1. Introduction

The work reported here is based on the joint paper [1] of the author and W. Runggaldier.

A large part of optimal stochastic control theory is concerned with diffusion models of the following type. Here, $x(\cdot)$ is the system, $z(\cdot)$ a reference process and $y(\cdot)$ an observation process.

\begin{align*}
\text{(1.1a) } & \quad dx = b_x(x,u)dt + \sigma_x(x)dW_x \\
\text{(1.1b) } & \quad dz = b_z(z)dt + \sigma_z(z)dW_z \\
\text{(1.1c) } & \quad dy = h(x,z)dt + \sigma(z)dW_z
\end{align*}

$u(\cdot)$ is a control - and takes values in a compact set $U$. It is appropriately non-anticipative with respect to whatever the data is (either $x(\cdot)$ or $y(\cdot)$). Typical cost functions are

\begin{align*}
\text{(1.2) } & \quad R_T(x,u) = \mathbb{E}_x \int_0^T k(x(s),u(s))ds, \\
\text{(1.3) } & \quad \gamma(u) = \frac{1}{T} \int_0^T R_T(x,u).
\end{align*}

The actual physical problem would almost never be of the type (1.1). The system (1.1) is usually chosen purely for simplicity in both the analysis and implementation. Its use on the physical problem must be justified, since it is
certainly not apriori obvious, even if (1.1) is a good approximation to a physical problem for each fixed nice control, that a control that is optimal or nearly optimal for (1.1) will be good or nearly optimal for the physical problem.

In this paper an outline of required justification is provided for a large class of physical models. We work with one particular continuous parameter form—driven by wide bandwidth noise— but the general method yields identical results for a wide variety of continuous and discrete parameter models. Basically, what we require is that for each 'nice' control the models converge weakly to (1.1) as the 'bandwidth' of the driving noise goes to infinity. 'Bandwidth' is used in a loose intuitive sense only—the processes need not even be stationary in general.

The 'physical' model. References [2],[4] concern the filtering problem. We work with the control problem only here. Let \( \xi(t) \) be a stationary process and for \( \varepsilon > 0 \), define \( \xi^\varepsilon(t) = \xi(t/\varepsilon^2) \). For \( x^\varepsilon(0) = x \) given, define \( x^\varepsilon(.) \) by

\[
(1.4) \quad x^\varepsilon = b(x^\varepsilon,u) + \mathcal{S}(x,\xi^\varepsilon) + g(x,\xi^\varepsilon)/\varepsilon, \quad x \in \mathbb{R}^n,
\]

for appropriate functions \( b, \mathcal{S} \) and \( g \). Here \( E\mathcal{S}(x,\xi) = E(g(x,\xi)) = 0 \). The scaling in (1.4) is a common way of modelling wide bandwidth noise. If the spectral density of \( \xi(.) \) exists and is \( S(w) \), then that of \( \xi(\cdot/\varepsilon^2)/\varepsilon \) is \( S(\varepsilon^2w) \). Hence if \( S(\cdot) \) is continuous and non-zero at \( w = 0 \), the spectrum converges to that of white noise as \( \varepsilon \to 0 \). The scaling is also the correct one from the point of view of weak convergence theory. Define

\[
(1.5) \quad R^\varepsilon_T(x,u) = E \int_0^T k(x^\varepsilon(s), u(s))ds
\]

\[
(1.6) \quad \gamma^\varepsilon(u) = \lim_{T \to \infty} R^\varepsilon_T(x,u)/T.
\]

The general results hold for many other classes of models: e.g., noise