0^{n-1}, U_0^{n-1}, S_0^{n-2}) , \end{aligned}$$

and let

$$\begin{aligned} \psi_1(x_1, y_1, s_0) &= p(x_1, y_1, s_0 | Y_0, U_0) \\ \phi_0(x_0, y_0, s_{-1}) &= i(x_0, y_0) \sigma(s_{-1}) \\ \psi_0(x_0, y_0, s_{-1}) &= i(x_0, y_0) \sigma(s_{-1}) . \end{aligned}$$

We can write

$$H(Y_n | Y_0^{n-1}, U_0^{n-1}) = E[G_1(\phi_n)] ,$$

for the function $G_1 : \Pi \mapsto \mathbf{R}_+$ where $G_1(\phi)$ gives the entropy of second of the three random variables having joint distribution ϕ . Similarly, for $n \geq 2$ we can write

$$H(Y_n, S_{n-1} | Y_0^{n-1}, U_0^{n-1}, S_0^{n-2}) = E[G_2(\psi_n)] ,$$

for the function $G_2 : \Pi \mapsto \mathbf{R}_+$ where $G_2(\psi)$ gives the joint entropy of the last two of the three random variables having joint distribution ψ .

We have an update rule that gives (ϕ_{n+1}, ψ_{n+1}) from (ϕ_n, ψ_n) and U_n for all $n \geq 0$.

$$\begin{aligned} \phi_{n+1}(x_{n+1}, y_{n+1}, s_n) &= \\ \frac{\sum_{x_n, s_{n-1}} \tau(x_{n+1}, y_{n+1} | x_n, s_n, U_n) \sigma(s_n) \phi_n(x_n, Y_n, s_{n-1})}{\sum_{x_n, s_{n-1}} \phi_n(x_n, Y_n, s_{n-1})} \\ \psi_{n+1}(x_{n+1}, y_{n+1}, s_n) &= \\ \frac{\sum_{x_n} \tau(x_{n+1}, y_{n+1} | x_n, s_n, U_n) \sigma(s_n) \psi_n(x_n, Y_n, S_{n-1})}{\sum_{x_n} \psi_n(x_n, Y_n, S_{n-1})} . \end{aligned}$$

Once again, one can prove inductively that

$$\phi_n(x, y, s), \psi_n(x, y, s) \geq \hat{\delta} > 0 \forall x, y, s, \forall n. \quad (2)$$

We now consider the partially observed Markov decision problem with state process $((\phi_n, \psi_n))$ and control process (U_n) , with the evolution equation as above. The objective is to maximize the expected long run average reward

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \sum_{k=2}^n E[G_1(\phi_k) - G_2(\psi_k)] .$$

Note that the processes $\{\phi_n\}, \{\psi_n\}$ are themselves only partially observed because $\{Y_n\}, \{S_n\}$ are not observed. Thus we define μ_n to be the regular conditional law of (ϕ_n, ψ_n) given U_0^n for $n \geq 0$. Write the dynamics of $\{\phi_n\}, \{\psi_n\}$ described above as

$$\begin{aligned} \phi_{n+1} &= F_1(\phi_n, Y_n, U_n), \\ \psi_{n+1} &= F_2(\psi_n, Y_n, S_{n-1}, U_n), \end{aligned}$$

for suitably defined F_1, F_2 . Then $\{\mu_n\}$ is recursively given as

$$\mu_{n+1}(d\phi, d\psi) =$$

$$\int \sum_{y,s} P(F_1(\phi', y, U_n) \in d\phi, F_2(\psi', y, s, U_n) \in d\psi) \psi'(\mathcal{X}, y, s) \mu_n(d\phi', d\psi'). \quad (3)$$

Let $\tilde{\mu}_n(d\phi), \hat{\mu}_n(d\psi) \in \Pi_0$ denote the marginals of $\mu_n, n \geq 0$. Then these are seen to be given recursively by

$$\tilde{\mu}_{n+1}(d\phi) = \int \sum_y P(F_1(\phi', y, U_n) \in d\phi) \phi'(\mathcal{X}, y) \tilde{\mu}_n(d\phi'), \quad (4)$$

and

$$\hat{\mu}_{n+1}(d\psi) = \int \sum_{y,s} P(F_2(\psi', y, s, U_n) \in d\psi) \psi'(\mathcal{X}, y, s) \hat{\mu}_n(d\psi'). \quad (5)$$

The cost in turn can be rewritten as

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \sum_{k=2}^n \int \mu_k(d\phi, d\psi) (G_1(\phi) - G_2(\psi)), \quad (6)$$

or equivalently as

$$\liminf_{n \rightarrow \infty} \left(\frac{1}{n} \sum_{k=2}^n \int \tilde{\mu}_k(d\phi) G_1(\phi) - \frac{1}{n} \sum_{k=2}^n \int \hat{\mu}_k(d\psi) G_2(\psi) \right). \quad (7)$$

Thus we may consider either the problem ‘IIIb*’ of controlling $\{\mu_n\}$ given by the dynamics (3) with cost (6), or the problem ‘IIIb’ of controlling $\{\{\tilde{\mu}_n, \hat{\mu}_n\}\}$ given by the dynamics (4), (5) with cost (7). The latter turns out to be more convenient. Comparing problems IIb and IIIb, we see that every strategy in problem IIb maps one to one to a strategy in problem IIIb and vice versa - this can be argued as for II, III.

From the results of next section, we conclude that an optimal control for problem IIIb can be found where the control is chosen deterministically as a function of this information state. It follows then that a deterministic choice of control process achieves the optimal objective in problem IIb. Since the realized objective in problem Ib is identical to that in problem IIb if one works with deterministic controls, and since the objective in problem Ib can be no bigger than that in problem IIb, we conclude that a deterministic control strategy is optimal for problem Ib.

V. DYNAMIC PROGRAMMING

We begin with some preliminaries. Consider the set $H \stackrel{def}{=} \{p = [p_1, \dots, p_M] : p_i \in [\delta, 1] \cup \{0\} \forall i, \sum_j p_j = 1\}$ and the function $f : H \rightarrow \mathcal{R}^+$ defined by $f(p) = -\sum_j p_j \log p_j$.

Lemma 1 Let X_1, X_2 be random variables taking values in $D \stackrel{def}{=} \{1, 2, \dots, M\}$ on a probability space (Ω, \mathcal{F}, P) . Let $\hat{\mathcal{F}}_1, \hat{\mathcal{F}}_2$ be sub- σ -fields of \mathcal{F} , η_i the regular conditional laws of X_i given $\hat{\mathcal{F}}_i$ for $i = 1, 2$, and $\kappa(j|i), i, j \in D$, a probability kernel on D (i.e., $\kappa(j|i) \geq 0$ with $\sum_j \kappa(j|i) = 1$). Let $\zeta_i(k) \stackrel{def}{=} \sum_j \eta_i(j) \kappa(k|j), i = 1, 2; k \in D$, and suppose $\zeta_i \in H, i = 1, 2$, a.s. Then

$$|E[f(\zeta_1)] - E[f(\zeta_2)]| \leq \log\left(\frac{1}{\delta}\right) E\left[\sum_j |\kappa(j|X_1) - \kappa(j|X_2)|\right]. \quad (8)$$

Proof: The proof is suppressed for lack of space. The most important corollary for us is for $\kappa(j|i) \stackrel{def}{=}} I\{j = i\}$, where we have

$$|E[f(\zeta_1)] - E[f(\zeta_2)]| \leq 2 \log\left(\frac{1}{\delta}\right) P(X_1 \neq X_2). \quad (9)$$

Note that, in view of (1), (2), it follows that $\{\pi_n\}, \{\phi_n\}, \{\psi_n\}$ are H -valued processes. \square

Now consider the first problem with discounted cost:

$$V_\kappa(\pi) \stackrel{def}{=} \inf E\left[\sum_m \kappa^m G(\pi_m) | \pi_0 = \pi\right], \quad (10)$$

where $0 < \kappa < 1$ is the discount factor and the infimum is over all admissible (i.e., open loop) controls $\{u_n\}$. The following is standard [8]:

Lemma 2 V_κ is the unique bounded solution to the dynamic programming equation

$$V_\kappa(\pi) = \min_u [G(\pi) + \kappa \int p(d\pi' | \pi, u) V_\kappa(\pi')], \quad (11)$$

where $p(\cdot)$ is the transition kernel of the controlled nonlinear filter $\{\pi_m\}$. Furthermore, the control choice $u_n = v(\pi_n) \forall n$, where $v(\cdot)$ is the measurable selector of the argmin of the r.h.s. of (11), is optimal. \square

We are interested in deriving the dynamic programming equations for the ergodic control problem using the ‘vanishing discount’ limit of (11). We do so by combining ideas from [1], [2]. Fix $\pi^* \in H$ and consider $\bar{V}_\kappa(\pi) \stackrel{def}{=}} V_\kappa(\pi) - V_\kappa(\pi^*)$. Letting $\{u_n^*\}$ denote the control process optimal for $\pi_0 = \pi^*$, let $\{\pi_n\}, \{\pi_n^*\}$ denote the corresponding nonlinear filters under the common control $\{u_n^*\}$ and common $\{s_n\}$, but with different initial conditions π, π^* resp. Let $\{X_m, Y_m\}, \{X_m^*, Y_m^*\}$ resp. denote the corresponding state-observation pairs. Let $\tau \stackrel{def}{=} \min\{n \geq 0 : X_n = X_n^*, Y_n = Y_n^*\}$. We may set $X_n = X_n^*, Y_n = Y_n^* \forall n \geq \tau$, coupling the chains at the ‘coupling time’ τ . Then we can show (details suppressed)

$$V_\kappa(\pi) - V_\kappa(\pi^*) \leq 4 \log\left(\frac{1}{\delta}\right) E[\tau].$$

The roles of π and π^* may be interchanged, giving

$$|V_\kappa(\pi) - V_\kappa(\pi^*)| \leq 4 \log\left(\frac{1}{\delta}\right) E[\tau] \quad (12)$$

From (12), we have

$$\bar{V}_\kappa(\pi) \leq 4 \log\left(\frac{1}{\delta}\right) \max E[\tau] < \infty,$$

where the ‘max’ is over all initial laws π and the finiteness follows easily from our ‘aperiodicity’ condition. This establishes the boundedness of $\bar{V}_\kappa, 1 > \kappa > 0$.

The equicontinuity of $\bar{V}_\kappa, 1 > \kappa > 0$, can now be established as in [1]. We sketch the argument here. Under our assumptions, we have $P(\tau > n) \leq K \zeta^n$ for some $K > 0, \zeta \in (0, 1)$ regardless of the choices of π, π^* . Replace π, π^* by a generic π_1, π_2 with the corresponding processes $\{X_m^1, Y_m^1\}, \{X_m^2, Y_m^2\}$. (Recall that the control process is kept the same for the two.) By considering π_1, π_2 to be Dirac

at i, j resp. and \mathcal{X} as a subset of \mathcal{R} under any convenient embedding, we may see, by suitably redefining K if necessary, that $P(\tau > n) \leq K\zeta^n|i-j|$. As above, one then has for more general π_1, π_2 ,

$$|V_\kappa(\pi_1) - V_\kappa(\pi_2)| \leq 4 \log\left(\frac{1}{\delta}\right) \frac{K}{1-\zeta} E[|X_0^1 - X_0^2|].$$

Taking the maximum on the right hand side over all possible joint laws of (X_0^1, X_0^2) with marginals π_1, π_2 , one has

$$|V_\kappa(\pi_1) - V_\kappa(\pi_2)| \leq 4 \log\left(\frac{1}{\delta}\right) \frac{K}{1-\zeta} \rho(\pi_1, \pi_2),$$

where $\rho(\cdot, \cdot)$ is the Vasserstein metric. This proves the equicontinuity of V w.r.t. ρ .

Now rewrite (11) as

$$\bar{V}_\kappa(\pi) = \min_u [G(\pi) - (1-\kappa)V_\kappa(\pi^*) + \kappa \int p(d\pi'|\pi, u) \bar{V}_\kappa(\pi')], \quad (13)$$

Since $(1-\kappa)V_\kappa(\pi^*)$, $\kappa \in (0, 1)$, is bounded, we may invoke the Arzela-Ascoli theorem and drop to a subsequence $\kappa_n \uparrow 1$ along which $\bar{V}_{\kappa_n} \rightarrow V$, $(1-\kappa_n)V_{\kappa_n}(\pi^*) \rightarrow \beta$ for suitable V, β . Then passing to the limit in (13) along this subsequence, we have

$$V(\pi) = \min_u [G(\pi) - \beta + \int p(d\pi'|\pi, u) V(\pi')], \quad (14)$$

which is the dynamic programming equation for ergodic control. Counterparts of Theorems 4.1-4.2 of [1] are now true for the same reasons. These are recalled below:

Theorem 1 $(V(\cdot), \beta)$ solve (14) with $\beta =$ the optimal cost, independent of the initial condition. Furthermore, a stationary policy $v(\cdot)$ (resp., a stationary randomized policy $\varphi(\cdot)$) is optimal for any initial condition if

$$v(\pi) \in (\text{resp., support}(\varphi(\pi)) \subset) \text{Argmin}\left(\int p(d\pi'|\pi, \cdot) V(\pi')\right). \quad (15)$$

Conversely, if φ is an optimal randomized stationary policy, then (15) holds a.s. with respect to the corresponding stationary distribution of $\{\pi_n\}$. \square

A limited uniqueness claim for $V(\cdot)$ is possible as in *ibid*. The arguments for removing the aperiodicity condition involve coupling to the periodic cycles and explicitly handling the phase offset. These are suppressed for lack of space.

An analogous development is possible for the second problem, with the pair $(\tilde{\mu}_n, \hat{\mu}_n)$ in place of $\{\pi_n\}$. We sketch the proof below. An important point to note is that because of the separated nature of the cost function it is not necessary to impose the consistency requirement on $\{\phi_n\}, \{\psi_n\}$ implicit in their original definition. This is not a problem: the dynamics of $\{(\tilde{\mu}_n, \hat{\mu}_n)\}$ makes sense regardless of any such condition.

As in the earlier problem, we can show by a coupling argument for the corresponding discounted problem that

$$|\bar{V}_\kappa(\tilde{\mu}, \hat{\mu})| = |V_\kappa(\tilde{\mu}, \hat{\mu}) - V_\kappa(\tilde{\mu}^*, \hat{\mu}^*)| \leq 8 \log\left(\frac{1}{\delta}\right) \max E[\tau] < \infty,$$

where the meaning of the notation should be clear from the context. As before, this also leads to

$$|\bar{V}_\kappa(\tilde{\mu}_1, \hat{\mu}_1) - \bar{V}_\kappa(\tilde{\mu}_2, \hat{\mu}_2)| \leq K [\hat{\rho}(\tilde{\mu}_1, \tilde{\mu}_2) + \hat{\rho}(\hat{\mu}_1, \hat{\mu}_2)].$$

Here $\hat{\rho}$ is the appropriate Vasserstein metric. Namely

$$\hat{\rho}(\tilde{\mu}_1, \tilde{\mu}_2) = \inf E[\rho(\zeta_1, \zeta_2)],$$

where the infimum is over all joint laws of (ζ_1, ζ_2) such that ζ_i has the law $\tilde{\mu}_i$ for $i = 1, 2$, and $\hat{\rho}(\hat{\mu}_1, \hat{\mu}_2)$ is defined similarly. The rest is as before. One then has the counterpart of Theorem 1 as follows:

Theorem 2 There exists a solution $(V(\cdot), \beta)$ of the dynamic programming equation

$$V(\tilde{\mu}, \hat{\mu}) = \min_u \left[\int \tilde{\mu}(d\phi) G_1(\phi) - \int \hat{\mu}(d\psi) G_2(\phi) - \beta + \int q(d\tilde{\mu}' d\hat{\mu}' | (\tilde{\mu}, \hat{\mu}), u) V(\tilde{\mu}', \hat{\mu}') \right], \quad (16)$$

with $q(\cdot | \cdot, \cdot)$ *def* the transition kernel of the controlled Markov process $\{(\tilde{\mu}_n, \hat{\mu}_n)\}$. Furthermore,

- $\beta =$ the optimal cost, independent of the initial condition,
- a stationary policy $v(\cdot)$ (resp., a stationary randomized policy $\varphi(\cdot)$) is optimal for any initial condition if

$$v(\tilde{\mu}, \hat{\mu}) \in (\text{resp., support}(\varphi(\tilde{\mu}, \hat{\mu})) \subset)$$

$$\text{Argmin}\left(\int q(d\tilde{\mu}' d\hat{\mu}' | \tilde{\mu}, \hat{\mu}, \cdot) V(\tilde{\mu}', \hat{\mu}')\right). \quad (17)$$

- Conversely, if φ is an optimal randomized stationary policy, then (17) holds a.s. with respect to the corresponding stationary distribution of $\{\tilde{\mu}_n, \hat{\mu}_n\}$.

ACKNOWLEDGMENT

The work of the first author is supported in part by NSF grant CCF-0500234. The work of the second author is supported in part by a J. C. Bose Fellowship.

REFERENCES

- [1] R. Benzi, A. Sutera, and A. Vulpiani, "The Mechanism of Stochastic Resonance", *Journal of Physics, A*, Vol. 14, pp. L 453 -L457, 1981.
- [2] R. Benzi, G. Parisi, A. Sutera, and A. Vulpiani, "A Theory of Stochastic Resonance in Climatic Change", *SIAM Journal on Applied Mathematics*, Vol. 43, pp. 565 -578, 1983.
- [3] V. S. Borkar, "Average Cost Dynamic Programming Equations for Controlled Markov Chains with Partial Observations", *SIAM Journal on Control and Optimization*, Vol. 39, No. 3, pp. 673 -681, 2000.
- [4] Vivek S. Borkar, Sanjoy K. Mitter, and Sekhar Tatikonda, "Optimal Sequential Vector Quantization of Markov Sources", *SIAM Journal on Control and Optimization*, Vol. 40, No. 1, pp. 135 -148, 2001.
- [5] M. I. Freidlin, "Quasi-deterministic Approximation, Metastability, and Stochastic Resonance", *Physica D*, Vol. 137, pp. 333 -352, 2000.
- [6] L. Grammaticoni, P. Hanggi, P. Jung, and F. Marchesoni, "Stochastic Resonance", *Reviews in Modern Physics*, Vol. 70, pp. 223 -287, 1998.
- [7] B. Kosko and S. Mitaim, "Stochastic Resonance in Noisy Threshold Neurons", *Neural Networks*, Vol 16, pp. 755 -761, 2003.
- [8] Onésimo Hernández-Lerma, and Jean Bernard Lasserre, *Discrete-Time Markov Control Processes*, Springer-Verlag, New York, 1996.
- [9] S. Mitaim and B. Kosko, "Adaptive Stochastic Resonance in Noisy Neurons Based on Mutual Information", *IEEE Transactions on Neural Networks*, Vol. 15, No. 6, pp. 1526 -1540, 2004.
- [10] K. Wiesenfeld and F. Moss, "Stochastic Resonance and the Benefits of Noise: from ice ages to crayfish and SQUIDS", *Nature*, Vol. 373, pp. 33 -36, 1995.