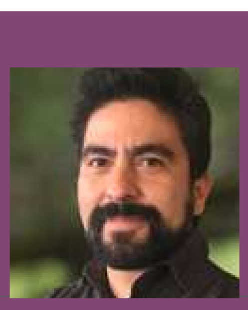
CONSTRAINED STOCHASTIC DIFFERENTIAL GAMES WITH ADDITIVE STRUCTURE: THE TOTAL EXPECTED PAYOFF

Anáhuac México

Beatris Adriana Escobedo Trujillo
Facultad de Ingeniería Civil.
Universidad Veracruzana
bescobedo@uv.mx



José Daniel López-Barrientos
Facultad de Ciencias Actuariales
daniel.lopez@anahuac.mx

Javier Garrido-Meléndez
Facultad de Ingeniería Civil.
Universidad Veracruzana
jgarrido@uv.mx

INTRODUCCIÓN

In this work, we use the approach presented in [1] to study a finite-horizon optimal control problem with restrictions. Our reward function is not required to be bounded (as in, for example [2]), although we certainly require it to comply with a Lyapunov-like regularity condition (see [3]). We intend to optimize the total reward criterion on a finite-time horizon subject to:

$$E_x^u\left[\int_0^T c(s,x(s),u(s))ds + c_1(T,x(T))\right] \leq E_x^u\left[\int_0^T \theta(s,x(s))ds + \theta_1(T,x(T))\right],$$

Where the functions inside the expectation are the cost rates, and the constraint rates.

MATERIAL Y MÉTODO

Assume the existence (in the sense of Markov-Feller — see [4]) of a stochastic differential system of the form: $dx(t) = b(x(t), u(t))dt + \sigma(x(t))dW(t)$, where the functions in the right-hand-side of this expression are the drift and diffusion coefficients of the system, and W is a d-dimensional Brownian motion. With this in mind, we use the dynamic programming and the Lagrange multipliers techniques (see [5]) to characterize optimal strategies in the spirit of the Topology of relaxed controls introduced in [6].

RESULTADOS

We present the traditional verification results and an extension of the Kawaguchi's example on the control of pollution accumulation (see [7]).

Theorem. For each fixed negative Lagrange multiplier, and all positive time, the finite-horizon optimal reward $J_T^*(t,\cdot,\cdot,r^\lambda)$

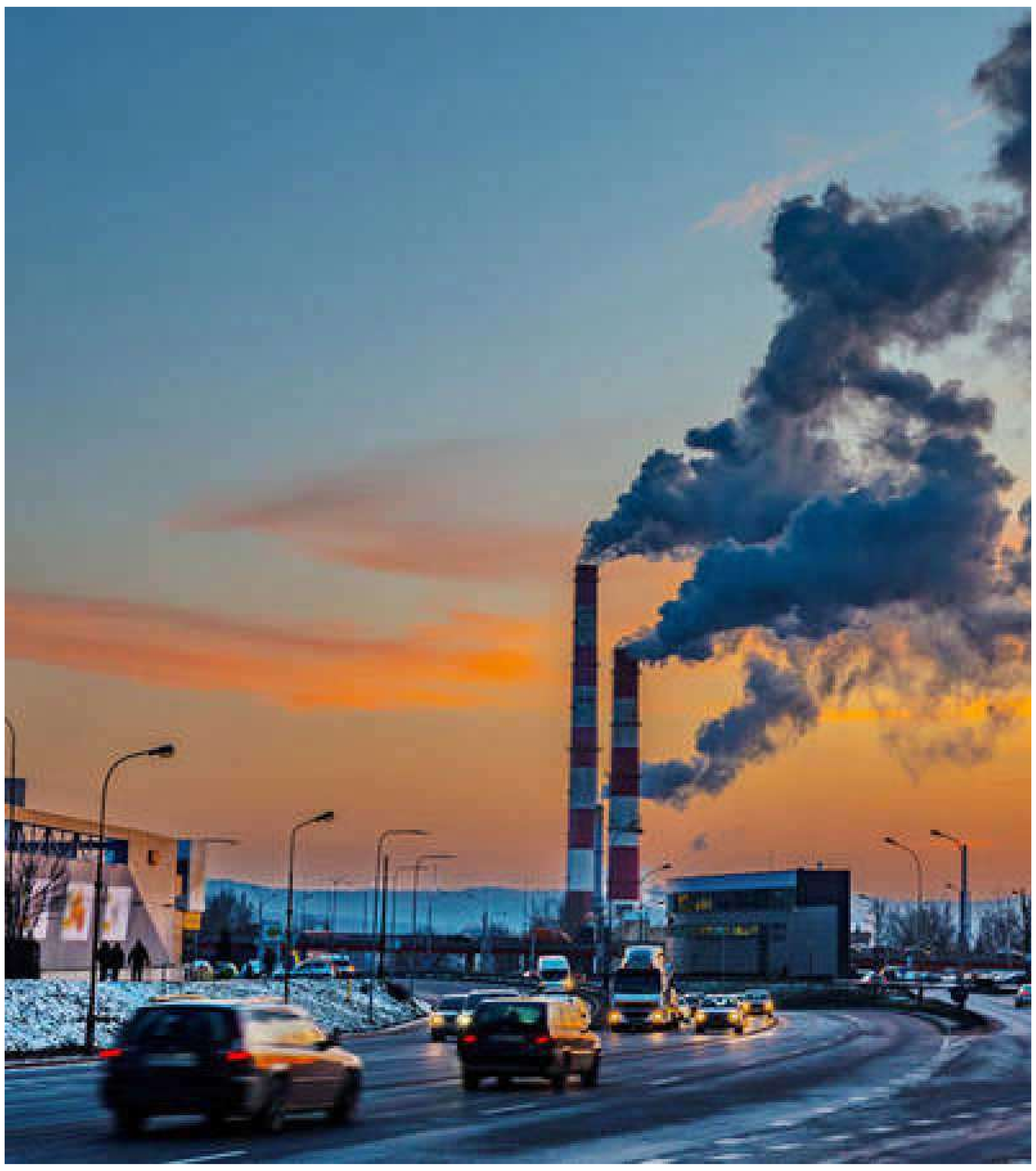
belongs to certain Sobolev space, and verifies the *total reward Hamilton-Jacobi-Bellman (HJB) equation;* that is,

$$0 = \sup_{\pi\in\Pi}\left\{r^\lambda(x,\pi) + \partial_t J_T^*(t,x,r^\lambda) + \mathbb{L}^\pi J_T^*\left(t,x,r^\lambda\right)\right\} \text{ for all } (t,x) \in [0;T]\times\mathbb{R}^n.$$

Conversely, if for all positive times within the horizon, some function satisfies the HJB equation, then this function coincides with the objective function. Moreover, if there exists a stationary policy f that maximizes the right-hand-side equation HJB, then this policy is finite-horizon optimal.

DISCUSIÓN

In this work we present a finite-horizon constrained control problem, which we solve with an extension of the Lagrange multipliers technique, and dynamic programming. We managed to give conditions to characterize the solution of the HJB equation, and to interpret it in terms of a pollution accumulation problem.



REFERENCIAS

1. Escobedo-Trujillo BA, Alaffita Hernández FA, López-Martínez R. Constrained stochastic differential games with additive structure: Average and discount payoffs. Journal of Dynamics and Games. 2018;5(2):109–41.
2. Fleming WH, Souganidis PE. On the existence of value functions of two player, zero-sum stochastic differential games. Indiana University Mathematics Journal. 1989;38(2):293–314.
3. Meyn SP, Tweedie RL. Stability of Markovian Processes III: Foster-Lyapunov Criteria for Continuous-Time Processes. Advances in Applied Probability [Internet]. 1993;25(3):518–48.
4. Arapostathis A, Ghosh MK, Borkar VS. Ergodic Control of Diffusion Processes. Cambridge University Press; 2011.
5. Jasso-Fuentes H, Escobedo-Trujillo BA, Mendoza-Pérez AF. The Lagrange and the vanishing discount techniques to controlled diffusions with cost constraints. Journal of Mathematical Analysis and Applications [Internet]. 2016 May 15 [cited 2019 Sep 7];437(2):999–1035.
6. Borkar VS. A topology for Markov controls. Applied Mathematics and Optimization. 1989;20:55–62.
7. Kawaguchi K, Morimoto H. Long-run average welfare in a pollution accumulation model. Journal of Economic Dynamics and Control. 2007;31(2):703–20.