DEPARTMENT OF INTERNATIONAL AND EUROPEAN ECONOMIC STUDIES

ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS

Spatiotemporal robust control in infinite dimensional spaces

William Brock

Anastasios Xepapadeas

Athanasios Yannacopoulos

Working Paper Series

14-05

April 2014

Spatiotemporal robust control in infinite dimensional spaces *

W. A. Brock[1], A. Xepapadeas[2] and A. N. Yannacopoulos [3],
(1) Department of Economics,
University of Wisconsin, Madisson,
(2) Department of International and European Studies,
Athens University of Economics and Business, xepapad@aueb.gr,
(3) Department of Statistics,
Athens University of Economics and Business, ayannaco@aueb.gr

April 11, 2014

Abstract: We formulate infinite dimensional stochastic robust optimal control problems, motivated by applications arising in interconnected economic systems, or spatially extended economies. We study in detail linear quadratic problems and nonlinear problems. We derive optimal robust controls and identify conditions under which concerns about model misspecification at specific cite(s) could cause regulation to break down, to be very costly, or to induce pattern formation and spatial clustering. We call sites associated with these phenomena hot spots.
Keywords: Infinite dimensional stochastic control, robust control, Operator and matrix Riccati equation, Hot spot formation, Pattern formation, Hamilton-Jacobi Bellman equation.
AMS classification:
JEL classification:

1 Introduction

There are several applications in economics in which temporal models (i.e. models taking into account the evolution of economic quantities as depending only on time) are not enough and the spatial dimension has to be taken into consideration. Such areas, include spatial or environmental economics, etc. It is by now well understood that the interplay of the temporal with the spatial dimensions may introduce interesting phenomena into our models, related to pattern formation behaviour, agglomeration formation or absense of it etc. Such behaviour is known to persist even in controlled systems, and may serve as a way of understanding certain phenomena in the economics of space. While important from the point of view of economic theory, the introduction of space into the models turns them into infinite dimensional dynamical systems, whose behaviour is by far more complex and technically difficult (both from the analytic and the numerical point of view) to study. The introduction of space, turns our standard ordinary differential or difference equation models into partial differential or integral equation models, which are seriously more difficult to analyse and present far richer qualitative behaviour than their temporal (finite dimensional) counterparts.

*This research has been co-financed by the European Union (European Social Fund ESF) and Greek national funds through the Operational Program "Education and Lifelong Learning" of the National Strategic Reference Framework (NSRF) - Research Funding Program: "ARISTEIA Athens University of Economics and Business - Spatiotemporal Dynamics in Economics". W. Brock is grateful for financial and scientific support received from The Center for Robust Decision Making on Climate and Energy Policy (RDCEP) which is funded by a grant from the National Science Foundation (NSF) through the Decision Making Under Uncertainty (DMUU) program. He is also grateful to the Vilas Trust for financial support. None of the above are responsible for any errors, opinions, or shortcomings in this article.

The situation becomes worse, since in most cases economic models are contaminated by noise, due either to insufficient knowledge or imperfect measurement, so the economic quantities now become random variables, whose evolution is provided by stochastic differential equations. The introduction of the spatial variable, turns the relevant quantities into random fields rather than stochastic processes, and this may introduce soem very interesting new phenomena in the model but at the cost of more technical complications, both analytically and numerically.

To make the situation even worse, in most cases we know that there is noise contaminating the state of the system (i.e. there exist fluctuations around a mean state) but we do not know the exact statistical distribution of this noise term, i.e. we do not have a particular model concerning the noise that we may trust. Therefore, the best we can do is consider a whole universe of possible models, each with some degree of credibility, and try to obtain a decision which is consistent with the whole universe of models rather than with a single "reference" model. One we to do that, is to take this decision which shields us against the worst possible scenario, and this can be conveniently re-expressed in terms of a min-max decision problem, akin to a game theoretic model. This approach has been studied very deeply by Hansen and Sargent (Hansen and Sargent (2001)) who have proposed ways of obtaining robust decision rules employing techniques from the theory of robust control.

The important question that arises from such studies is regarding the interplay of model uncertainty, risk, and spatio-temporal variability of the state variables of the model. How can we obtain robust rules in such cases and how would the robust rules vary in space, can they be expected to turn spatially homogeneous policies. Furthermore, are there cases that robust rules cannot be formulated (so that the decision maker is simply overwhelmed by uncertainty) and therefore, control may turn inefficient, if not harmful? Such questions are important not only from the quantitative point of view but also from the philosophical point of view, regarding the futility or not of the act of modeling.

A first attempt to the study of such questions has been made in Brock et al. (2014) where a finite dimensional model, which was an approximation of a spatio-temporal connected economic model was studies under the effects of control, risk and uncertainty. The model was essentially a compartmental model in which the state of the system was considered as spatially homogeneous within certain cells, but there was a variability of the state across cells, and there was interaction between the cells as well. The robust control was formulated as a stochastic differential game and was treated in terms of the Hamilton-Jacobi-Bellman-Isaacs equation. One of the important findings was (a) the breakdown of the solution of the robust control problem as a result of increased uncertainty concerning the true model, that was allowed to vary locally, i.e., we allowed for different levels of uncertainty across cells and (b) the observations that deviations from the desired final state often presented a spatial pattern, i..e there were cells which performed worse than others when the distance from a predefined target was considered. This behaviour has been called hot spot formation, and has been thoroughly analyzed in Brock et al. (2014) for finite dimensional systems in the above simplified form, and criteria for the emergence of such behaviour have been obtained. The results have been applied in the study of distance dependent utility functions for consumption (Brock et al. (2014)) or optimal management of fisheries.

It is the aim of the present paper to study the full spatial temporal problem robust control problem in its infinite dimensional version, without resorting to the oversimplified compartmental model studied in Brock et al. (2014). Our approach uses the theory of Hamilton-Jacobi-Bellman-Isaacs equations on infinite dimensional separable Hilbert spaces and it is interesting to note that modulo technical difficulties confirms the findings of our finite dimensional model, as far as the breakdown of rocust policy rules is concerned.

2 Modeling a spatial economy under uncertainty

2.1 The controlled state equation

Consider the economy as being located on a discrete lattice $\mathfrak{L}$, finite or infinite e.g. $\mathfrak{L} = (\mathbb{Z}_N)^d$ or $\mathfrak{L} = (\mathbb{Z})^d$ respectively. By the term "economy" at this point we consider a collection of state variables $x = \{x_n\}$, $n \in \mathfrak{L}$. For fixed n, $x_n \in \mathbb{R}$ and it corresponds to the state of the economy at lattice site n. Since the infinite lattice is the technically more involved we deal with this case here, setting without loss of generality $\mathfrak{L} = \mathbb{Z}$, all the results being valid in the finite dimensional case under straightforward modifications. We therefore consider the state variable x as taking values on a sequence space. To keep our discussion within a Hilbert space setting we choose to work with the sequence space $\ell^2 = \{\{x_n\},\ \sum_{n\in\mathbb{Z}} x_n^2 < \infty\}$ or when explicitly stated so with a weighted version of this space. This space is a Hilbert space with a norm derivable from the inner product $\langle x, y\rangle = \sum_{n\in\mathbb{Z}} x_n y_n$. Given this economy we consider a social planning problem modelled as an optimal linear regulator problem (e.g.Ljungqvist and Sargent (2004)). The optimal linear regulator problem refers to the optimization of a quadratic objective defined over the whole lattice by exerting on each lattice site a control $u_n \in \mathbb{R}$ where the control for the whole economy is described as a sequence $u = \{u_n\}$, $n \in \mathbb{Z}$ such that $u \in \ell^2$.[1]

The economy evolves in time and this is modelled by considering the state of the economy as described by a function $\check{x} : I \to \ell^2$ such that $\check{x}(t) = \{x_n(t)\}$, $n \in \mathfrak{L}$, where $x_n(t)$ is the state of the system at site n at time t. To ease notation we will use the notation x for this function and similarly u for the control exerted on the system. In this paper, we are interested in an infinite horizon economy and thus we assume $I = \mathbb{R}_+$. The evolution of the state of the economy in time is subject to statistical fluctuations (noise) modelled in terms of a stochastic process $w = \{w_n\}$, $n \in \mathbb{Z}$, which is considered as a cylindrical Wiener process on a suitable filtered probability space $(\Omega, \{\mathcal{F}_t\}_{t\in\mathbb{R}_+}, \mathcal{F}, P)$ (see e.g., Carmona and Tehranchi (2006)). The introduction of noise turns the state of the system for a fixed time t into an ℓ^2-valued random variable, thus the state of the system can be described as an ℓ^2-valued stochastic process. We assume that this stochastic process is the solution of an infinite dimensional stochastic differential equation of the form

$$dx_n = (\sum_m a_{nm} x_m + \sum_m b_{nm} u_m) dt + \sum_m c_{nm} dw_m,\ \ n \in \mathbb{Z}$$

where the last term, describing the fluctuations of the state due to the stochasticity, is understood in the sense of the Itō theory of stochastic integration. In compact form this can be expressed as

$$dx = (\mathsf{A}x + \mathsf{B}u)\, dt + \mathsf{C}dw \tag{1}$$

where $\mathsf{A}, \mathsf{B}, \mathsf{C} : \ell^2 \to \ell^2$ are linear operators, related to the doubly infinite matrices with elements a_{nm},b_{nm}, c_{nm}, respectively. The state equation (1) is an infinite dimensional Ornstein-Uhlenbeck equation on the Hilbert space ℓ^2.

At this point we make some comments concerning the economic intuition behind the state equation (1). Our model is a "spatial" economy where it is considered that the state of the economy at point m has an effect at the state of the economy at point n. This effect is quantified through an influence "kernel" (or rather a discretized version of an influence kernel) which assumes the form of a double sequence $\mathsf{A} = (a_{nm})$ (if one prefers it may be viewed as a doubly infinite matrix). The entry a_{nm} provides a measure of the influence of the state of the system at point m to the state of the system at point n. Network effects knowledge spillovers can be modelled for example through a proper choice of A. For instance, if the economies do not interact at all then $\mathsf{A} = a_{nm} = \delta_{n,m}$ where $\delta_{n,m}$ is the Kronecker delta. If only next neighborhood effects are possible then a_{nm} is non-zero only if m is a

[1] The generalization to vector valued state and control variables $x^n \in \mathbb{R}^{d_1}$ and $u_n \in \mathbb{R}^{d_2}$ on each site $n \in \mathbb{Z}$ requires the use of the sequence spaces $\ell^2(\mathbb{R}^{d_i})$, $i = 1, 2$ rather than $\ell^2 := \ell^2(\mathbb{R})$) and is straightforward.

neighbor of n. Such an example is the discrete Laplacian. Similarly, the controls at different point of the lattice u_m are assumed to have an effect at the state of the system at site n, through the term $\sum_m b_{nm} u_m$. For example in a model of a spatial fishery, fishing effort at a given site may affect fish biomass at another sites through biomass movements. A similar interpretation for this term holds as for the term $\sum_m a_{nm} x_m$. We will identify the doubly infinite matrices $\mathsf{A} = (a_{nm})$ and $\mathsf{B} = (b_{nm})$ with operators denoted by the same symbol, acting from $\ell^2 \to \ell^2$ or if needed, between properly weighted versions of these spaces. Under simple summability conditions these operators may be bounded or compact.

Finally, the interpretation of the third term $\sum_m c_{nm} dw_m$ is a term that tells us how the uncertainty at site m is affecting the uncertainty concerning the state of the system at site n. The double sequence (or doubly infinite matrix) $\mathsf{C} = (c_{nm})$ can be thought of as the spatial autocorrelation operator for the system.

If the system is finite, e.g. if $\mathfrak{L} = \mathbb{Z}_N$ then the above matrices are finite $N \times N$ matrices which of course may still be treated as representations of operators.

2.2 Model uncertainty

Assume now that there is some uncertainty concerning the "true" statistical distribution of the state of the system. This corresponds to a family of probability measures $\mathcal{Q}$ such that each $Q \in \mathcal{Q}$ corresponds to an alternative stochastic model (scenario) concerning the state of the system. Considering for the time being finite horizon T, we restrict ourselves to measures which are equivalent with P (i.e. having the same null sets) such that the Radon-Nikodym derivatives dQ/dP are defined through an exponential martingale of the type employed in Girsanov's theorem,

$$\left.\frac{dQ}{dP}\right|_{\mathcal{F}_T} = \exp\left(\int_0^T \sum_n v_n(t) dw_n(t) - \frac{1}{2}\int_0^T \sum_n v_n^2(t) dt\right)$$

where $v = \{v_n\}$, $n \in \mathbb{Z}$ is an ℓ^2-valued stochastic process which is measurable with respect to the filtration $\{\mathcal{F}_t\}$ satisfying the Novikov condition $\exp(\int_0^T \sum_n v_n^2(t)dt) < \infty$. If the process v can be chosen so that this condition is true for all T, then the result in the infinite horizon limit follows by appropriately passing to the limit $T \to \infty$. Furthermore, the same theorem guarantees that $\bar{w}_n(t) = w_n(t) - \int_0^t v_n(s)ds$ is a Q-Brownian motion for all $n \in \mathbb{N}$, where the drift term v_n may be considered as a measure of the model misspecification at lattice site n. Thus, the infinite dimensional version of Girsanov's theorem (see e.g. Carmona and Tehranchi (2006)) shows that the adoption of the family $\mathcal{Q}$ of alternative measures concerning the state of the system, leads to a family of different equations for the state variable

$$dx_n = (\sum_m a_{nm} x_m + \sum_m b_{nm} u_m + \sum_m c_{nm} v_m) dt + \sum_m c_{nm} d\bar{w}_m, \quad n \in \mathbb{Z}.$$

The state variables $x = \{x_n\}$ depend on the choice of $u = \{u_n\}$ and $v = \{v_n\}$ therefore, $x = x^{u,v}$, however we choose to avoid this notation for simplicity. We therefore tacitly assume that x indicates the state of the system when the measure Q corresponding to the "information drift" $v = \{v_n\}$ and the control procedure $u = \{u_n\}$ is adopted. In compact form this equation becomes the infinite dimensional Ornstein-Uhlenbeck equation

$$dx = (\mathsf{A}x + \mathsf{B}u + \mathsf{C}v)dt + \mathsf{C}d\bar{w}, \tag{2}$$

where for notational convenience the superscripts u, v are omitted from x.

The following proposition guarantees the well posedness of the state equation (2).

Proposition 1. *Assume that*

(i) A *is the infinitesimal generator of a C_0- contraction semigroup on the Hilbert space $H = \ell^2_\rho$ for a suitable choice of weight sequence ρ.*

(ii) B, C *are bounded operators when considered in the above fuctional setting.*

Then, for any $T > 0$, the state equation (2) is well posed in $L^2(\Omega \times [0,T]; \ell^2)$ for every $u, v \in L^2(\Omega \times [0,T]; \ell^2_\rho)$. Furthermore, if $x(t;y)$ is the solution of the state equation for the initial condition y, the transition semigroup

$$P_t\phi(y) := \mathbb{E}[\phi(x(t;y))], \;\; t \geq 0, \;\; x \in \mathbb{H}, \;\; \phi \in C_b(\mathbb{H})$$

has the Feller property. By $C_b(\mathbb{H})$ we denote the set of functions $\phi : \mathbb{H} \to \mathbb{R}$, which are uniformly continuous and bounded.

Proof: The proof follows by standard arguments see e.g. Da Prato and Zabczyk (1996) Theorem 5.5.8 and references therein. ■

Remark 1. The assumptions of Proposition 1 can be met quite easily for general types of the interaction matrix $a = \{a_{ij}\}$. In fact general conditions on a are given so that A is an m-dissipative operator in a variety of sequence spaces of ℓ^2 type (weighted or not) are given in Prato and Zabczyk (1995); see also Da Prato and Zabczyk (1996).

Remark 2. The fact that the solution $x(t;y)$ gives rise to a transition semigroup on $C_b(\mathbb{H})$ guarantees the Markovian property of the solution of the state equation. This is indispensable for the treatment of the optimal control problem, using the Hamilton-Jacobi-Bellman approach. Recall, that this approach relies heavily on the dynamic programming principle which in turns has to assume the Markov property of the solution.

If certain more restrictive conditions hold, then the solution of the infinite dimensional state equation 2) can be very well behaved and in particular may have properties akin to those of finite dimensional systems. This condition is the famous controllability condition (proposed by Zabczyk, see e.g. Da Prato and Zabczyk (1996) or Da Prato and Zabczyk (2002), Ch. 6, condition (6.2.3) p. 104) that requires

$$e^{t\mathsf{A}}(\mathbb{H}) \subset \mathsf{C}_t^{1/2}(\mathbb{H}), \quad \forall\, t > 0, \tag{3}$$

where the operator family C_t is defined by

$$\mathsf{C}_t x = \int_0^t e^{s\mathsf{A}} \mathsf{C} e^{s\mathsf{A}^*} x ds, \quad x \in \mathbb{H}, \; t \geq 0.$$

This condition is related to the null controllability of a related deterministic control system and as we shall see later on it plays an important role in connecting the solution of the Hamilton-Jacobi-Bellman equation to the solution of the robust optimal control problem under consideration[2]. If the controllability relation (3) holds, then, the solution of the infinite dimensional Ornstein-Uhlenbeck equation is more regular, and in particular generates a strongly Feller transition semigroup (that maps bounded functions to smooth functions). In this respect the system, eventhough it is infinite dimensional does not present "pathological" properties but behaves as a finite dimensional one. For a detailed account of these results see e.g. Theorem 6.2.2 in Da Prato and Zabczyk (2002).

Condition (3) can be characterized quite easily for a large class of operators A and C, in terms of their spectrum, as the next example shows.

[2]It is very important to realize that the HJB equation, written down most of the times for optimal control problems is a formal equation, which may or may not be related to the value function of a particular optimal control problem. This is a problem that may arise either in finite or in infinite dimensions. One is entitled to use the HJB equation to characterize the solution of an optimal control problem, only if certain restrictive conditions hold, that allow us to do so. For the particular problem at hand, it turns out that the Zabczyk controllability condition, is such a condition that legitimates the connection of the solutions of the HJB equation with the solution of the relevant optimal control problem.

Example 1 (Da Prato and Zabczyk (2002), Example 6.2.11)**.** Assume that A, C are diagonalizable operators such that ther exists an orthonormal basis $\{\phi_n\}$ such that $\mathsf{A}\phi_n = \alpha_n\phi_n$ and $\mathsf{C}\phi_n = \gamma_n\phi_n$. A fundamental condition for well posedness of problem (2) is that

$$\sum_n \frac{\gamma_n}{\alpha_n} < \infty$$

This condition guarantees that the family of operators C_t is such that its action on any $x \in \mathbb{H}$ leads to a function $x(\cdot) : \mathbb{R}_+ \to \mathbb{H}$ which is integrable with respect to time (which clearly is essential in defining the solution of the differetial equation (2)). When examined carefully this condition provides a balance between the action of the noise (as modelled with the use of the covariance operator C) and the action of the determistic interaction between lattice sites (as modelled with the use of the operator A). The physical meaning of this condition is that very rough noise effects (for instance the case where $\mathsf{C} = I$, the identity operator) can be counterbalanced if the deterministic interaction between lattice sites has spatial regularizing effects. Furthermore, the controllability condition (3) becomes

$$\sup_n \frac{\alpha_n}{\gamma_n(e^{2\alpha_n t} - 1)} < \infty, \quad t \geq 0$$

which again can be interpreted as some sort of balance between noise effects and deterministic inter-site economic efects.

As a particular case consider the case where $\mathsf{A}\phi_n = -n^{\alpha}\phi_n$, and $\mathsf{Q}\phi_n = n^{-\gamma}\phi_n$. Then, if $\alpha+\gamma > 1$ holds the controllability condition holds as well. The choice of α and γ, further provides important information on the regularization properties of the Feller semigroup generated by the solution of equation (2). This is important for the properties of the solution of the related HJBI equation.

2.3 The control objective

We now define the control objective. Let us first fix a model, i.e. let us assume that the drift v is fixed. Then, the control procedure is designed so that the distance form a desired target, chosen without loss of generality to be $x^0 = 0$ the zero sequence, is minimized at the minimum possible cost, as measured by the amplitude of the control variable u. Therefore, having chosen the state variable x as given by the solution of the dynamic equation (2) the decision maker's goal is to choose the control procedure u so as to solve the stochastic control problem[3]

$$\min_u \mathbb{E}_Q \left[\int_0^{\infty} e^{-rt} \sum_{n,m} \left(p_{nm} x_n(t) x_m(t) + q_{nm} u_n(t) u_m(t) \right) dt \right]$$

or in compact form

$$\min_u \mathbb{E}_Q \left[\int_0^{\infty} e^{-rt} (\langle \mathsf{P}x(t), x(t) \rangle + \langle \mathsf{Q}u(t), u(t) \rangle) dt \right]$$

where $\langle \cdot, \cdot \rangle$ is the inner product in the Hilbert space ℓ^2 and $\mathsf{P}, \mathsf{Q} : \ell^2 \to \ell^2$ are symmetric positive operators, whose infinite matrix representation is $\mathsf{P} = \{p_{nm}\}$ and $\mathsf{Q} = \{q_{nm}\}$ respectively.

In the special case of diagonal operators $p_{nm} = p\,\delta_{nm}$, $q_{nm} = q\,\delta_{nm}$ this functional assumes the simplified form

$$\min_u \mathbb{E}_Q \left[\int_0^{\infty} e^{-rt} \sum_n p(x_n(t))^2 + q(u_n(t))^2) dt \right]$$

[3]To simplify the exposition we do not consider. without loss of generality, cross product terms in the quadratic objective. Cross product terms can be eliminated by a change in units see for example Magill (1977b).

The first sum, can be considered as the total deviation of the states of the system at each site from the desired state 0 whereas the second sum is the total control exerted on the system in the effort to drive it to 0. We emphasize that because of the linearity of the system the choice of 0 as the target state is without any loss of generality whatsoever. This problem is solved under the adoption of the measure Q, related to the drift v, i.e. it is solved under the dynamic constraint (2). This will provide a solution leading to a value function $V(x_0; v)$; corresponding to the minimum deviation obtained for the model Q_v under the minimum possible effort. Being uncertain about the true model, the decision maker will opt to choose this strategy that will work in the worst case scenario; this being the one that maximizes $V(x_0; v)$, the minimum over all u having chosen v, over all possible choices for v. Therefore, the robust control problem to be solved is of the general form

$$\min_u \max_v \mathbb{E}_Q \left[\int_0^\infty e^{-rt} \sum_n \sum_m (p_{nm} x_n(t) x_m(t) + q_{nm} u_n(t) u_m(t) - \theta r_{nm} v_n(t) v_m(t)) dt \right] \tag{4}$$

or in compact form

$$\min_u \max_v \mathbb{E}_Q \left[\int_0^\infty e^{-rt} (\langle (\mathsf{P}x)(t), x(t) \rangle + \langle (\mathsf{Q}u)(t), u(t) \rangle - \theta \langle (\mathsf{R}v)(t), v(t) \rangle) dt \right], \tag{5}$$

subject to the dynamic constraint (2), where $\theta > 0$ and $\mathsf{R} = \{r_{nm}\}$ is a symmetric positive operator. The third term corresponds to a quadratic loss function related to the "cost" of model misspecification. Quadratic loss functions are rather common in statistical decision theory, mainly on account of their connection with entropy (see Proposition 2).

In the special case where $p_{nm} = p\,\delta_{nm}$, $q_{nm} = q\,\delta_{nm}$, $r_{nm} = \delta_{nm}$ the cost functional simplifies to

$$\min_u \max_v \mathbb{E}_Q \left[\int_0^\infty e^{-rt} \sum_n p(x_n(t))^2 + q(u_n(t))^2 - \theta \sum_n (v_n(t))^2) dt \right]$$

subject to the dynamic constraint (2). The new term in the functional, is related to our aversion for model misspecification.

Another important special case is the case where $p_{nm} = p_n\,\delta_{nm}$, $q_{nm} = q_n\,\delta_{nm}$, $r_{nm} = \theta_n \delta_{nm}$ and $\theta = 1$. Then the control functional simplifies to

$$\min_u \max_v \mathbb{E}_Q \left[\int_0^\infty e^{-rt} \sum_n p_n(x_n(t))^2 + q_n(u_n(t))^2 - \sum_n \theta_n (v_n(t))^2) dt \right]$$

subject to the dynamic constraint (2). This version of the control functional introduces localized concerns on the deviation from particular state targets, on the cost of required control as well as on the cost of model misspecification.

Remark 3 (Interpretation as a differential game)**.** One particularly intuitive way of viewing this problem is as a two player game, the first player is the decision maker while the second player is nature who has *control* over the uncertainty. The first player chooses her actions so as to minimize the distance of the state of the system from a chosen target at the minimum possible cost, whereas the second player is a considered by the first player as a malevolent player who tries to *mess up* the first players efforts. This interpretation allows us to use the Hamilton-Jacobi-Bellman-Isaacs equation approach for the solution of the robust control problem.

2.4 Relation with entropic constrained robust control

The optimization problem (5) subject to (2) for various choices of the operator R is related to entropic constrained robust control. We present here two examples where this assertion holds. The first example is related with a "global" in space entropy constraint, while the second example is related with a "localized" in space entropy constraint, which may be even more relevant in the robust control of spatially varying interconnected economic systems.

Proposition 2 (Global entropy constraints)**.** *The optimization problem (5) subject to (2), for the choice $R = I$, is related to a robust control problem with an entropic constraint of the form*

$$\inf_u \sup_{Q\in\mathcal{Q}} \mathbb{E}_Q \left[\int_0^\infty e^{-rt} (\langle \mathsf{P}x(t), x(t)\rangle + \langle \mathsf{Q}u(t), u(t)\rangle) dt \right],$$
$$\textit{subject to } \ \mathcal{H}(P \mid Q) < H_0$$

and the dynamic constraint (2), where by $\mathcal{H}(P \mid Q)$ we denote the Kullback-Leibler entropy of the probability measures P and Q.

Proof: Consider first a finite horizon problem with horizon $T > 0$. Within the class of models considered, for any $T > 0$, an application of the Girsanov theorem yields that the likelihood of the models is given by the Radon-Nikodym derivative of the measures P and Q in the form of equation (**??**) as long as the Novikov condition holds. A quick calculation yields that the relative entropy of P and Q is given by

$$\mathcal{H}(Q \mid P) := \mathbb{E}_Q \left[\ln\left(\frac{dQ}{dP}\right)\right] = \frac{1}{2}\int_0^T \sum_n v_n^2(t) dt$$

We now consider the robust optimization problem

$$\sup_{Q\in\mathcal{Q}} J(x, u; v)$$
$$\text{subject to} \qquad \mathcal{H}(Q \mid P) \le H_0$$

and the dynamic constraint (2) where $J(x, u; v) := \mathbb{E}_Q \left[\int_0^T e^{-rt} \left(\langle \mathsf{P}x(t), x(t)\rangle + \langle \mathsf{Q}u(t), u(t)\rangle\right) dt\right]$. The entropic constraint means that we are only considering models (i.e., measures Q) whose deviation in terms of the relative entropy from the "true" model (i.e., the measure P) is less than H_0. Taking into account the representation of the entropy in terms of $\{v_n\}$ and using Lagrange multipliers for the equivalent minimization problem $\inf_{Q\in\mathcal{Q}}(-J(x, u; v))$ we see that a solution of the relative entropy constraint problem is equivalent to the solution of

$$\inf_{Q\in\mathcal{Q}} -J(x, u; v) + \theta(\mathcal{H}(Q|P) - H_0),$$

subject to (2) where $\theta \in \mathbb{R}_+$ plays the role of the Lagrange multiplier. This is of course equivalent to the maximization problem

$$\sup_{Q\in\mathcal{Q}} J(x, u; v) - \theta(\mathcal{H}(Q|P) - H_0),$$

subject to (2) which using the representation of the relative entropy in terms of $\{v_n\}$ reduces to

$$\sup_{\{v_n\}} \mathbb{E}_Q \left[\int_0^T e^{-rt} \left(\langle \mathsf{P}x(t), x(t)\rangle + \langle \mathsf{Q}u(t), u(t)\rangle - \theta \sum_n v_n^2(t) \right) dt \right],$$

subject to (2). The above reasoning may explain the negative coefficient in front of the terms $\{v_n\}$. Taking the limit as $T \to \infty$ leads to the required result. ■

Generalizations of Proposition 2 for the case where more localized constraints with respect to model uncertainty are taken into account may be considered.

We motivate this by the following discussion: Assume that we are interested in the effect of uncertainty not on the Wiener process $w = \{w_n\}$ (the primary risk factors in our model) as such but rather on the process $W = \mathsf{T}w$ where $\mathsf{T} : \ell^2 \to \ell^2$ is an appropriate operator. This means that we are not interested on the effect of model uncertainty on $\{w_n\}$ but on the linear combination $\{\sum_m t_{nm} w_m\}$, which is assumed to reflect more accurately the effect of noise on the state variable. Various choices for the operator T are possible. One obvious choice is $\mathsf{T} = \mathsf{C}$, this means that the policy maker is interested in specifying the uncertainty so that she may understand its effect on the state of the system at each lattice site. Another obvious choice is to take T defined as $\mathsf{T}w = w_\ell$ where $\ell \in \mathbb{Z}$; this choice means that the policy maker is not worried about the uncertainty with respect to the noise term in general, but only as far as the uncertainty at site m is concerned.

Since w is a Wiener process under P, it follows that $w_n(t) \sim N(0,t)$ while under Q (as a consequence of Girsanov's theorem) $w_n(t) \sim N(-v_n, t)$. Therefore, by the properties of the normal distribution $W_m \sim N(0, \sum_m t_{nm} t_{nm})$ under the measure P, whereas it is distributed as $W_m \sim N(-\sum_m t_{nm} v_m, \sum_m t_{nm} t_{nm})$ under the measure Q. Therefore, if we only consider the marginal measures $\bar{P}_n$ and $\bar{Q}_n$ which are related with the distributions of the random variable W_n under the measures P and Q respectively, we may consider the "localized" entropy for the two measures as

$$\bar{d}_n := \mathcal{H}(\bar{Q}_n, \bar{P}_n) = \int_0^T \left(\sum_m t_{nm} v_m(t) \right)^2 dt. \tag{6}$$

The case where T is chosen such that $\mathsf{T}w = w_n$ (i.e. $\mathsf{T} = \pi_n$ the projection onto the lattice site n), then the local entropy is the entropy of the marginal measures P_n, Q_n which give the distribution of the component w_n given that w is distributed with the measure P and Q respectively. In this case

$$d_n := \mathcal{H}(Q_n, P_n) = \int_0^T v_n^2(t) dt. \tag{7}$$

The robust control problem

$$\inf_u \sup_{Q \in \mathcal{Q}} J(x; u, v)$$

subject to the localized entropy constraints $\bar{d}_n \leq H_n$ (where $\bar{d}_n$ is defined in (6)) or $d_n \leq H_n$ (where d_n is defined in (7) may lead to optimal control problems of the form discussed here for proper choice of the operator R. As an illustration we provide the following proposition, which of course can be generalized to other choices for the constraints.

Proposition 3. *The optimization problem (5) subject to (2), for the choice* $\mathsf{R} = \mathsf{D}$*, where* D *is a diagonal operator with representation* $d_{nm} = \theta_n \delta_{nm}$ *is related to a robust control problem with an entropic constraint of the form*

$$\inf_u \sup_{Q \in \mathcal{Q}} \mathbb{E}_Q \left[\int_0^\infty e^{-rt} (\langle \mathsf{P}x(t), x(t) \rangle + \langle \mathsf{Q}u(t), u(t) \rangle) dt \right],$$
$$\textit{subject to } \ \mathcal{H}(P_n \mid Q_n) < H_n, \ \ n \in \mathbb{Z}$$

and the dynamic constraint (2), where by $\mathcal{H}(P_n \mid Q_n)$ *we denote the Kullback-Leibler entropy of the marginal probability measures* P_n *and* Q_n *(see equation 7)).*

Proof: The proof follows the same steps as that of Proposition 2 only that now we need the Lagrangian

$$L = \sup_v \left\{ J(x; u, v) + \sum_n \theta_n \left(\mathcal{H}(\bar{Q}_n \mid \bar{P}_n) - H_n \right) \right\}$$

where $\{\theta_n\}$ is a sequence of Lagrange multipliers needed in order to guarantee that the localized entropic constraints hold. The major difference with Proposition 2 is that here we need a sequence of Lagrange multipliers rather than a single Lagrange multiplier, since now the constraints are more than one. Using the expression (7) for the entropies, we obtain the stated results. ■

Remark 4. The introduction of the local entropic constraints means that the policy maker is concerned on the effect of model uncertainty on W rather than on w, and her concerns differ at various lattice points. The concern of the policy maker on uncertainty at lattice site n is quantified by H_n, the smaller H_n is the less model uncertainty is she willing to accept for lattice site n. This assumption is not unreasonable as certain lattice points may be considered as more crucial than others therefore specific care should be taken for them.

Remark 5. In the robust control problem of Proposition 3 the maximizing adversarial agent - Nature - chooses a $\{v_n(t)\}$ while $\theta_n \in (\underline{\theta}_n, +\infty]$, $\underline{\theta}_n > 0$, is a penalty parameter restraining the maximizing choice of Nature. As noted above θ_n is associated with the Lagrange multiplier of the entropy constraint at each site. In the entropy constraint H_n is the maximum misspecification error that the decision maker is willing to consider given the existing information about the system at site n[4]. The lower bound $\underline{\theta}_n$ is a so-called breakdown point beyond which it is fruitless to seek more robustness because the maximizing agent is sufficiently unconstrained so that she/he can push the criterion function to $+\infty$ despite the best response of the minimizing agent. Thus when $\theta_n < \underline{\theta}_n$ for a specific site robust control rules cannot be attained. In our terminology this site is a candidate for a "nucleus" of a hot spot since misspecification concerns for this site will break down robust control for the whole spatial domain. On the other hand when $\theta_m \to \infty$ or equivalently $H_m = 0$ there are no misspecification concerns for this site and the benchmark model can be used. The effects of spatial connectivity can be seen in this extreme example. The spatial relation of site m with site n breaks down regulation from both sites. If site m was spatially isolated from n there would have been no problem with regulation at m.

Remark 6. Alternative equivalent problems can be formulated. For instance one may consider utility maximization problems in lieu of distance from a target minimization problems. In such cases the agent wishes to maximize her utility while nature, the malevolent player, acts so as to minimize it. This corresponds to maximizing the worst case utility which formally leads to an equivalent problem with the max and the min interchanged. For uniformity and clarity of presentation we work throughout with the distance from a target minimization interpretation of the problem (min / max) and emphasize that all our results may be easily modified to work for the utility maximization interpretations.

3 Translation invariant systems: Closed form solution

In this section we treat a special case of the robust control problem, which allows a solution in closed form. As discussed in Remark 9 the results in this section apply under rather restricted conditions[5] however, the closed form solution allows us to obtain a good intuition concerning the qualitative behavior of the solution, which will guide us in the treatment of the general case in later sections.

Assume that the operators A, B and C are discrete convolution type operators. This is an assumption which essentially states that $a_{nm} = a_{n-m}$, i.e. the effect that a site m has at site n depends only on the distance between n and m and not on the actual positions of the sites. Therefore we assume that the operators A, B and C are translation invariant. This assumption allows us to make a great simplifying step towards the resolution of the problem. We employ the discrete Fourier transform on the lattice $\mathfrak{L}$, denoted by $\mathfrak{F}$. For a detailed account of the Fourier transform the reader may consult the appendix in Section 8.2. The Fourier transform has the property of turning a convolution operator

[4] If the decision maker can use physical principles and statistical analysis to formulate bounds on the relative entropy of plausible probabilistic deviations from her/his benchmark model, these bounds can be used to calibrate the parameters H_n (Athanassoglou and Xepapadeas (2011)).

[5] On the other hand are examples where the translation invariance property is relevant and useful, as in the longitudinal control and string stability of vehicular platoons, with $\mathfrak{L} = \mathbb{Z}$ (e.g. Bamieh et al. (2002) Curtain et al. (2008)). On more general terms translation invariance is quite commonly used in Interacting Particle Systems (IPS) type statistical mechanics models. Structures like IPS models are quite popular in modeling social interactions.

into a multiplication operator, i.e. $\mathfrak{F}(\mathsf{A}u) = \mathfrak{F}(\mathsf{A})\mathfrak{F}(u)$ where by $\mathfrak{F}(\mathsf{A})$ we denote the Fourier transform of the (doubly infinite) matrix A. To ease notation we will use the convention $\hat{u}_k := \mathfrak{F}(u)(k)$ where now k takes values on the dual lattice. A similar notation with the hats will hold for all other involved quantities.

As the rationale for this section is simply to help us develop our intuition, and we plan to consider the problem in full generality in subsequent sections using techniques which are generally applicable, we will make a few more simplifying assumptions. We will assume that our physical space is the finite dimensional lattice $\mathbb{Z}_N$, so that the dynamical system is defined on the Hilbert space $\ell^2(\mathbb{Z}_N)$ and furthermore we restrict our attention to the class of vectors in $\ell^2(\mathbb{Z}_N)$ such that their Fourier transforms are real valued vectors (as seen in the relevant appendix (see Section 8.2 and in particular Example 9). The case of the infinite lattice is considered in the Appendix, Section 8.6.

Definition 1. For $m = 0, \cdots, N-1$ consider the following vectors which are elements of $\ell^2(\mathbb{Z}_N)$:

$$\mathfrak{C}^{(m)} := \Re(\mathfrak{E}^{(m)}) = \left(1, \cos\left(2\pi \frac{m}{N}\right), \cdots, \cos\left(2\pi \frac{nm}{N}\right), \cdots, \cos\left(2\pi \frac{(N-1)m}{N}\right)\right),$$

and define

$$\mathcal{X}_R := span(\mathfrak{C}^{(m)}; m = 0, \cdots, N-1) \subset \ell^2(\mathbb{Z}_N).$$

Remark 7. The space $\mathcal{X}_R$ contains vectors with specific symmetry patterns. For simplicity assume that $N = 2 * n + 1$ is odd. Since $\cos\left(2\pi \frac{(N-r)m}{N}\right) = \cos\left(2\pi \frac{rm}{N}\right)$ for all $r = 1, \cdots, n$, any element x of $\mathcal{X}_R$ is such that $x(0)$ is arbitrary whereas $x(1) = x(N-1)$, $x(2) = x(N-2)$, $\cdots$, $x(n) = x(n+1) = x(2n+1-n)$.

The following lemma is useful:

Lemma 1.

(i) If $x \in \mathcal{X}_R$ then $\Im(\hat{x}) = 0$.

(ii) Let A be a symmetric matrix corresponding to a convolution operator (a circulant matrix) such that the first column of the matrix A, $a^{(1)} \in \mathcal{X}_R$. If $x \in \mathcal{X}_R$ then $\Im(\mathfrak{F}(A \star x)) = 0$.

Proof: The proof of (i) follows immediately from the linearity of the discrete Fourier transform and the properties of the vectors $\mathfrak{C}^{(m)}$. Then (ii) follows from the fact that $\mathfrak{F}(A \star x) = \mathfrak{F}a^{(1)}\,\mathfrak{F}x$, and each of the vectors involved in this product are real valued (by (i)). ■

We are now ready to state the assumptions needed for this section. We emphasize that these assumptions are only used here in order to provide a simple completely worked out example in order to motivate the general discussion that will be developed in subsequent sections of this paper.

Assumption 1.

(i) The operators A, B, C are translation invariant (such that they correspond to discrete convolution operators and are represented by circulant matrices).

(ii) The first column of the matrix representation of these operators are vectors which belong to $\mathcal{X}_R$.

(iii) The initial condition $x_0 \in \mathcal{X}_R$ and the stochastic process $w \in \mathcal{X}_R$.

Remark 8. Out of the above assumptions only (i) is essential for the treatment of the control problem using the Fourier transform. Assumptions (ii) and (iii) are adopted simply to make sure that the resulting dynamical system in Fourier space is real valued and thus facilitate the analysis. The results stated here, e.g., the treatment of the control problem using the Riccati equation by no means is restricted to the real valued case, and can be extended in the case where the resulting dynamical system is complex valued, by simple separation of the real and the imaginary parts. However, this would render the algebra rather involved, obscuring the main points regarding the qualitative behavior of the system, that we wish to stress here.

Example 2. There are many interesting operators arising in realistic models that satisfy Assumption 1 (ii). The discrete Laplacian is an example of such an operator. Furthermore copies of the identity operator are such operators as well. Therefore, an example that falls in this category is the case of system (2) with $\mathsf{A} = \Delta_d$, the discrete Laplacian operator defined (in 1 dimension) as $(\mathsf{A}x)_n = x_{n+1} - 2x_n + x_{n-1}$ and $\mathsf{B} = bI$, $\mathsf{C} = cI$. This leads to diffusive effects on the lattice, but localized control and uncertainty effects. Other options are possible.

Proposition 4. *Let $\mathsf{P} = p\,I$, $\mathsf{Q} = q\,I$, where $I : \ell^2(\mathbb{Z}_N) \to \ell^2(\mathbb{Z}_N)$ is the identity operator. Under Assumption 1 (i) the control system (5) under the dynamic constraint (2) decouples in Fourier space and becomes*

$$\min_{\{\hat{u}_k\}} \max_{\{\hat{v}_k\}} \mathbb{E}_Q \left[\int_0^\infty e^{-rt} \sum_k p(\hat{x}_k(t))^2 + q(\hat{u}_k(t))^2 - \theta \sum_n (\hat{v}_k(t))^2) dt \right]$$

subject to the decoupled state equations

$$d\hat{x}_k(t) = (\hat{a}_k \hat{x}_k(t) + \hat{b}_k \hat{u}_k(t) + \hat{c}_k \hat{v}_k(t))dt + \hat{c}_k \sigma_k \mathfrak{w}_k(t), k \in \mathfrak{L},$$

where $\hat{a}_k$, $\hat{b}_k$, $\hat{c}_k$ are the components of the Fourier transform of the first column of the matrix representation of the operators A, B, C respectively, and

$$\sigma_k^2 := 1 + 4 \sum_{r=1}^{n} \cos^2 \left(2\pi \frac{rk}{N} \right).$$

and $\mathfrak{w}_k$ is a standard Brownian motion under the measure Q.

If furthermore Assumptions 1 (ii) and (iii) hold then the robust control problem

$$\min_{\{\hat{u}_k\} \in \mathcal{X}_R} \max_{\{\hat{v}_k\} \in \mathcal{X}_R} \mathbb{E}_Q \left[\int_0^\infty e^{-rt} \sum_k p(\hat{x}_k(t))^2 + q(\hat{u}_k(t))^2 - \theta \sum_n (\hat{v}_k(t))^2) dt \right] \tag{8}$$

*subject to the dynamic constraints (**??**) admits real valued solutions in Fourier space.*

Proof: Applying the Fourier transform $\mathfrak{F}$ on the equations for the state variables yields

$$d\hat{x}_k(t) = (\hat{a}_k \hat{x}_k(t) + \hat{b}_k \hat{u}_k(t) + \hat{c}_k \hat{v}_k(t))dt + \hat{c}_k \hat{w}_k(t), k \in \mathfrak{L}. \tag{9}$$

Since $w \in \mathcal{X}_R$ it must possess the spatial symmetry of the elements of $\mathcal{X}_R$, therefore it consists of $n+1$ independent Wiener processes $w_0, w_1, \cdots, w_n$ and is of the form

$$w(t) = (w_0, w_1, \cdots, w_n, w_n, w_{n-1}, \cdots, w_1)$$

whose Fourier transform is real valued and equal to the vector $\hat{w}$ with coordinates

$$\hat{w}_k(t) = w_0 + 2 \sum_{r=1}^{n} \cos \left(2\pi \frac{rk}{N} \right)$$

where by a simple application of Lévy's characterization theorem it can be seen that $\{\hat{w}_k\} = \sigma_k \mathfrak{w}$ where $\mathfrak{w}$ is a standard Wiener process with respect to the measure Q and

$$\sigma_k^2 := 1 + 4 \sum_{r=1}^{n} \cos^2 \left(2\pi \frac{rk}{N} \right).$$

The system (??) is now a decoupled system and this greatly simplifies the presentation.

Assuming further that $\mathsf{P} = pI$ and $\mathsf{Q} = qI$ where $I : \ell^2 \to \ell^2$ is the identity operator, so that we may use the Plancherel theorem to restate the control functional with respect to the Fourier transformed variables. According to this result,

$$\sum_n u_n^2 = \frac{1}{N}\sum_k [\mathfrak{F}(u)(k)]^2 = \frac{1}{N}\sum_k \hat{u}_k^2$$

where the first summation takes place in the lattice $\mathfrak{L}$ whereas the second summation takes place in the dual lattice. We have used the fact that we restrict our problem to control variables $u \in \mathcal{X}_R$ so that all the quantities involved in the Plancherel formula are real valued. In a similar fashion we may deal with the other quadratic terms.

Therefore, one may restate the control functional in Fourier space as

$$\mathcal{J} := \frac{1}{N}\,\mathbb{E}_Q\left[\int_0^\infty e^{-rt}\sum_k p(\hat{x}_k(t))^2 + q(\hat{u}_k(t))^2 - \theta\sum_k (\hat{v}_k(t))^2)dt\right]$$

where using the Plancherel theorem we have replaced the summation over the primary lattice with the summation over the dual lattice. Notice that the effects of the size of the lattice (the $\frac{1}{N}$ terms) factor out and have a uniform effect over all the terms of the control functional.

The control problem then becomes

$$\min_{\{\hat{u}_k\}}\max_{\{\hat{v}_k\}}\mathbb{E}_Q\left[\int_0^\infty e^{-rt}\sum_k p(\hat{x}_k(t))^2 + q(\hat{u}_k(t))^2 - \theta\sum_n (\hat{v}_k(t))^2)dt\right] \tag{10}$$

subject to the decoupled state equations (??). ■

The decoupling of the system in Fourier space greatly facilitates its treatment and allows for explicit solutions.

Proposition 5. *The solution of the robust control problem (??) subject to state constraints (??) is equivalent to the solution of the decoupled problems*

$$\min_{\hat{u}_k}\max_{\hat{v}_k}\mathbb{E}_Q\left[\int_0^\infty e^{-rt} p(\hat{x}_k(t))^2 + q(\hat{u}_k(t))^2 - \theta\sum_k (\hat{v}_k(t))^2)dt\right], \quad k \in \mathfrak{L} \tag{11}$$

subject to the state constraint (??) for each individual $k \in \mathfrak{L}$.

Proof: The proof is straightforward and is omitted. ■

Remark 9. The Fourier approach considered here is rather limited as to the class of problems it is applicable to. The two major limitations are (a) the assumption that the operators involved are translation invariant and (b) the use of the Plancherel formula (isometry) to turn the control functional from a mapping of the primal lattice to a mapping of the dual lattice. The second requirement limits considerably the type of control problems we are allowed to treat in this manner. As a result, only minor generalizations of the results of this section are allowed to systems of more general forms. For instance, under further restrictions on the operators, one could treat using Fourier transforms the localized entropic constrained problem introduced in Proposition 3 by defining the new variables $\bar{v}_n = \sqrt{\theta_n}v_n$ and rewriting the functional into a form where the Plancherel isometry holds. This is equivalent to transforming v into $\bar{v} = \mathsf{D}v$, where D is a diagonal operator with representation $d_{nm} = \sqrt{\theta_n}\delta_{nm}$. However, this transformation changes the state equation as well, therefore care should be taken so that the operator $\mathsf{C}\mathsf{D}^{-1}$ and C are at the same time translation invariant, so that the Fourier transform of the state equation is also diagonal. This remark shows the difficulties in generalizing the Fourier transform approach to systems of more general form. These difficulties are overcome in Section 5 where the general linear quadratic problem is treated via a different approach and not through the Fourier transform.

We will consider the solution of the above problems (both primal and dual) using dynamic programming techniques, through the use of the Hamilton-Jacobi-Bellman-Isaacs (HJBI) equation (see e.g. Isaacs (1999), Hansen and Sargent (2001), Anderson et al. (2003) and references therein).

Proposition 6 (Solution of primal problem)**.** *The solution of the primal problem*

$$\min_{\hat{u}_k} \max_{\hat{v}_k} \mathbb{E}_Q \left[\int_0^\infty e^{-rt} p(\hat{x}_k(t))^2 + q(\hat{u}_k(t))^2 - \theta(\hat{v}_k(t))^2) dt \right], \tag{12}$$

subject to (??) for each k is given by the optimal state equation

$$d\hat{x}_k^* = R_k \hat{x}_k^* dt + \hat{c}_k \sigma_k d\mathfrak{w}_k$$

where

$$R_k := \hat{a}_k - \frac{\hat{b}_k^2 M_{2,k}}{2q} + \frac{\hat{c}_k^2 M_{2,k}}{2\theta}.$$

and $M_{2,k}$ is the solution of

$$\left(\frac{\hat{c}_k^2}{2\theta} - \frac{\hat{b}_k^2}{2q} \right) M_{2,k}^2 + (2\,\hat{a}_k - r)\, M_{2,k} + 2p = 0. \tag{13}$$

The optimal controls are given by the feedback laws

$$\hat{u}_k^* = -\frac{\hat{b}_k M_{2,k}}{2q} \hat{x}_k^*, \qquad \hat{v}_k^* = \frac{\hat{c}_k\, M_{2,k}}{2\theta} \hat{x}_k^*.$$

Proof: Fix $k \in \mathbb{Z}$ and let V_k be the value function corresponding to this choice.

Let $\mathcal{L}_k : C^2(\mathbb{R}) \to C(\mathbb{R})$ be the generator operator of the diffusion process $\{\hat{x}_k(t)\}$, $t \in \mathbb{R}_+$ defined by

$$(\mathcal{L}_k \Phi)(\hat{x}_k) = (\hat{a}_k \hat{x}_k + \hat{b}_k \hat{u}_k + \hat{c}_k \hat{v}_k) \frac{\partial \Phi}{\partial \hat{x}_k} + \frac{1}{2} \hat{c}_k^2 \sigma_k^2 \frac{\partial^2 \Phi}{\partial \hat{x}_k^2}.$$

The relevant Hamilton-Jacobi-Belman-Isaacs (HJBI) equation becomes

$$r V_k = \bar{H} \left(\hat{x}_k, \frac{\partial V_k}{\partial \hat{x}_k}, \frac{\partial^2 V_k}{\partial \hat{x}_k^2} \right)$$

where for any function Φ of sufficient regularity the Hamiltonian $\bar{H}\left(\hat{x}_k, \frac{\partial \Phi}{\partial \hat{x}_k}, \frac{\partial^2 \Phi}{\partial \hat{x}_k^2}\right)$ is defined by

$$\bar{H} \left(\hat{x}_k, \frac{\partial \Phi}{\partial \hat{x}_k}, \frac{\partial^2 \Phi}{\partial \hat{x}_k^2} \right) := \inf_{\hat{u}_k} \sup_{\hat{v}_k} \left(p\hat{x}_k^2 + q\hat{u}_k^2 - \theta \hat{v}_k^2 + \mathcal{L}_k \Phi \right),$$

and the optimization problems in the definition of the Hamiltonian are considered as static optimization problems over $\hat{v}_k \in \mathcal{V}_k \subset \mathbb{R}$, $\hat{u}_k \in \mathcal{U}_k \subset \mathbb{R}$, for fixed k, where $\mathcal{U}_k$, $\mathcal{V}_k$ are appropriate subsets of $\mathbb{R}$ [6].

We first calculate $\bar{H}(x, \Phi_x, \Phi_{xx})$ for any function Φ, where we use the shorthand notation $\Phi_x = \frac{\partial \Phi}{\partial \hat{x}_k}$ and $\Phi_{xx} = \frac{\partial^2 \Phi}{\partial \hat{x}_k^2}$ for simplicity.

[6] Or of $\mathbb{R}^d$ without loss of generality.

The solution of the static optimization problem is given by the first order condition $\hat{v}_k^* = \frac{\hat{c}_k}{2\theta}\Phi_x$. This corresponds to a maximum value which becomes

$$\Psi := \frac{\hat{c}_k^2 \sigma_k^2 \Phi_{xx}}{2} + q\,\hat{u}_k^2 + p\,\hat{x}_k^2 + b\,\Phi_x\,\hat{u}_k + \hat{a}_k\,\Phi_x\,\hat{x}_k + \frac{\hat{c}_k^2 \Phi_x^2}{4\theta}.$$

We now minimize the function Ψ with respect to $\hat{u}_k$. The first order condition for the minimum gives

$$\hat{u}_k = -\frac{\hat{b}_k\,\Phi_x}{2\,q}$$

which upon substitution gives

$$\bar{H}(x, \Phi_x, \Phi_{xx}) = \frac{\hat{c}_k^2 \sigma_k^2\,\Phi_{xx}}{2} + p\,\hat{x}_k^2 + \hat{a}_k\,\Phi_x\,\hat{x}_k - \frac{\hat{b}_k^2 \Phi_x^2}{4\,q} + \frac{\hat{c}_k^2\,\Phi_x^2}{4\theta}.$$

The HJBI equation thus assumes the form

$$\frac{\hat{c}_k^2 \sigma_k^2\,V_{xx}}{2} + p\,\hat{x}_k^2 + \hat{a}_k\,V_x\,\hat{x}_k - \frac{\hat{b}_k^2 V_x^2}{4\,q} + \frac{\hat{c}_k^2\,V_x^2}{4\theta} = rV$$

which is a nonlinear second order differential equation.

We look for a solution of the special form

$$V(\hat{x}_k) = \frac{M_{2,k}}{2}\hat{x}_k^2 + M_{1,k}\hat{x}_k + M_{0,k}.$$

Substituting into the HJBI equation and matching coefficients of different orders of $\hat{x}_k$ we obtain that the coefficient $M_{i,k}$, $i = 0, 1, 2$ are given by

$$\begin{aligned}
&\left(\frac{\hat{c}_k^2}{2\theta} - \frac{\hat{b}_k^2}{2q}\right) M_{2,k}^2 + (2\,\hat{a}_k - r)\,M_{2,k} + 2p = 0,\\
&M_{1,k} = 0,\\
&M_{0,k} = \frac{\hat{c}_k^2 \sigma_k^2 M_{2,k}}{2r}.
\end{aligned}$$

The value function is thus obtained as long as the solution of the above quadratic equation is obtained.

We now substitute this expression into the equations for the optimal controls to obtain

$$\hat{v}_k^* = \frac{\hat{c}_k\,M_{2,k}}{2|\theta|}\hat{x}_k, \quad \hat{u}_k^* = -\frac{\hat{b}_k M_{2,k}}{2q}\hat{x}_k.$$

Therefore, upon substitution into the state equation we see that the optimal state is given by the solution of the stochastic differential equation

$$d\hat{x}_k^* = R_k \hat{x}_k^* dt + \hat{c}_k d\hat{w}_k$$

where

$$R_k := \hat{a}_k - \frac{\hat{b}_k^2 M_{2,k}}{2q} + \frac{\hat{c}_k^2\,M_{2,k}}{2\theta}$$

This completes the proof. ∎

Proposition 7 (Solution of the dual problem)**.** *The solution of problem*

$$\max_{\hat{v}_k} \min_{\hat{u}_k} \mathbb{E}_P[\int_0^\infty e^{-rt} p(\hat{x}_k(t))^2 + q(\hat{u}_k(t))^2 - \theta(\hat{v}_k(t))^2)dt], \;\; k \in \mathfrak{L} \tag{14}$$

*subject to (**??**) for any $k \in \mathbb{Z}$ coincides with that of the primal problem (12) as given by Proposition 6 and there is no duality gap.*

Proof: The value function of the dual problem $V^\sharp$ satisfies the Hamilton-Jacobi-Bellman equation

$$\underline{H}(\hat{x}_k, V^\sharp_x, V^\sharp_{xx}) = rV^\sharp$$

where for any function Φ, of sufficient regularity, the Hamiltonian $\underline{H}(x, \Phi_x, \Phi_{xx})$ is defined by

$$\underline{H}(x, \Phi_x, \Phi_{xx}) := \sup_{\hat{v}_k} \inf_{\hat{u}_k} \{p\hat{x}_k^2 + q\hat{u}_k^2 - \theta\hat{v}_k^2 + \mathcal{L}_k\Phi\}$$

where $\mathcal{L}_k$ is the generator operator of the diffusion process, defined by (**??**). As before the optimization problem is a static one. A quick calculation shows that for any function Φ,

$$\underline{H}(x, \Phi_x, \Phi_{xx}) = \bar{H}(x, \Phi_x, \Phi_{xx})$$

thus leading to the same HJBI equation as for the primal problem. The result then follows retracing the steps in the proof of Proposition 6. ■

Remark 10. The above two propositions simply provide candidates for the solution of the problem. Whether these candidates are indeed solutions and whether the solution is a saddle point depends on the choice of the parameter θ, as will become clear in the next section.

Remark 11 (Certainty equivalent)**.** Suppose that instead of the stochastic problem treated here we treat instead the control problem with the deterministic state equation

$$d\hat{x}_k = (\hat{a}_k\hat{x}_k + \hat{b}_k\hat{u}_k + \hat{c}_k\hat{v}_k)dt$$

and the same quadratic cost functional (where of course now the expectation is redundant). This is a deterministic linear quadratic optimal control system. The solution of the relevant robust control problem is governed by the Hamilton-Jacobi-Isaacs equation, where the operator $\mathcal{L}_k$ is now replaced by the first order operator $\mathcal{L}_k^0$ with the following action on the value function: $\mathcal{L}_k^0 V = (\hat{a}_k\hat{x}_k + \hat{b}_k\hat{u}_k + \hat{c}_k\hat{v}_k)V_x$. Then, working in the same fashion as in Propositions 6 and 7 we see that the optimal policy for the deterministic problem **coincides** with that of the stochastic problem. Therefore, as far as the form of the feedback law of the optimal policies are concerned stochastic effects play no role. This has been called by Hansen and Sargent the certainty equivalent. However, one should be extremely cautious with that, since in the stochastic case the optimal policy is a stochastic process (through the dependence of u^* and v^* on x which is a stochastic process. On the contrary, the optimal policy in the deterministic case is deterministic (through the dependence of u^* and v^* on x which is a deterministic process). This qualitative behavior shows in the calculation of the value function, which for the certainty equivalent problem works out to be $V^0 = \frac{M_{2,k}}{2}\hat{x}_k^2 < V$.

4 Hot spot formation in translation invariant systems

In this section we study the validity and the qualitative behavior of the controlled system. We will call the qualitative changes of the behavior of the system **hot spots**. We will define three types of hot spots:

- ▷ **Hot spot of type I:** This is a breakdown of the solution procedure, i.e., a set of parameters where a solution to the above problem does not exist.
- ▷ **Hot spot of type II:** This corresponds to the case where the solution exists but may lead to spatial pattern formation, i.e., to spatial instability similar to the Turing instability.
- ▷ **Hot spot of type III:** This corresponds to the case where the cost of robustness becomes more that what is offering us, i.e., where the relative cost of robustness may become very large.

In what follows we discuss the formation of hot spots in the case of finite lattices $\mathbb{Z}_N$; the mechanism for hot spot formation in the infinite lattice is similar and certain remarks will be made when necessary.

4.1 Hot spots of type I

The breakdown of the solution procedure can be seen quite easily by the following simple argument. As seen in the proof of Proposition 6 the value function assumes a simple quadratic form, as long as the algebraic quadratic equation

$$\left(\frac{\hat{c}_k^2}{2\theta} - \frac{\hat{b}_k^2}{2q}\right) M_{2,k}^2 + (2\,\hat{a}_k - r)\, M_{2,k} + 2p = 0. \tag{15}$$

admits real valued solutions, at least one of which is positive. The positivity of the real root is needed since, by general considerations in optimal control, the value function must be convex. If the above algebraic quadratic equation does not admit at least one positive real valued solution this is an indication of breakdown of the existence of a solution to the robust control problem which will be called a hot spot of Type I.

Proposition 8 (Type I hot spot creation:)**.** *Hot spots of Type I may be created in one of the following two cases:*

(I_A) Either,

$$(2\hat{a}_k - r)^2 < 8\,p\left(\frac{\hat{c}_k^2}{2\theta} - \frac{\hat{b}_k^2}{2q}\right), \tag{16}$$

(I_B) Or,

$$(2\hat{a}_k - r)^2 > 8\,p\left(\frac{\hat{c}_k^2}{2\theta} - \frac{\hat{b}_k^2}{2q}\right), \quad \left(\frac{\hat{c}_k^2}{2\theta} - \frac{\hat{b}_k^2}{2q}\right) > 0, \;\; 2\hat{a}_k - r > 0. \tag{17}$$

Hot spots of this type may arise either due to low values of θ, or due to high values of q or low values of r. For example, they may arise either if

$$\theta < \frac{p\hat{c}_k^2}{\left(\hat{a}_k - \frac{r}{2}\right)^2 + \frac{p}{q}\hat{b}_k^2}, \quad k \in \mathbb{Z}_N.$$

or if

$$\theta > \frac{p\hat{c}_k^2}{\left(\hat{a}_k - \frac{r}{2}\right)^2 + \frac{p}{q}\hat{b}_k^2}, \quad \frac{q}{\theta} > \frac{\hat{b}_k^2}{\hat{c}_k^2}, \quad r < 2\hat{a}_k, \quad k \in \mathbb{Z}_N.$$

In particular hot spots are expected to occur in the limit as $\theta \to 0$ while they are not expected to occur in the limit as $\theta \to \infty$.

Proof: Let us rewrite the above equation in the simpler form

$$M_{2,k}^2 + \frac{\mathfrak{a}}{\mathfrak{R}} + \frac{2p}{\mathfrak{R}} = 0$$

where

$$\mathfrak{a} := 2\hat{a}_k - r, \;\; \mathfrak{R} := \frac{\hat{c}_k^2}{2\theta} - \frac{\hat{b}_k^2}{2q}.$$

assuming of course that $\mathfrak{R} \neq 0$. The condition for having two real solutions of equation (15) is that the discriminant is positive,

$$\Delta := \left(\frac{\mathfrak{a}}{\mathfrak{R}}\right)^2 - 8\frac{p}{\mathfrak{R}} > 0$$

which is equivalent to the condition $\mathfrak{a}^2 > 8\,p\,\mathfrak{R}$. If $\mathfrak{R} < 0$ this condition always holds, whereas if $\mathfrak{R} > 0$ is will hold for particular modes which of course depend on the parameter values. We now look for the solutions. We will consider two cases and the relevant subcases:

- **Case A** $\mathfrak{R} < 0$. Then $\sqrt{\Delta} \geq |\mathfrak{a}/\mathfrak{R}|$.
- **Case A1** If $\mathfrak{a} > 0$ then there is only one positive solution which is

$$M_{2,k} = \frac{1}{2}\left(-\frac{\mathfrak{a}}{\mathfrak{R}} + \sqrt{\Delta}\right)$$

- **Case A2** if $\mathfrak{a} < 0$ then there is only one positive solution (which is exactly of the same form as in Case A1),

$$M_{2,k} = \frac{1}{2}\left(-\frac{\mathfrak{a}}{\mathfrak{R}} + \sqrt{\Delta}\right)$$

- **Case B** $\mathfrak{R} > 0$. Then $\sqrt{\Delta} \leq \mathfrak{a}/\mathfrak{R}|$.
- **Case B1** If $\mathfrak{a} > 0$ then there is two negative solutions none of which is acceptable on account of loss of convexity of the value function, therefore this is a hot spot of Type I.
- **Case B2** if $\mathfrak{a} < 0$ then there as two positive solutions

$$0 \leq M_{2,k}^{(1)} = \frac{1}{2}\left(-\frac{\mathfrak{a}}{\mathfrak{R}} - \sqrt{\Delta}\right) \leq M_{2,k}^{(2)} = \frac{1}{2}\left(-\frac{\mathfrak{a}}{\mathfrak{R}} + \sqrt{\Delta}\right).$$

  Out of these two we should keep the smaller one $M_{2,k}^{(1)}$ which of course will give the minimum value function (the other choice will correspond to a "second best" or suboptimal solution).

Therefore, summarizing a hot spot of Type I may arise if

(I_A) Either, $\mathfrak{a}^2 < 8\,p\,\mathfrak{R}$

(I_B) Or, $\mathfrak{a}^2 > 8\,p\,\mathfrak{R}$, $\mathfrak{R} > 0$ and $\mathfrak{a} > 0$.

or in terms of the original notation if

(I_A) Either,

$$(2\hat{a}_k - r)^2 < 8\,p\left(\frac{\hat{c}_k^2}{2\theta} - \frac{\hat{b}_k^2}{2q}\right), \tag{18}$$

(I_B) Or,

$$(2\hat{a}_k - r)^2 > 8\,p\left(\frac{\hat{c}_k^2}{2\theta} - \frac{\hat{b}_k^2}{2q}\right), \quad \left(\frac{\hat{c}_k^2}{2\theta} - \frac{\hat{b}_k^2}{2q}\right) > 0, \;\; 2\hat{a}_k - r > 0. \tag{19}$$

Conditions (18) and (19) for the occurrence of hot spots of Type I are multiparameter conditions which are easily checked once a particular system is selected for study but when trying to infer general qualitative aspects concerning the optimal path there is not a simple or unique way of interpreting them. They may hold for some $k \in \mathbb{Z}$, meaning that the robust control procedure will break down in the particular site of the dual lattice, thus bringing down the successful control procedure of the whole system due to the coupling effects. One may call that mechanism a transmission of breakdown. Another way to look at these conditions is to use them as selection criteria for the parameters of the system not related to the operators A, B, C for which a hot spot will definitely occur. For example case I_A can be translated to

$$\theta < \frac{p\hat{c}_k^2}{\left(\hat{a}_k - \frac{r}{2}\right)^2 + \frac{p}{q}\hat{b}_k^2}, \quad k \in \mathbb{Z}_N.$$

which will hold for every site in the dual lattice as long as

$$\theta < \theta_{cr} := \min_{k \in \mathbb{Z}_N} \left\{ \frac{p \hat{c}_k^2}{\left(\hat{a}_k - \frac{r}{2}\right)^2 + \frac{p}{q} \hat{b}_k^2} \right\}.$$

Therefore, if θ is too small, smaller than the critical value θ_{cr} then the robust control mechanism breaks down and Type I hot spots will certainly occur. However, this is not the only possible case. As Case I_B shows, if

$$\theta > \frac{p \hat{c}_k^2}{\left(\hat{a}_k - \frac{r}{2}\right)^2 + \frac{p}{q} \hat{b}_k^2}, \quad q > \theta \, \frac{\hat{b}_k^2}{\hat{c}_k^2}, \quad r < 2\hat{a}_k, \quad k \in \mathbb{Z}_N,$$

then a hot spot of Type I will arise. This will be true if

$$\theta > \max_{k \in \mathbb{Z}_N} \left\{ \frac{p \hat{c}_k^2}{\left(\hat{a}_k - \frac{r}{2}\right)^2 + \frac{p}{q} \hat{b}_k^2} \right\}, \quad q > \max_{k \in \mathbb{Z}_N} \left\{ \theta \, \frac{\hat{b}_k^2}{\hat{c}_k^2} \right\}, \quad r < 2 \min_{k \in \mathbb{Z}_N} \{\hat{a}_k\},$$

meaning that hot spots of Type I may also arise for high values of θ if either the cost of control is high or is the discount factor is low enough. ■

Remark 12 (Hot spot of type I and loss of convexity)**.** As mentioned above, a hot spot of Type I represents breakdown of the solvability of the optimal control problem. We argue that this represents some sort of loss of convexity of the problem thus leading to non existence of solution. To illustrate this point more clearly let us take the limit as $\theta \to 0$ which corresponds to hot spot formation. For such values of θ, the particular ansatz employed for the solution breaks down and in fact as $\theta \to 0$ we expect $M_{2,k} \to 0$ so that the quadratic term in the value function will disappear. This leads to loss of strict concavity of the functional, which may be seen as follows: The functional contains a contribution from $\hat{v}_k$ through the dependence of $\hat{x}_k$ on $\hat{v}_k$ which contributes a quadratic term of positive sign in $\hat{v}_k$. The robustness term, which is proportional to $-\theta$ contributes a quadratic term of negative sign in $\hat{v}_k$. For large enough values of θ the latter term dominates in the functional and guarantees the strict concavity, therefore, leading to a well defined maximization problem. In the limit of small θ the former term dominates and thus turn the functional into a convex functional leading to problems with respect to the maximization problem over $\{\hat{v}_k\}$. We call this breakdown of concavity in v, which lead to loss of convexity of the value function in x, for small values of θ a hot spot of type I. When this happens, there is a duality gap, since the assumptions of the min-max theorem do not hold. In terms or regulatory objectives this means that concerns about model misspecification make regulation impossible.

The following examples show some interesting limiting situations:

Example 3. Assume that A is the discrete Laplacian whereas B and C are copies of the identity operator. This corresponds to the case that there is diffusive coupling in the state equation but controls as well as the uncertainty have purely localized effects. A quick calculation shows that in this case $a_k = \alpha\left(1 + 2\cos\left(\frac{2\pi k}{N}\right)\right)$ where α is the diffusion coefficient whereas $b_k = \beta$ and $c_k = \gamma$ for every $k \in \mathbb{Z}_N$ where β and γ is a measure for the control and the uncertainty respectively. In this particular case, the quadratic equation becomes

$$\left(\frac{\gamma^2}{2\theta} - \frac{\beta^2}{2q}\right) M_{2,k}^2 + \left(2\,\alpha\left(1 + 2\cos\left(\frac{2\pi k}{N}\right)\right) - r\right) M_{2,k} + 2p = 0.$$

which must have a real valued solution for every k. There will not exist real valued solutions if

$$\Delta := \left(2\,\alpha\left(1 + 2\cos\left(\frac{2\pi k}{N}\right)\right) - r\right)^2 - 8\,p\left(\frac{\gamma^2}{2\theta} - \frac{\beta^2}{2q}\right) < 0$$

or equivalently after some algebra

$$\left(\left(1+2\cos\left(\frac{2\pi k}{N}\right)\right)^2 - \frac{r}{2\alpha}\right) < \frac{p}{\alpha^2}\left(\frac{\gamma^2}{\theta} - \frac{\beta^2}{q}\right).$$

This is the condition for generation of a hot spot of Type I in this particular example. If this condition holds for some $k \in \mathbb{Z}_N$, this particular k is a candidate for such a hot spot. We may spot directly that this cannot hold for any $k \in \mathbb{Z}_N$ if the right hand side of this inequality is negative, i.e., when $\theta > \theta_{cr} := q\frac{\gamma^2}{\beta^2}$, therefore hot spots of this type will never occur for large enough values of θ. The critical value of θ for the formation of such hot spots will depend on the relative magnitude of uncertainty over control . For $\theta < \theta_{cr}$ then a hot spot of Type I may occur for the modes k such that

$$\left(1+2\cos\left(\frac{2\pi k}{N}\right)\right)^2 \leq \frac{r}{2\alpha} + \rho$$

or equivalently for k such that

$$\left(1+2\cos\left(\frac{2\pi k}{N}\right)\right)^2 \leq \left(\frac{r}{2\alpha} + \rho\right)^{\frac{1}{2}}$$

where $\rho^2 = \frac{p}{\alpha^2}\left(\frac{\gamma^2}{\theta} - \frac{\beta^2}{q}\right)$.

Example 4. The opposite case is when A is again the discrete Laplacian while B and C are multiples of matrices containing 1 in the diagonal and the same entry ν in every other position. This means that the controls as well as the uncertainty has a globalized effect to all lattice points, in the sense that the controls even at remote lattice sites have an effect at each lattice point. Then $\hat{b}_k = \beta\,\delta_{k,0}$, $\hat{c}_k = \gamma\,\delta_{k,0}$, i.e., the Fourier transform is fully localized and is a delta function. Then, for $k = 0$ the quadratic equation becomes

$$\left(\frac{\gamma^2}{2\theta} - \frac{\beta^2}{2q}\right) M_{2,0}^2 - (6\alpha - r)M_{2,0} + 2p = 0$$

while for $k \neq 0$ the quadratic term vanishes yielding

$$-\left(2\alpha\left(1+2\cos\left(\frac{2\pi k}{N}\right)\right) - r\right) M_{2,0} + 2p = 0$$

4.2 Hot spots of type II

We now consider the spatial behavior of the optimal path, as given by the Itō stochastic differential equation

$$d\hat{x}_k^* = R_k \hat{x}_k^* dt + \hat{c}_k d\hat{w}_k \tag{20}$$

The optimal path is a random field, thus leading to random patterns in space, some of which may be short lived and generated simply by the fluctuations of the Wiener process. We thus look for the spatial behavior of the mean field as describable by the expectation $\hat{X}_k := \mathbb{E}_Q[\hat{x}_k^*]$. By standard linear theory $\hat{X}_k(t) = \hat{X}_k(0)\exp(R_k t)$ and this means that for the modes $k \in \mathbb{Z}_N$ such that $R_k \geq 0$ we have temporal growth and these modes will dominate the long term temporal behavior. On the contrary modes k such that $R_k < 0$ decay as $t \to \infty$ therefore such modes correspond to (short term) transient temporal behavior, not likely to be observable in the long term temporal behavior. The above discussion implies that the long time asymptotic of the solution in Fourier space will be given by

$$\hat{X}_k(t) \simeq \begin{cases} \hat{x}_k(0)\,\exp(R_k t), & k \in \mathcal{P} := \{k \in \mathbb{Z}_N \,:\, R_k \geq 0\} \\ 0 & \text{otherwise} \end{cases}$$

To see what this pattern will look like in real space, we simply need to invert the Fourier transform, thus obtaining a spatial pattern of the form

$$X_n(t) := \mathbb{E}_Q[x_n(t)] = \sum_{k \in \mathcal{P}} \hat{x}_k(0) \exp(R_k t) \cos\left(2\pi \frac{k}{N} n\right).$$

The above discussion therefore leads us to a very important conclusion, which is of importance to economic theory of spatially interconnected systems:

> If as an effect of the robust optimal control procedure exerted on the system there exist modes $k \in \mathbb{Z}_N$ such that $R_k > 0$, then this will lead to spatial pattern formation which will create spatial patterns of the form (**??**). As we will see there are cases what such patterns will not exist in the uncontrolled system and will appear as an effect of the control procedure. We will call such patterns an **optimal robustness induced spatial instability** or hot spot of Type II.

The economic significance of this result should be stressed. We show the emergence of a spatial pattern formation instability, which can be triggered by the optimal control procedures exerted on the system; in other words emergence of spatial clustering and agglomerations in the economy caused by uncertainty aversion and robust control. This observation can further be extended in the case of nonlinear dynamics, in the weakly nonlinear case. When the dynamics are nonlinear and the state the emergence of hot spots of Type II and optimal robustness induced spatial instability should be linked to the spatial instability of a spatially uniform steady state corresponding to the linear quadratic approximation of a nonlinear system. This instability which can be thought as pattern formation precursor will induce the emergence of spatial clustering. As time progresses and the linearized solution (**??**) grows beyond a certain critical value (in terms of a relevant norm) then the deviation from the homogeneous steady state is so large that the linearized dynamics are no longer a valid approximation. Then the nonlinear dynamics will take over and as an effect of that some of the exponentially growing modes will be balanced thus leading to more complicated stable patterns. At any rate even in the nonlinear case the mechanism described here will be a Turing type pattern formation mechanism explaining the onset of spatial patterns in the economy. For more details concerning the connection (similarities and differences) of the optimal robustness induced instability with the Turing mechanism for pattern formation see Remark 17. For a full discussion of the nonlinear case (beyond the particular case of translation invariant operators) see Sections 6 and 8.8.

The next proposition identifies which modes can lead to hot spot of Type II formation (optimal robustness induced spatial instability) and in this way through equation (**??**) identifies possible spatial patterns that can emerge in the spatial economy.

Proposition 9 (Pattern formation for the primal problem)**.** *There exist pattern formation behavior for the primal problem if there exist modes k such that $R_k > 0$, i.e., if there exist modes k such that*

$$\frac{1}{2}\left(r - \sqrt{r^2 + 8\,p\left(\frac{\hat{c}_k^2}{2\theta} - \frac{\hat{b}_k^2}{2q}\right)}\right) \leq \hat{a}_k \leq \frac{1}{2}\left(r + \sqrt{r^2 + 8\,p\left(\frac{\hat{c}_k^2}{2\theta} - \frac{\hat{b}_k^2}{2q}\right)}\right),$$
$$r^2 + 8\,p\left(\frac{\hat{c}_k^2}{2\theta} - \frac{\hat{b}_k^2}{2q}\right) \geq 0. \tag{21}$$

Proof: The expectation $X_k^* := \mathbb{E}_Q[x_k^*]$, $k \in \mathbb{Z}$ of the optimal path follows that linear deterministic ODE $dX_k^*(t) = R_k X_k^*(t)\,dt$, $k \in \mathbb{Z}$. Thus pattern formation occurs for these $k \in \mathbb{Z}$ such that $R_k > 0$. Let us now try to express R_k in a form which reveals in a more clear fashion the actual dynamics of the optimal path. We will use the notation of Section 4.1, introducing again the quantities $\mathfrak{a}$ and

$\mathfrak{R}$ (see equation (??)). In terms of these quantities we rewrite $R_k = \hat{a}_k + \mathfrak{R}\, M_{2,k}$ where $M_{2,k}$ is the positive solution of the quadratic equation (of course we assume that we do not have occurrence of a Type I hot spot). We are limited in cases $A1$, $A2$ and $B2$ of Section 4.1. A quick calculation shows that in cases $A1$ and $A2$,

$$R_k = \frac{1}{2}(r + \mathfrak{R}\sqrt{\Delta})$$

whereas in case $B2$,

$$R_k = \frac{1}{2}(r - \mathfrak{R}\sqrt{\Delta})$$

where

$$\Delta := \left(\frac{\mathfrak{a}}{\mathfrak{R}}\right)^2 - 8\frac{p}{\mathfrak{R}} > 0.$$

Note that in any case $R_k < \frac{r}{2}$ which is of course expected since on the optimal path the functional is finite therefore possible exponential growth of a mode cannot exceed $e^{\frac{r}{2}t}$. Thus pattern formation type behavior in the optimal path will correspond to cases where $0 \leq R_k < \frac{r}{2}$. Since the right hand side of the inequality always holds, we just consider the left hand side. A simple but tedious algebraic calculation (nor reproduced here) shows that in any of the above cases $R_k \geq 0$ implies $\mathfrak{a}^2 \leq r^2 + 8\,p\,\mathfrak{R}$ (which can only hold as long as $r^2 + 8\pi\,\mathfrak{R} \geq 0$) and this is equivalent to

$$\frac{1}{2}\left(r - \sqrt{r^2 + 8\,p\,\mathfrak{R}}\right) \leq \hat{a}_k \leq \frac{1}{2}\left(r - +\sqrt{r^2 + 8\,p\,\mathfrak{R}}\right).$$

Thus modes satisfying this condition will lead to pattern formation. This condition translated to the original parameters of the problem yields the pattern formation condition

$$\frac{1}{2}\left(r - \sqrt{r^2 + 8\,p\left(\frac{\hat{c}_k^2}{2\theta} - \frac{\hat{b}_k^2}{2q}\right)}\right) \leq \hat{a}_k \leq \frac{1}{2}\left(r + \sqrt{r^2 + 8\,p\left(\frac{\hat{c}_k^2}{2\theta} - \frac{\hat{b}_k^2}{2q}\right)}\right),$$

$$r^2 + 8\,p\left(\frac{\hat{c}_k^2}{2\theta} - \frac{\hat{b}_k^2}{2q}\right) \geq 0$$

■

It is interesting to see what is the behavior of the system as a function of parameters with respect to pattern formation and the qualitative behavior of the optimal path.

Remark 13 (Pattern formation and the discount factor)**.** Note that this pattern formation behavior is in full accordance with the fact that our state equation is the optimal path for the linear quadratic control problem. Since it solves this problem it is guaranteed that $I := \mathbb{E}_Q[\int_0^\infty e^{-rt}\hat{x}_k^2(t)dt]$ is finite therefore $\hat{x}_k(t)$ can at most grow as $e^{\frac{r}{2}t}$, otherwise the quantity I would be infinite. This is verified explicitly in the proof of the above proposition where it is shown that $R_k \leq \frac{r}{2}$ for every $k \in \mathbb{Z}_N$. Therefore, all possible patterns may at most exhibit growth rates less or equal to $r/2$. In the limit as $r \to 0$ i.e. in the limit of small discount rates pattern formation is becoming increasingly difficult in the linear quadratic model since growing patterns will be suppressed by the control procedures.

Proposition 10 (Stabilizing or destabilizing effects of control)**.** *The robust control procedure may either have a stabilizing or destabilizing effect with respect to pattern formation. in the sense that they may either stabilize an unstable mode of the uncontrolled system or on the contrary facilitate the onset of instabilities.*

In particular,

(i) If $\frac{q}{\theta} < \frac{\hat{b}_k^2}{\hat{c}_k^2}$ then the robust control procedure has a stabilizing effect

(ii) If $\frac{q}{\theta} > \frac{\hat{b}_k^2}{\hat{c}_k^2}$ then the robust control procedure has a destabilizing effect

In case (ii) we may then talk about robust control caused pattern formation, in the sense that we obtain a growing mode leading to a pattern which would not have appeared in the uncontrolled system.

Proof: To support the above claim we need to compare the threshold in $\hat{a}_k$ for the onset of instability in the uncontrolled system $a^{(0)}_{k,cr}$ and the relevant quantity $a^{(c)}_{k,cr}$. Of course for the uncontrolled system $a^{(0)}_{k,cr}$, whereas by Proposition 9 we see that $a^{(0)}_{k.cr} := \frac{1}{2}\left(r - \sqrt{r^2 + 8\,p\left(\frac{\hat{c}_k^2}{2\theta} - \frac{\hat{b}_k^2}{2q}\right)}\right)$. If $a^{(c)}_{k,cr} > a^{(0)}_{k,cr} = 0$ then control has a stabilizing effect over mode k (since the effect of control is to make it more difficult for this mode to develop instability by raising the instability threshold) while if on the contrary $a^{(c)}_{k,cr} < a^{(0)}_{k,cr} = 0$, then control has a de-stabilizing effect over mode k (since the effect of control is to make it easier for this mode to develop instability by lowering the instability threshold). Then by simple algebra the claim arises. ■

Remark 14. However, it should be emphasized that for any parameter values the robust control imposes an upper bound of $\frac{r}{2}$ for any unstable mode (see also Remark 13) whereas this does not hold for the uncontrolled system, for which unstable modes may have any growth rate, determined purely by the spectrum of A (the largest growth rate will correspond to the largest positive eigenvalue of A in the finite dimensional case). Therefore, the effect of control on a mode which is unstable will be to temper its growth rate and "trim" it to the maximum value $\frac{r}{2}$.

Remark 15 (The $\theta \to \infty$ limit)**.** As seen by Proposition 10 in the $\theta \to \infty$ limit, the control has a stabilizing effect on unstable modes of the uncontrolled system.

Remark 16 (The $\theta \to 0$ limit)**.** Similarly, by Proposition 10 in the $\theta \to 0$ limit, the robust control has a destabilizing effect on modes of the uncontrolled system which are "marginal" to be stable i.e. with $\hat{\alpha}_k$ negative but close to zero.

Remark 17 (Similarities and Differences with Turing instability)**.** This is similar to Turing instability leading to pattern formation but with a very important difference! In contrast to Turing instability which is observed in an uncontrolled forward Cauchy problem, this instability is created in an optimally controlled problem in the infinite horizon. This has important consequences and repercussions both from the conceptual as well as from the practical point of view. On the conceptual level, a controlled system is related to a system that somehow its final state (at $t \to \infty$ in our case) is predescribed. Therefore, our result is an "extension" of Turing instability in a forward-backward system and not just to a forward Cauchy problem, as is the case for the Turing instability. On the practical point of view, the optimal control nature of the problem we study here induces serious constraints on the growth rate of the allowed patterns which has a strict upper bound is related only to the discount factor of the model and not on the operator A. This is not the case for the standard Turing pattern formation mechanism, in which the growth rate upper bound is simply related to the spectrum of the operator A.

4.3 Hot spot of type III: The cost of robustness

The value function is of the form $V_k = \frac{M_{2,k}}{2}\hat{x}_k^2 + \frac{\hat{c}_k^2 M_{2,k}}{2r}$. This gives us the minimum possible deviation from the desired goal and it is made up from contributions by three terms:

- ▷ the term proportional to p in the cost functional which corresponds to the cost related to the deviation from the desired target,

- ▷ the term proportional to q in the cost functional which corresponds to the cost related to the cost of the control u needed to drive the system to the desired target and
- ▷ the term proportional to θ in the cost functional which corresponds to the cost of robustness (cost in the misspecification of the model).

The value functions depends on all these three contributions and this may be clearly seen since $M_{2,k}$ is in fact a function of the parameters p, q, θ.

An interesting question is which is the relevant importance of each of these contributions in the overall value function. Does one term dominates over the others or not?

A simple answer to this question will be given by the elasticity of the value function with respect to these parameters, i.e., by the calculation of the quantities $\frac{1}{V}\frac{\partial V}{\partial p}$, $\frac{1}{V}\frac{\partial V}{\partial q}$ and $\frac{1}{V}\frac{\partial V}{\partial \theta}$. It is easily seen that these elasticities are independent of $\hat{x}_k$ and reduce to $\frac{1}{M_{2,k}}\frac{\partial M_{2,k}}{\partial p}$, $\frac{1}{M_{2,k}}\frac{\partial M_{2,k}}{\partial q}$ and $\frac{1}{M_{2,k}}\frac{\partial M_{2,k}}{\partial \theta}$, respectively. Whenever one of these quantities tends to infinity, that means that the contribution of the relevant procedure dominates the control problem[7]

In particular whenever $\frac{1}{M_{2,k}}\frac{\partial M_{2,k}}{\partial \theta} \to \infty$, then we say that the cost of robustness becomes more expensive that what it offers, and we will call that a hot spot of type III. This quantity can be calculated directly from the solution of the quadratic equation (13) through straightforward but tedious algebraic manipulations, which we choose not to reproduce here.

However, an illustrative partial case, which allows some insight on the nature of hot spots of type III is the following:

Differentiating (13) with respect to θ yields

$$-\frac{\hat{c}_k^2}{2\theta^2}M_{2,k}^2 + 2\left(\frac{\hat{c}_k^2}{2\theta} - \frac{\hat{b}_k^2}{2q}\right) M_{2,k}\frac{\partial M_{2,k}}{\partial \theta} + (2\,\hat{a}_k - r)\frac{\partial M_{2,k}}{\partial \theta} = 0.$$

Dividing by $M_{2,k}^2$ we obtain

$$-\frac{\hat{c}_k^2}{2\theta^2} + 2\left(\frac{\hat{c}_k^2}{2\theta} - \frac{\hat{b}_k^2}{2q}\right)\frac{1}{M_{2,k}}\frac{\partial M_{2,k}}{\partial \theta} + (2\,\hat{a}_k - r)\frac{1}{M_{2,k}^2}\frac{\partial M_{2,k}}{\partial \theta} = 0.$$

Let us now take the particular case where $2\hat{a}_k = r$, so that

$$\frac{1}{M_{2,k}}\frac{\partial M_{2,k}}{\partial \theta} = \frac{\hat{c}_k^2}{4\theta\left(\frac{\hat{c}_k^2}{2\theta} - \frac{\hat{b}_k^2}{2q}\right)}$$

which becomes infinite for values of θ such that $\theta \to \frac{q\hat{c}_k^2}{\hat{b}_k^2}$. The general case $2\hat{a}_k \neq r$ may present similar phenomena.

4.4 Non translation invariant systems

The methodology employed in this section to provide closed form solutions used the translation invariant property of the dynamical system, which allowed the use of the discrete Fourier transform. This is a symmetry property of the system (commutation of the vector field with the translation operator) which has as a result that the spatial operators are convolutions and therefore the discrete Fourier

[7] This interpretation arises from observation that close to a point (p_0, q_0, θ_0) the value function behaves as

$$V_k \simeq \left.\frac{\partial V_k}{\partial p}\right|_{p=p_0}(p-p_0) + \left.\frac{\partial V_k}{\partial q}\right|_{q=q_0}(q-q_0) + \left.\frac{\partial V_k}{\partial \theta}\right|_{\theta=\theta_0}(\theta-\theta_0).$$

transform may be used to turn this convolution into a product in Fourier space. This situation may be generalized for other symmetry groups and may lead to interesting generalizations for systems which are not translation invariant but invariant under other more complicated symmetries. In this case the tools of harmonic analysis on groups (see e.g. Rudin (1990)) may be used and generalized Fourier transforms may be defined in terms of the Haar measure[8]. In terms of this generalized Fourier transform, the system will decouple thus allowing for use of the proposed method in more general settings (see e.g. Bamieh et al. (2002) for a related discussion).

5 The general linear quadratic control problem

We now relax the simplifying (and restrictive) assumptions concerning the translation invariance property of the operators $\mathsf{A}, \mathsf{B}, \mathsf{C}$ as well as the overly restrictive assumption that $\mathsf{P} = pI$ and $\mathsf{Q} = qI$.

We now consider instead the solution of the general linear quadratic robust control problem (5) under the state constraint (2), and comment on the possibility of hot spot formation working in real space directly rather than in Fourier transform space. The general form of the problem allows the study of a wider range of economic applications (see, e.g., Section **??** for an illustration of the applicability of the general problem). The relaxation of translation invariance leads to significant complications, and to the inability to derive solutions in closed form. For a simple illustration of the complexity of the problem see Appendix, Section 8.7.

5.1 Solution in terms of operator Riccati equation

The problem may be treated using an infinite dimensional Hamilton-Jacobi-Bellman-Isaacs equation, which is solvable in terms of an operator Riccati equation.

Theorem 1. *If the problem (5) under the dynamic constraint (2) has a solution, for arbitrary $x \in \mathbb{H}$, then the optimal controls are of the feedback control form*

$$u = -\mathsf{Q}^{-1}\mathsf{B}^*\mathsf{H}^{sym}x, \quad v = \frac{1}{\theta}\mathsf{R}^{-1}\mathsf{C}^*\mathsf{H}^{sym}x, \tag{22}$$

and the optimal state satisfies the Ornstein-Uhlenbeck equation

$$dx = (\mathsf{A} - \mathsf{B}\mathsf{Q}^{-1}\mathsf{B}^*\mathsf{H}^{sym} + \frac{1}{\theta}\mathsf{C}\mathsf{R}^{-1}\mathsf{C}^*\mathsf{H}^{sym})x\,dt + \mathsf{C}dW \tag{23}$$

where H^{sym} is the solution of the operator Riccati equation

$$\mathsf{H}^{sym}\mathsf{A} + \mathsf{A}^*\mathsf{H}^{sym} - \mathsf{H}^{sym}\mathsf{E}^{sym}\mathsf{H}^{sym} - r\mathsf{H}^{sym} + \mathsf{P} = 0 \tag{24}$$

and $\mathsf{E}^{sym} := \frac{1}{2}(\mathsf{E} + \mathsf{E}^*)$ is the symmetric part of $\mathsf{E} := \mathsf{B}\mathsf{Q}^{-1}\mathsf{B}^* - \frac{1}{\theta}\mathsf{C}\mathsf{R}^{-1}\mathsf{C}^*$.

Proof: We first obtain the relevant Hamilton-Jacobi-Bellman-Isaacs equation for this infinite dimensional stochastic differential game. This will be an elliptic equation on an infinite dimensional Hilbert space (and in particular ℓ^2). The generator for the Ornstein-Uhlenbeck equation is of the form

$$\mathcal{L}V = \sum_n \left(\sum_m (a_{nm}x_m + b_{nm}u_m + c_{nm}v_m) \right) \frac{\partial}{\partial x_n} + \sum_n \sum_m \sum_k c_{nk}c_{mk} \frac{\partial^2}{\partial x_n \partial x_m} \tag{25}$$

or in compact form

$$\mathcal{L}V = \langle \mathsf{A}x + \mathsf{B}u + \mathsf{C}v, DV \rangle + Tr(\mathsf{C}\mathsf{C}^* D^2 V)$$

[8] The Haar measure is a generalization of the Lebesgue measure, which is invariant under the symmetry group

where DV is the sequence (in ℓ^2) with $(DV)_n = \frac{\partial}{\partial x_n} V$ and D^2V is the infinite matrix (operator) with elements $(D^2V)_{nm} = \frac{\partial^2}{\partial x_n \partial x_m} V$. These are the infinite dimensional generalizations (in the Hilbert space ℓ^2) of the gradient and the Hessian matrix respectively. With Tr we denote the trace of the operator involved whereas with the superscript tr we denote the transposition. This is an elliptic operator in the infinite dimensional Hilbert space ℓ^2 (see Da Prato and Zabczyk (2002) or Cerrai (2001) for a general introduction to such operators).

We now construct the Hamiltonian. We start with the function $H : \ell^2 \times \ell^2 \times \ell^2 \to \mathbb{R}$ defined as

$$H(V; x, u, v) = \mathcal{L}V + \langle \mathsf{P}x, x\rangle + \langle \mathsf{Q}u, u\rangle - \theta\langle \mathsf{R}v, v\rangle$$

We need to obtain the upper Hamiltonian and lower Hamiltonians defined respectively as

$$\bar{H} := \sup_u \inf_v H(V; x, u), \quad \underline{H} := \inf_v \sup_u H(V; x, u).$$

Let us present in some relative detail the construction of $\underline{H}$. The maximization over u is a quadratic optimization problem over the Hilbert space ℓ^2. The first order condition is easily calculated to be

$$\frac{1}{2}(\mathsf{Q} + \mathsf{Q}^*)u = -\mathsf{B}^* DV$$

where by the symmetry of Q and the positive definite property is is seen that

$$u = -\mathsf{Q}^{-1}\mathsf{B}^* DV$$

We work similarly with the minimization problem over v which is again a quadratic optimization problem in ℓ^2 whose first order conditions yield

$$\frac{\theta}{2}(\mathsf{R}^* + \mathsf{R})v = \mathsf{C}^* DV$$

which upon invoking the symmetry of the operator R and the positive definite property yields

$$v = \frac{1}{\theta}\mathsf{R}^{-1}\mathsf{C}^* DV$$

We now insert these expressions for u and v back into H to obtain

$$\begin{aligned}\underline{H} = &\left\langle \mathsf{A}x - \mathsf{B}\mathsf{Q}^{-1}\mathsf{B}^* DV + \frac{1}{\theta}\mathsf{C}\mathsf{R}^{-1}\mathsf{C}^* DV, DV \right\rangle + Tr(\mathsf{C}\mathsf{C}^* D^2V)\\ &+ \frac{1}{2}\langle \mathsf{P}x, x\rangle + \frac{1}{2}\langle \mathsf{Q}\mathsf{Q}^{-1}\mathsf{B}^* DV, \mathsf{Q}^{-1}\mathsf{B}^* DV\rangle - \frac{1}{2\theta}\langle \mathsf{R}\mathsf{R}^{-1}\mathsf{C}^* DV, \mathsf{R}^{-1}\mathsf{C}^* DV\rangle\end{aligned}$$

which upon rearrangement yields

$$\underline{H} = \langle \mathsf{A}x, DV\rangle - \frac{1}{2}\left\langle \left(\mathsf{B}\mathsf{Q}^{-1}\mathsf{B}^* - \frac{1}{\theta}\mathsf{C}\mathsf{R}^{-1}\mathsf{C}^*\right) DV, DV \right\rangle + \frac{1}{2}\langle \mathsf{P}x, x\rangle + Tr(\mathsf{C}\mathsf{C}^* D^2V)$$

The lower Hamilton-Jacobi-Bellman-Isaacs then becomes

$$\langle \mathsf{A}x, DV\rangle - \frac{1}{2}\left\langle \left(\mathsf{B}\mathsf{Q}^{-1}\mathsf{B}^* - \frac{1}{\theta}\mathsf{C}\mathsf{R}^{-1}\mathsf{C}^*\right) DV, DV \right\rangle + \frac{1}{2}\langle \mathsf{P}x, x\rangle + \frac{1}{2}Tr(\mathsf{C}\mathsf{C}^* D^2V) = rV$$

This is a nonlinear elliptic equation in ℓ^2, the solution of which will provide the optimal controls.

On account of the quadratic nature of the problem we seek a solution of the following ansatz

$$V(x) = \frac{1}{2}\langle \mathsf{H}x, x\rangle + \langle j, x\rangle + K$$

where H is an operator, j is an element of the Hilbert space ℓ^2 and $K \in \mathbb{R}$. For this choice[9]

$$DV = \frac{1}{2}(\mathsf{H} + \mathsf{H}^*)x + j, \quad D^2 V = \frac{1}{2}(\mathsf{H} + \mathsf{H}^*).$$

We now substitute into the HJBI equation and match powers of x.

(a) The quadratic terms yield the operator Riccati equation

$$\frac{1}{2}(\mathsf{H}^* + \mathsf{H})\mathsf{A} - \frac{1}{8}(\mathsf{H}^* + \mathsf{H})\mathsf{E}(\mathsf{H} + \mathsf{H}^*) + \frac{1}{2}(\mathsf{P} - r\mathsf{H}) = 0 \tag{26}$$

where

$$\mathsf{E} := \mathsf{B}\mathsf{Q}^{-1}\mathsf{B}^* - \frac{1}{\theta}\mathsf{C}\mathsf{R}^{-1}\mathsf{C}^*$$

the solution of which yields H.

(b) The terms which are linear in x provide a linear homogeneous equation for j which admits the trivial solution $j = 0$.

(c) Finally, the constant term K is

$$K = \frac{1}{2r} Tr \left(\mathsf{C}\mathsf{C}^* \frac{\mathsf{H} + \mathsf{H}^*}{2} \right).$$

A quick inspection of the solution shows that we only need the symmetric part of the operator H, $\mathsf{H}^{sym} := \frac{1}{2}(\mathsf{H} + \mathsf{H}^*)$, in order to specify the optimal control and the value function[10].

We therefore need to obtain a Riccati equation involving H^{sym} only.

Taking the adjoint of equation (26) and adding we obtain the following symmetric Riccati equation for $\mathsf{H}^{sym} := \frac{1}{2}(\mathsf{H} + \mathsf{H}^*)$ in the form

$$\mathsf{H}^{sym}\mathsf{A} + \mathsf{A}^*\mathsf{H}^{sym} - \mathsf{H}^{sym}\mathsf{E}^{sym}\mathsf{H}^{sym} - r\mathsf{H}^{sym} + \mathsf{P} = 0$$

where $\mathsf{E}^{sym} := \frac{1}{2}(\mathsf{E} + \mathsf{E}^*)$ is the symmetric part of E. In the above calculation we have explicitly taken into account the symmetry of P.

The determination of the optimal controls requires only the solution of the symmetric Riccati operator equation (27). Once H^{sym} is known the feedback control laws are given by

$$u = -\mathsf{Q}^{-1}\mathsf{B}^*\mathsf{H}^{sym}x, \quad v = \frac{1}{\theta}\mathsf{R}^{-1}\mathsf{C}^*\mathsf{H}^{sym}x, \tag{27}$$

where now x is assumed to be the current state of the system. ■

Remark 18. The operator Riccati equation (27) is the generalization of the quadratic algebraic equation (13) in the case where the operators A, B and C are not translation invariant, and thus amenable to analysis using the Fourier transform. Furthermore, in the case where the state space is finite dimensional (i.e., in the case of finite lattices) the operator Riccati equation assumes the form of a matrix Riccati equation.

Clearly, by Theorem 1 the solvability and the properties of the solution for the optimal control problem is reduced to the solvability and the properties of the solution of the operator Riccati equation (27).

Proposition 11. *Let $m = ||\mathsf{A}||$ defined as $m = \{\sup\langle \mathsf{A}x, x\rangle, \ ||x||_{\ell^2} = 1\}$ and assume that $m < r/2$. Then, for small enough values of $||\mathsf{E}||$ and $||\mathsf{P}||$ the operator Riccati equation (27) admits a unique bounded strong solution.*

[9] If H is a symmetric operator, then the above expressions simplify to $DV = \mathsf{H}x + j$, $D^2V = \mathsf{H}$, however, we do not make this assumption at this point.

[10] It is easily seen that $\langle \mathsf{H}x, x\rangle = \langle \mathsf{H}^{sym}x, x\rangle$.

Proof: By further defining the operator $\tilde{\mathsf{A}} = \mathsf{A} - \frac{r}{2} I$ the symmetric operator Riccati equation simplifies to

$$\mathsf{H}^{sym}\tilde{\mathsf{A}} + \tilde{\mathsf{A}}^{tr}\mathsf{H}^{sym} - \mathsf{H}^{sym}\mathsf{E}^{sym}\mathsf{H}^{sym} + \mathsf{P} = 0 \tag{28}$$

This is in the standard form of operator Riccati equation studied in the literature (see e.g., Bensoussan et al. (1992) or Da Prato (2002)). The spectrum of the operator $\tilde{\mathsf{A}}$ is in the interval $[-m - \frac{r}{2}, m - \frac{r}{2}]$ whereas the spectrum of the operator $-\tilde{\mathsf{A}}$ is in the interval $[\frac{r}{2} - m, m + \frac{r}{2}]$. If $m < \frac{r}{2}$ then $d := dist(spec(\tilde{\mathsf{A}}), spec(-\tilde{\mathsf{A}})) > 0$. Then according to Theorem 3.7 in Albeverio et al. (2003) (whose proof is based on the Banach contraction theorem) equation (28) has a unique solution. ■

Remark 19. The "smallness" condition on $||\mathsf{E}||$ and $||\mathsf{P}||$ is made explicit via the Banach contraction argument in the proof of Theorem 3.7 in Albeverio et al. (2003). In particular, for the existence of a strong solution we need $||\mathsf{E}|| + ||\mathsf{P}|| < d$. It can be seen that this condition breaks down for small enough values of θ, which in fact is the analogue of the hot spot of Type I that was obtained before. for the restricted class of models involving translation invariant operators, using the Fourier expansion method (see Proposition 8).

As long as we can guarantee the solvability of the relevant HJBI equation through the solvability of the equivalent operator Riccati equation, what is finally left to do is to check that this solution and the feedback control law it leads to is indeed the optimal solution. This is a verification theorem and is needed in order to check the assumption we have included in Theorem 1 concerning the existence of solution to the robust optimal control problem. This is not simply a mathematical caprice, it is an important step in the analysis needed to verify that the candidate for a solution offered by the solution of the HJBI equation is indeed a solution of the control problem. This step is especially important even for the simple linear quadratic regulator problem in infinite dimensional spaces.

Furthermore, we will need to guarantee that the solution to the operator Riccati equation has positivity properties, as this will allow us to prove that the optimal path defined by the relevant feedback law has minimizing properties for the cost functional. These conditions are related to the properties of the operators A, B, C and in particular properties related to stabilizability or observability.

5.2 Hot spot formation in general linear quadratic systems

The various hot spots that were obtained explicitly for the translation invariance case, can be generalized for the general linear quadratic case.

Concerning hot spots of Type I, these are related to the breakdown of the minimax problem involved, for small values of the parameter θ. In fact, our results concerning the solvability of the relevant Riccati equation are pointing in this direction. Therefore, hot spots of Type I do exist in general linear quadratic systems, and are indeed related to model mispecification costs.

Similarly, hot spots of Type II will also exist. It can be seen that if the operators involved in the Riccati equation are diagonalizable then the Riccati equation admits a solution H^{sym} is a diagonalizable operator as well. Then, by the properties of diagonalizable operators the operators involved in the feedback laws are diagonalizable as well. A quick inspection shows that in this case the spectral theorem holds for the operator $\mathcal{R}$, therefore, the pattern formation behavior for the optimal path may now be obtained by spectral analysis. The hot spots of Type II will correspond to these eigenfunctions of the operator $\mathcal{R} := \mathsf{A} - \mathsf{B}\mathsf{Q}^{-1}\mathsf{B}^*\mathsf{H}^{sym} + \frac{1}{\theta}\mathsf{C}\mathsf{R}^{-1}\mathsf{C}^*\mathsf{H}^{sym}$ that have positive eigenvalues. A priori estimates of the spectrum may help us rule out the possibility of the emergence of hot spots.

The above general formulae for the feedback controls simplify in certain particular cases of interest. For instance, assume that we are interested in the localized entropic constraints problem introduced in Proposition 3. Then, since R is a diagonal operator, it can easily be seen that R^{-1} is also diagonal and has the representation $r_{nm}^{-1} = \theta_n^{-1}\delta_{nm}$. Therefore, the optimal feedback control v can be expressed as

$$v_n = \frac{1}{\theta_n}(\mathsf{C}^*\mathsf{H}^{sym}x)_n,$$

where,

$$(\mathsf{C}^*\mathsf{H}^{sym}x)_n = \sum_m \mathsf{f}_{nm}x_m$$

where f_{nm} are determined as long as the operators C^*, H^{sym} are known. This form clarifies the effect of model misspecification in particular sites; the behavior of the system at lattice site n depends inversely proportionally ($\frac{1}{\theta_n}$ contribution) on the Lagrange multiplier of the localized entropic constraint at this site, as well as on the state of the system at neighboring sites through the term $(\mathsf{C}^*\mathsf{H}^{sym}x)_n$. This semi-explicit form allows us to understand the effect of model misspecification in certain sites. For instance sites with very large values of θ_n that in the absence of robust control tend to be unstable will remain so. Sites with small values of θ_n, but larger than a critical cutoff value, that in the absence of model misspecification tend to be unstable may be stabilized as an effect of robustness. On the other hand extremely small values of θ_n may destabilize the system. This qualitative picture is in some sense a generalization of the arguments concerning hot spot formation, from the limited case of translation invariant operators to the general linear quadratic case.

The above remarks can help us understand the emergence of hot spots in the general linear model. Assume for simplicity that C is diagonal and that the spatial domain is finite so that $\theta = (\theta_0, ..., \theta_{N-1})$ is the vector of local misspecification concerns. The low θ's will correspond to locations with the higher concerns. If one or more of these low θ's are such that the "smallness" condition on $||\mathsf{E}||$ and $||\mathsf{P}||$ is violated then local concerns will cause global regulation to break down.

In the same way if the low θ's are such that the operator $\mathcal{R}$ has positive eigenvalues then local concerns may induce global spatial clustering through the mechanism described in section 4.2.

Pattern formation in the general linear quadratic system can also emerge through a 'non-Turing' mechanism. We can write the mean field for the optimal state as

$$dx = \left[\mathsf{A} - \mathsf{B}\mathsf{Q}^{-1}\mathsf{B}^*\mathsf{H}^{sym} + \frac{1}{\theta}\mathsf{C}\mathsf{R}^{-1}\mathsf{C}^*\mathsf{H}^{sym}\right] x = \mathcal{R}x \tag{29}$$

Assume that matrix A is invertible but matrix $\mathcal{R}$ which embodies optimization and misspecification concerns is not invertible. In this case the steady state equation $0 = \mathcal{R}x$ will have more than one solutions. This means that there will be vectors $x \neq \mathbf{0}$ that will satisfy $0 = \mathcal{R}x$. These vectors will be $\ker(\mathcal{R})$. If $\ker(\mathcal{R})$ consists of vectors which spatially nonuniform then pattern formation emerges. This is pattern formation mechanism is however a non-Turing mechanism. Of course to examine whether such a mechanism exists a detailed analysis of matrix $\mathcal{R}$ and its null space is required. For example given the parametrization of the system one could ask the question whether a vector of misspecification concerns θ exists, such that $\mathcal{R}$ is not invertible. If such a vector exists then the specific misspecification concerns could induce pattern formation in the general linear quadratic model.

6 Nonlinear systems

6.1 General form of the controlled system

Consider now the nonlinear system

$$dx = (\mathsf{A}x + \mathsf{F}(x) + \mathsf{B}u)dt + \mathsf{C}dw$$

where $\mathsf{A} : \ell^2 \to \ell^2$ is a linear operator and $\mathsf{F} : \ell^2 \to \ell^2$ is in general a nonlinear operator and C is the covariance operator. The simplest choice for the nonlinear term F may be $\mathsf{F}(x) = (f_1(x_1), f_2(x_2), \cdots)$ in which case the nonlinear effects are purely local, however this is be no means a necessary restriction[11]. The functions $\{f_i\}$ will be assumed to be twice differentiable, together with their derivatives and satisfying dissipativity conditions. The control acts on the system through the linear operator $\mathsf{B} : \ell^2 \to \ell^2$. Unless otherwise explicitly stated the linear operators A, B, C are assumed to satisfy the conditions stated in Proposition 1.

The robust form of the system, using the Girsanov theorem is

$$dx = (\mathsf{A}x + \mathsf{F}(x) + \mathsf{B}u + \mathsf{C}v)dt + \mathsf{C}dw. \tag{30}$$

Under the stated assumptions for the nonlinearity an analogue of Proposition 1 can be shown to hold using standard techniques (see e.g. Prato and Zabczyk (1995), Da Prato and Zabczyk (1996), Da Prato and Zabczyk (2002), Cerrai (2001)).

We now consider a control functional of the form

$$J = \mathbb{E}_Q \left[\int_0^\infty e^{-rt} (\mathsf{U}(x(t)) + \mathsf{K}(u(t)) - \mathsf{T}(v(t)))dt \right]$$

where $\mathsf{U} : \ell^2 \to \mathbb{R}$ is a measure of distance from a desired target (a disutility function), $\mathsf{K} : \ell^2 \to \mathbb{R}$ is a cost function for the control and $\mathsf{T} : \ell^2 \to \mathbb{R}$ is a cost function for the robustness. All three functions are assumed convex. The robust control problem thus becomes

$$\min_u \max_v \mathbb{E}_Q \left[\int_0^\infty e^{-rt} (\mathsf{U}(x(t)) + \mathsf{K}(u(t)) - \mathsf{T}(v(t)))dt \right]$$

subject to the nonlinear state equation (30). By $\mathsf{K}^{\mathcal{L}} : \mathbb{H}^* \to \mathbb{R}$ we denote the Fenchel-Legendre transform of K defined by

$$\mathsf{K}^{\mathcal{L}}(p) := \sup_{x \in \mathbb{H}} [\langle p, x \rangle - \mathsf{K}(x)],$$

where by the Riesz representation we assume that the dual space $\mathbb{H}^* \simeq \mathbb{H}$ (in this particular case $\mathbb{H} = \ell^2$).

6.2 Solution in terms of the HJBI equation

The nonlinear optimal control problem may be treated in terms of a fully nonlinear Hamilton-Jacobi-Bellman-Isaacs equation.

Theorem 2. *The Hamilton-Jacobi-Bellman-Isaacs equation associated with the robust control problem (**??**) subject to the constraint (30) is the infinite dimensional nonlinear PDE*

$$\langle \mathsf{A}x + \mathsf{F}(x), DV \rangle + Tr(\mathsf{C}\mathsf{C}^* D^2 V) + \mathsf{U}(x) - \mathsf{K}^{\mathcal{L}}(-\mathsf{B}^* DV) + \mathsf{T}^{\mathcal{L}}(\mathsf{C}^* DV) = rV \tag{31}$$

where $\mathsf{K}^{\mathcal{L}}$, $\mathsf{T}^{\mathcal{L}}$ are the Fenchel-Legendre transforms of K and $\mathsf{T}^{\mathcal{L}}$ respectively. Given a solution of this equation $V : \mathbb{H} \to \mathbb{R}$ of sufficient regularity the associated closed loop system is the nonlinear infinite dimensional Ornstein-Uhlenbeck system

$$dx = (\mathsf{A}x + \mathsf{F}(x) - D\mathsf{K}^{\mathcal{L}}(\mathsf{B}^* DV(x)) + D\mathsf{T}^{\mathcal{L}}(\mathsf{C}^* DV(x)))dt + \mathsf{C}dw \tag{32}$$

Proof: The generator operator for the infinite dimensional diffusion process defined by the solution of (30) is the linear operator $\mathcal{L}$ whose action on a suitably smooth function Φ is given by

$$\mathcal{L}\Phi = \sum_n \left(\sum_m (a_{nm}x_m + b_{nm}u_m + c_{nm}v_m) + f_n(x_n) \right) \frac{\partial}{\partial x_n}\Phi + \sum_n \sum_m \sum_k c_{nk}c_{mk} \frac{\partial^2}{\partial x_n \partial x_m}\Phi \tag{33}$$

[11] The general form $\mathsf{F}(x) = (f_1(x), f_2(x), \cdots)$ where at each site the nonlinear effects depends on the state of the system at all lattice sites is easily handled within the framework presented here.

or in compact form

$$\mathcal{L}\Phi = \langle \mathsf{A}x + \mathsf{F}(x) + \mathsf{B}u + \mathsf{C}v, D\Phi\rangle + Tr(\mathsf{C}\mathsf{C}^* D^2\Phi)$$

We now consider the Hamiltonian

$$H(V; x, u, v) = \mathcal{L}V + \mathsf{U}(x) + \mathsf{K}(u) + \mathsf{T}(v),$$

which may be rewritten as a sum of three terms

$$\begin{aligned}
H(V; x, u, v) &= H_1(V; x) + H_2(V; u) + H_3(V; v),\\
H_1(V; x) &:= \langle \mathsf{A}x + \mathsf{F}(x), DV\rangle + Tr(\mathsf{C}\mathsf{C}^* D^2 V) + \mathsf{U}(x),\\
H_2(V; u) &:= \langle \mathsf{B}u, DV\rangle + \mathsf{K}(u),\\
H_3(V; v) &:= \langle \mathsf{C}v, DV\rangle + \mathsf{T}(v),
\end{aligned}$$

Note that

$$\begin{aligned}
\inf_u H_2(V; u) &= \inf_u\left(\langle \mathsf{B}u, DV\rangle + \mathsf{K}(u)\right) = \inf_u\left(\langle u, \mathsf{B}^* DV\rangle + \mathsf{K}(u)\right)\\
&= -\sup_u\left(\langle u, (-\mathsf{B}^* DV)\rangle - \mathsf{K}(u)\right) = -\mathsf{K}^{\mathcal{L}}(-B^{tr}DV),
\end{aligned}$$

where $\mathsf{K}^{\mathcal{L}}$ denotes the Fenchel-Legendre transform of K. By the theory of the Fenchel-Legendre transform (see e.g., Aubin and Ekeland (1984)) $p \in \partial\mathsf{K}(u)$ is equivalent to $u \in \partial\mathsf{K}^{\mathcal{L}}(p)$ where ∂ denotes the subdifferential operator, therefore, by the regularity assumptions imposed on K, the minimizer is

$$u = -D\mathsf{K}^{\mathcal{L}}(\mathsf{B}^* DV) \tag{34}$$

Similarly,

$$\sup_v H_2(V; v) = \sup_v\left(\langle \mathsf{C}u, DV\rangle - \mathsf{T}(u)\right) = \mathsf{T}^{\mathcal{L}}(\mathsf{C}^* DV) \tag{35}$$

where $\mathsf{T}^{\mathcal{L}}$ is the Fenchel-Legendre transform of T. By similar arguments as above, the maximizer is

$$v = D\mathsf{T}^{\mathcal{L}}(\mathsf{C}^* DV) \tag{36}$$

The Hamilton-Jacobi-Bellman-Isaacs equation assumes the form

$$\langle \mathsf{A}x + \mathsf{F}(x), DV\rangle + Tr(\mathsf{C}\mathsf{C}^* D^2 V) + \mathsf{U}(x) - \mathsf{K}^{\mathcal{L}}(-\mathsf{B}^* DV) + \mathsf{T}^{\mathcal{L}}(\mathsf{C}^* DV) = rV$$

which is an infinite dimensional second order fully nonlinear partial differential equation for the value function V. The nonlinearity enters through the Fenchel-Legendre transforms $\mathsf{K}^{\mathcal{L}}$ and $\mathsf{T}^{\mathcal{L}}$ of K and T respectively. Assume the existence of a regular enough solution (see Theorem 3) and substitute (34) and (36) into (30) to obtain the closed loop control system (32). ■

The solvability of the infinite dimensional HBJI equation is provided in the next theorem.

Theorem 3. *Assume that either A is the generator of an analytic semigroup or that A is the generator of a C_0 semigroup such that $||\mathsf{Q}^{-1/2}\exp(t\mathsf{A})|| \le C\,t^{-\delta}$ for some $\delta \in (0,1)$, $t \ge 0$. Assume furthermore that F is a locally Lipschitz nonlinear operator. Then, there exists a critical discount factor r_{cr} such that for $r > rcr$ the HBJI equation has a unique solution $V \in D(\mathcal{L})$.*

Proof: The proof follows by generalizing results either of Da Prato and Zabczyk (2002) or Cerrai (2001), for stochastic control problems, to the case of stochastic differential games. We do not provide the detailed proof here for the sake of brevity. The proof uses techniques from the theory of maximal dissipative operators. ■

Remark 20. If we replace the local Lipschitz condition on the nonlinearity with a global Lipschitz condition then we may put $r_{cr} = 0$. However, the global Lipschitz property is a rather strong assumption. A particular case where is holds is, e.g., when F is generated by a quadratic nonlinearity for bounded u which vanishes for $u > R$ for some R.

Remark 21. The conditions on the operator A are not too strict. The discretized Laplacian A_h for example satisfies these conditions on finite dimensional approximations of the Hilbert space L^2. Another case where this is true is when C has bounded inverse and A has spectrum $\{a_n\}$ satisfying certain conditions such as for example $\sum^n \frac{a_n^2}{(e^{2a_n t}-1)^2} < \infty$ for all $t \geq 0$ (see, e.g., Da Prato and Zabczyk (2002)). Such conditions hold for second order differential operators of elliptic type such as the Laplacian in bounded domains, but it may also hold for the (unbounded) discrete Laplacian on infinite graphs.

6.3 Hot spot formation in nonlinear systems

We now consider the possibility of hot spot formation in the nonlinear robust control system. Let us assume a steady state $x_0 \in \mathbb{H}$ for the averaged over all realizations of the noise closed loop system, which presents no spatial patterns. This means that x_0 is such that $\mathsf{A}x_0 = 0$ and furthermore

$$0 = \mathsf{F}(x_0) - D\mathsf{K}^{\pounds}(\mathsf{B}^* DV(x_0)) + D\mathsf{T}^{\pounds}(\mathsf{C}^* DV(x_0))$$

for a sufficiently smooth solution of the HJBI equation. We now consider a small perturbation of x_0 in the form $x = x_0 + \epsilon z$ where ϵ is a small real number and $z \in \mathbb{H}$, and see how the closed loop system (32) evolves under this perturbation. A relevant question is whether z develops any spatial variability which is interpreted as a hot spot of type II.

The following proposition provides some answer to this question.

Proposition 12. *Assume that V is a C^2 solution of (31). If $\mathsf{K}^{\pounds}$ and $\mathsf{T}^{\pounds}$ are C^2 then the perturbation z is the solution of the linear Ornstein-Uhlenbeck equation*

$$dz = (\mathsf{A}z + D\mathsf{F}(x_0)z - D^2\mathsf{K}^{\pounds}\,\mathsf{B}^* D^2V(x_0)z + D^2\mathsf{T}^{\pounds}\,\mathsf{C}^* D^2V(x_0)z)dt + \mathsf{C}dw$$

*The type II hot spots correspond to the unstable modes of the equation (**??**), i.e., to eigenfunctions of the operator*

$$\mathcal{R} := \mathsf{A} + D\mathsf{F}(x_0) - D^2\mathsf{K}^{\pounds}\,\mathsf{B}^* D^2V(x_0) + D^2\mathsf{T}^{\pounds}\,\mathsf{C}^* D^2V(x_0)$$

with positive eigenvalues.

Proof: Substitute $x = x_0 + \epsilon z$ into (32) and expand in powers of ϵ, using the Taylor expansion theorem for Hilbert space valued functions. According to that

$$DV(x_0 + \epsilon z) = DV(x_0) + \epsilon D^2V\, z + O(\epsilon^2)$$

where $D^2V : \mathbb{H} \to \mathbb{H}$ is a symmetric operator corresponding to the generalization of the Hessian matrix. Therefore, by the assumed regularity of $\mathsf{K}^{\pounds}$

$$D\mathsf{K}^{\pounds}(\mathsf{B}^* DV(x)) = D\mathsf{K}^{\pounds}(\mathsf{B}^* DV(x_0)) + \epsilon D^2\mathsf{K}^{\pounds}\,\mathsf{B}^* \epsilon D^2V\, z + O(\epsilon^2)$$

and similarly for the other nonlinear terms. Inserting into the closed loop system yields the required result for the evolution of z. The rest follows by spectral theory considerations. ■

The value functions and the Legendre-Fenchel transforms satisfy convexity properties. This gives important information on the second derivatives $D^2\mathsf{K}^{\pounds}$, $D^2V(x_0)$, $D^2\mathsf{T}^{\pounds}$ and in particular assuming sufficient regularity they are positive operators. This property allows us at least to obtain some a priori estimates on the spectrum of the operator $\mathcal{R}$ and thus provide values on the parameters of the model which allow the generation of hot spot. In this respect, we may generalize some of the findings of the linear model in the nonlinear model.

7 Concluding remarks

We study robust control methods in a spatial domain where explicit spatial interactions are modelled by kernels and where concerns about model misspecification could be different across locations. We analyze linear quadratic problem. We derive closed form solutions for translation invariant systems in finite and infinite lattices but we also extent our results to general non translation invariant linear quadratic problems as well as to fully non linear systems. We show that misspecification concerns about specific cites could induce the emergence of hot spots which cause regulation to break down for the whole spatial domain. We also identify conditions for two more types of hot spots where location specific concerns could induce the emergence of spatial patters, or could render regulation very costly. We apply our methods to a problem of regulating in situ consumption when consumers are characterized by distance-dependent utility. We examine the emergence of local markets for in situ consumption and cases where location specific concerns could brake down regulation for the whole area, or could induce specific clustering.

Our results provide tools for studying optimal regulation of spatially interconnected systems when there are concerns about the specification of the model describing local processes describing the evolution of the system's states. Given the increasing interconnections and the localized uncertainties in the real world our approach could be appropriate for a wide class of economic problems characterized and connectivity, not necessarily spatial, since connectivity can be regarded with respect to other attributes, and by local uncertainties.

8 Appendix

8.1 The abstract formulation of the state equation

This section makes a few comments on the abstract formulation of the state equation. Since we are dealing with a (possibly) infinite dimensional dynamical system, this step is important in dealing with the question of well posedeness of the system, a property which is essential before using the model to derive any qualitative information concerning the economics.

As mentioned above the state space of the system is a version of the sequence space ℓ^2, the space of square summable sequences $\{x_n\}$. However, for certain states of the system which may be of economic interest this space may not be sufficient. As an example, consider an interconnected system on the infinite lattice $\mathbb{Z}$, and consider the case where all economies of the lattice are on the same state x, i.e., the spatially homogeneous state $x^{(0)} := \{x_n^{(0)}\}$, $x_n = x$ for all $n \in \mathbb{Z}$. Clearly, $x^{(0)} \neq \ell^2(\mathbb{Z})$, therefore, modelling our system in this state space we are unable to capture phenomena related to the spatially homogeneous state or its perturbations. This is rather annoying specially if we wish to consider agglomeration phenomena arising from weak pertrubation of the homogeneous state. Same thing happens for other states of importance to the economics of the system such as for instance patterns such that $x_n \to c$ as $n \to \infty$ (i.e. a pattern in space that at infinity our system reaches a steady state not equal to 0) or travelling wave type of patterns. What is to be done? We have to change the state space so as to accomodate states of economic importance on the infinite lattice such as the above mentioned examples. The right choice would be a weighted version of ℓ^2, the space ℓ^2_ρ, where $\rho : \mathbb{Z} \to \mathbb{R}$ is an appropriate weight function. The space ℓ^2_ρ is defined as

$$\ell^2_\rho := \{\, \{x_n\} \,:\, \sum_{n\in\mathbb{Z}} \rho_n x_n^2 < \infty\},$$

and this is a Hilbert space when equipped with the inner product

$$\langle x, y\rangle := \sum_{n\in\mathbb{Z}} \rho_n x_n y_n.$$

The appropriate choice of the weight function ρ allows us to accomodate "interesting" patterns in this state space, for example if the weight function is chosen so that $\sum_{n\in\mathbb{Z}} \rho(n) < \infty$ (i.e. $\rho := \{\rho(n)\} \in \ell^1$ then the homogeneous state $x^{(0)} \in \ell^2_\rho$ whereas $x^{(0)} \notin \ell^2$. Clearly for the appropriate choice of ρ, the following embedding holds $\ell^2 \subset \ell^2_\rho$. Weighted spaces can be very useful in the study of infinite dimensional lattice dynamical systems and their long time dynamics (see e.g. Karachalios and Yannacopoulos (2005) or Karachalios and Yannacopoulos (2007)).

We will now characterize the operators A, B, C as mappings of $\ell^2_\rho \to \ell^2_\rho$ for appropriate choices of the weight sequence ρ. We will work in terms of the generic operator $\mathsf{T} : \ell^2_\rho \to \ell^2_\rho$ whose action is defined as $(Tx)_n := \sum_m t_{nm}x_m$. In what follows T and t_{nm} will be a proxy for $\mathsf{A}, \mathsf{B}, \mathsf{C}$ and a_{nm}, b_{nm}, c_{nm} respectively.

Assumption 2. There exists exists a weight sequence[12] $\rho = \{\rho_n\}$ such that

(i) $\sup_n \sum_m |t_{nm}| < \infty$

(ii) $\sum_n \rho_n |t_{nm}| \leq C\,\rho_m$ for all $m \in \mathbb{Z}$.

Assumption 3. There exists a weight sequence[13] $\rho = \{\rho_n\}$ such that the double sequence $\{t_{nm}\}$ it holds that $\sum_n \sum_m \frac{\rho_n}{\rho_m} |t_{nm}|^2 < \infty$

Lemma 2.

(i) Let Assumption 2 hold for some weight sequence ρ. Then $\mathsf{T} : \ell^2_\rho \to \ell^2_\rho$ is a bounded operator.

(ii) Let Assumption 2 hold for the weight sequence $\rho = 1$ and consider a weight sequence $\phi \in \ell^{\frac{q-1}{q-2}}$ for $q > 2$. Then $\mathsf{T} : \ell^2 \to \ell^2$ is a bounded operator while $\mathsf{T} : \ell^2 \to \ell^2_\phi$ is a compact operator.

(iii) Let Assumption 3 hold, for the choice $\{\rho_n\} = \{1\}$. Then $\mathsf{T} : \ell^2 \to \ell^2$ is a compact operator.

(iv) Let Assumption 3 hold. Then $\mathsf{T} : \ell^2_\rho \to \ell^2_\rho$ is a compact operator.

(iv) Let Assumption 2 hold. Then the operators $\mathsf{T} : \ell^2 \to \ell^2_\rho$ generators a uniformly continuous semigroup of linear operators $\{S(t)\}_{t\in\mathbb{R}_+} = \{e^{tT}\}_{t\in\mathbb{R}_+}$ on ℓ^2_ρ.

Proof: (i) This is essentially a special case in Hilbert space of Proposition 12.2.1, p. 228 in Da Prato and Zabczyk (1996). The result follows by appropriate application of the Cauchy-Schwartz inequality according to which

$$\left(\sum_m a_{nm}x_m\right)^2 \leq \left(\sum_m |a_{nm}||x_m|^2\right)\left(\sum_m |a_{nm}|\right).$$

We then multiply each of the above inequalities by ρ_n and add over all n to obtain

$$\sum_n \rho_n \left(\sum_m a_{nm}x_m\right)^2 \leq C\left(\sum_m |a_{nm}|\right)\sum_n \rho_n |x_n|^2$$

from which the claim follows.

(ii) Since $\phi \in \ell^{\frac{q-1}{q-2}}$, $q > 2$, it can be seen that $\ell^2 \subset \ell^2_\phi$. Therefore, any element of ℓ^2 can be considered as an element of ℓ^2_ϕ and we may thus consider a mapping $\iota : \ell^2 \to \ell^2_\phi$, called an embedding such that for any element of $x \in \ell^2$, $y := \iota x \in \ell^2_\phi$ and y is nothing else but x only that now this is

[12] The choice $\rho = \{1\}$ is allowed.
[13] The choice $\rho = \{1\}$ is allowed.

considered as an element of the larger space ℓ^2_ϕ. This embedding map $\iota : \ell^2 \to \ell^2_\phi$ is compact (see Lemma 2.1 in Karachalios (2006)). This is an important observation since it implies that a bounded sequence in ℓ^2 has a convergent subsequence in the larger space ℓ^2_ϕ.

By (i) $\mathsf{T} : \ell^2 \to \ell^2$ is a bounded operator. That means, for every bounded sequence $\{x^{(k)}\} \subset \ell^2$, $\{\mathsf{T}x^{(k)}\}$ is a bounded sequence in ℓ^2. But from the compact embedding lemma (Lemma **??**) this bounded sequence in ℓ^2 has a convergent subsequence in ℓ^2_ϕ. Therefore $\mathsf{T} : \ell^2 \to \ell^2_\phi$ is a compact operator.

(iii) Consider a bounded sequence $\{x^k\} \subset \ell^2$. We need to show that under Assumption **??** the sequence $\{\mathsf{T}x^{(k)}\}$ is Cauchy in ℓ^2. A straightforward application of the Cauchy-Schwarz inequality for sums yields

$$||\mathsf{T}x^{(k)} - \mathsf{T}x^{(r)}||^2_{\ell^2} = \sum_n \left(\sum_m t_{nm}(x^{(k)}_m - x^{(r)}_m) \right)^2 \leq \left(\sum_n \sum_m t^2_{nm} \right) \left(\sum_m (x^{(k)}_m - x^{(r)}_m)^2 \right)$$

which upon definition of $\hat{t} := \sum_n \sum_m t^2_{nm}$ becomes

$$||\mathsf{T}x^{(k)} - \mathsf{T}x^{(r)}||^2_{\ell^2} \leq \hat{t}\, ||x^{(k)} - x^{(r)}||_{\ell^2}. \tag{37}$$

Since $x^{(k)}$ is bounded in ℓ^2 then $x^{(k)} - x^{(r)}$ is also bounded therefore for every $N \in \mathbb{N}$ there exists $\epsilon > 0$ such that

$$\sum_{|n|>N} |x^{(k)}_n - x^{(r)}_n|^2 < \epsilon \tag{38}$$

(the series is convergent therefore the tail of the series can be made as small as possible). For course the same argument (boundedness of the ℓ^2 norm) implies that there are constants C_1, C_2 such that for every N,

$$\sum_{|n|\leq N} |x^{(k)}|^2 < C_1 \quad \sum_{|n|\leq N} |x^{(k)}_n - x^{(r)}_n|^2 < C_2.$$

We now consider $\bar{x}^{(k)} = (\cdots, 0, x^{(k)}_{-N}, \cdots, x^{(k)}_N, 0, \cdots)$ as a sequence in the finite dimensional space $\mathbb{R}^{2N+1}$. Since $\{\bar{x}^{(k)}\} \subset \mathbb{R}^{2N+1}$ is bounded and we are in finite dimensional space, it has a convergent subsequence (denoted the same for simplicity) therefore this subsequence is Cauchy. That implies that for every $\epsilon > 0$ there exists $M \in \mathbb{N}$ such that

$$\sum_{|n|\leq N} |x^{(k)}_n - x^{(r)}_n|^2 < \epsilon, \quad k, r > M. \tag{39}$$

We now write 37 (for the above subsequence) noting that

$$||\mathsf{T}x^{(k)} - \mathsf{T}x^{(r)}||^2_{\ell^2} \leq \hat{t} \left(\sum_{|n|\leq N} |x^{(k)}_n - x^{(r)}_n|^2 + \sum_{|n|>N} |x^{(k)}_n - x^{(r)}_n|^2 \right), \tag{40}$$

and taking $k, r > M$ both terms can be made smaller than ϵ, so that the subsequence $\mathsf{T}x^{(k)}$ is Cauchy in ℓ^2. Therefore $\mathsf{T} : \ell^2 \to \ell^2$ is compact.

(iv) We may repeat the above calculation in weighted spaces. In particular, for any weight sequence $\phi = \{\phi_n\}$ the weighted Cauchy-Schwartz inequality gives

$$\left(\sum_m t_{nm} x_m \right)^2 \leq \sum_m (\phi^{-2}_m |t^2_{nm}) \left(\sum_m \phi^2_m x^2_m \right), \quad n \in \mathbb{Z},$$

which upon multiplying by another weight sequence $\rho = \{\rho_n\}$ and adding over all $n \in \mathbb{Z}$ yields,

$$\sum_n \rho_n \left(\sum_m t_{nm} x_m \right)^2 \leq \left(\sum_n \sum_m \rho_n \phi_m^{-2} |t_{nm}|^2 \right) \left(\sum_m \phi_m^2 x_m^2 \right),$$

as long as all the sums are converging. Choosing $\phi_n = \sqrt{\rho_n}$, $n \in \mathbb{Z}$, the above gives

$$||\mathsf{T}x||_{\ell^2_\rho} \leq C_\rho \, ||x||_{\ell^2_\rho},$$

where

$$C_\rho^2 := \sum_n \sum_m \frac{\rho_n}{\rho_m} |t_{nm}|^2.$$

As long as $C_\rho < \infty$ the arguments in (iii) above may be repeated to yield the compactness of $\mathsf{T} : \ell^2_\rho \to \ell^2_\rho$.

(v) Since T is a bounded operator the result is standard and the representation

$$S(t) = I + t\,\mathsf{T} + \cdots + \frac{t^n}{n!}\mathsf{T}^n + \cdots = \sum_{n=0}^{\infty} \frac{t^n}{n!}\mathsf{T}^n,$$

holds. ■

Whether Assumption 2 holds or not depends on the range of the interactions as well as on the strength of the interactions. If the interactions coefficients t_{nm} are bounded, short range interactions, i.e., if only lattice sites m relatively close to n have a effect on the state at n, imply that the assumption holds. The range of the interactions is taken into account in the choice of the weight sequence ρ. For instance if we consider the discrete Laplacian then Assumption 2 holds for $\rho = 1$. In general $\rho_n < 1$ is needed to take into account long range interactions.

However there are interesting applications of the above framework in which the strength of the interaction may not necessarily be bounded. Consider for instance the case where the state space of the system is a separable Hilbert space $\mathbb{H}$, with an orthonormal basis $\{e_n\}$. As an example consider the Hilbert space $L^2([-\pi, \pi])$ of square integrable functions $x : [-\pi, \pi] \to \mathbb{R}$ and the standard Fourier basis. Then any operator $\mathsf{T} : D(\mathsf{T}) \subset \mathbb{H} \to \mathbb{H}$ (possibly unbounded) may have an (infinite) matrix representation in terms of the basis $\{e_n\}$ as $t_{nm} = \langle \mathsf{T}e_n, e_m \rangle$. This operator may then be considered as an operator $\hat{\mathsf{T}} : \ell^2_\phi \subset \ell^2 \to \ell^2$ for a properly chosen weight sequence ϕ (here ℓ^2_ϕ plays the role of the domain of the operator T, $D(\mathsf{T})$). If the original operator T is not bounded then the infinite matrix representation $\hat{\mathsf{T}}$ will not be a bounded operator either. When working with the new operator $\hat{\mathsf{T}}$ the framework developed in this paper can be used to obtain results for spatially interacting economies in continuous space.

Example 5. As a concrete example assume that the operator T is the standard Laplace operator, acting on a function $f : [-\pi, \pi] \to \mathbb{R}$ as $(\mathsf{T}f)(z) = \frac{\partial^2 f}{\partial z^2}$. Then, using the Fourier basis $e_n = \exp(in\pi z)$ and expressing $f(z) = \sum_n x_n e_n$ we may "understand" any element $f \in \mathbb{H} = L^2([-\pi, \pi])$ as the element $x = \{x_n\} \in \ell^2$. The action of the operator T on $f \in \mathbb{H}$ may thus be understood as the action of the (infinite) matrix operator $\hat{\mathsf{T}}$ on ℓ^2 where $(\hat{\mathsf{T}}x)_n = -n^2 x_n$, $n \in \mathbb{N}$. This operator is obviously unbounded. Its action is also defined not on every $x \in \ell^2$ but only for those x such that the series $\sum_n n^2 x_n$ converges in ℓ^2. This subset corresponds to $D(\mathsf{T})$ which is clearly a subset $\ell^2_\phi \subset \ell^2$ for properly selected weight sequence ϕ. This operator however generates a strongly continuous operator semigroup, $S(t) := e^{t\,\mathsf{T}}$ of linear bounded operators and furthermore if $x(t) := S(t)x_0$ then $x'(t) = \mathsf{T}x(t)$ and $x(0) = x_0$. Furthermore, this semigroup has smoothing properties.

For the case of more general (i.e. not bounded) operators T the following approach may be useful. Let us rewrite the action of the operator T in a slightly differen(but equivalent) form as

$$(\mathsf{T}x)_n = \sum_{m \in \mathfrak{N}_n} \bar{t}_{nm} \, (x_n - x_m)$$

where by $\mathfrak{N}_n$ we denote the neighbours of site n, i.e. all those $m \in Z$ such that $t_{nm} \neq 0$ and $m \neq n$. The set $\mathfrak{N}_n$ thus contains all the sites m interacting with n.

Example 6. As an example consider the discrete Laplacian $\mathsf{T} : \ell^2 \to \ell^2$ defined by $(\mathsf{T}x)_n = x_{n+1} - 2x_n + x_{n-1}$ (i.e. for given n, $t_{nm} = 1\delta_{m,n+1} - 2\delta_{m,n} + 1\,\delta_{m,n-1}$). For this case $\mathfrak{N}_n := \{n-1, n+1\}$. Then the discrete Laplacian can be expressed in equivalent form as $(\mathsf{T}x)_n = (x_{n+1} - x_n) + (x_{n-1} - x_n)$ (i.e. for given $\bar{t}_{nm} = \delta_{m,n-1} + \delta_{m,n+1}$).

With the operator T (expressed in the above form) we may define the bilinear form $\mathfrak{a}(\cdot,\cdot) : \ell^2 \times \ell^2 \to \mathbb{R}$ by

$$\mathfrak{a}(x,y) = \sum_n \sum_{m \in \mathfrak{N}_n} \bar{t}_{nm}(x_m - x_n)(y_m - y_n)$$

which may be used to define the operator T using the representation formula

$$\langle \mathsf{T}x, y \rangle = \mathfrak{a}(x,y), \quad \forall\, y \in \ell^2.$$

This definition makes sense as long as the bilinear form $\mathfrak{a}$ satisfies certain properties (see Assumption 4 below) and allows the proper (variational) definition of a large class of operators T which includes unbounded operators as well.

Assumption 4. Let ϕ , ρ be two weight sequences, such that $\ell^2_\phi \subseteq \ell^2_\rho$, the embedding being dense, and let the bilinear form $\mathfrak{a} : \ell^2_\phi \times \ell^2_\phi \to \mathbb{R}$ satisfy the following two properties:

(i) Continuity, i.e. there exists a costant C such that

$$|\mathfrak{a}(x,y)| \leq C \, ||x||_{\ell^2_\phi} ||y||_{\ell^2_\phi}, \;\; \forall\, x, y \in \ell^2_\phi.$$

(ii) Coercivity, i.e. there exists a constant $C > 0$, $C_0 \geq 0$ such that

$$|\mathfrak{a}(x,x)| + C_0 \, ||x||^2_{\ell^2_\rho} \geq C \, ||x||^2_{\ell^2_\phi}, \;\; \forall\, x \in \ell^2_\phi.$$

The following result follows easily from a classic abstract result in semigroup theory (see e.g. Dautray et al. (2000) Chapter XVII, § 6, Proposition 3, p. 380)

Lemma 3. *If Assumption 4 holds then the operator* T *generates a stronly continuous semigroup of linear operators, which is holomorphic (analytic) in* ℓ^2_ρ.

Assumption 4 is a very useful and important assumption since it allows us to define (fractional) powers of the operator T, and as a consequence provides certain very important smoothing properties (with respect to spatial structure) of the semigroup $S(t)$ generated by T. In particular (see e.g. Dautray et al. (2000) Chapter XVII, § 6, Proposition 4, p. 383) under Assumption 4 we have that the operator $\mathsf{T}^\alpha S(t)$ is bounded for all t and in particular $||\mathsf{T}^\alpha S(t)|| \leq C_\alpha \, t^{-\alpha} \, e^{-\delta t}$, $t > 0$, for $\alpha > 0$ and some $\delta \geq 0$. This property allows for smoothing of possible "rough" spatial data as a result of the action of the semigroup and this smoothing property as we shall see may turn to be extremely useful with regards to the well posedness of the resulting Hamilton-Jacobi-Bellman equation.

Remark 22. In assumption 4, one of the two weight functions may be chosen to be 1, if the structure of the operator permits it, in which case the corresponding Hilbert space is ℓ^2.

Example 7. An important class of operators that fall with the general class of operators that will be addressed in this work is that of diagonalizable operators. An example of a diagonalizable operator is the Laplacian, but of course there is a wide range of other important operators that fall in this class. For instance, any self adjoint operator with compact resolvent will be a diagonalizable operator. Such operators will necessarily be unbounded, however, they will be generators of C_0 semigroups, that will also be compact and analytic.

Example 8. Compact symmetric operators fall also within the general class of diagonalizable operators, however some care should be taken concerning the properties of the semigroups they generate. The kernel integral operators used in the modelling of agglomeration in spatial economics (Brock, Xepapadeas, and Yannacopoulos (Brock et al.)) as well as their discrete counterparts discussed here fall in this category. As compact operators, they generate strongly continuous semigroups, however, their properties with respect to spatial regularization properties may be limited.

8.2 Discrete Fourier Transform

Definition 2. Given a finite sequence of complex numbers $x = \{x_n\}$, $n = 0, \cdots, N-1$ we may define the discrete Fourier transform of the sequence as the sequence $\hat{x} - \mathfrak{F}x$ defined by

$$\hat{x}_k = (\mathfrak{F}x)_k := \sum_{n=0}^{N-1} x_n \exp\left(-i2\pi \frac{k}{N} n\right).$$

Definition 3. The inverse discrete Fourier transform is defined by

$$x_n = (\mathfrak{F}^{-1}\hat{x})_n := \frac{1}{N} \sum_{k=0}^{N-1} \hat{x}_k \exp\left(i2\pi \frac{k}{N} n\right).$$

The discrete Fourier transform has a nice interpretation in terms of the Fourier matrix

$$\mathbb{F}_N := \begin{pmatrix} 1 & 1 & 1 & 1 & \cdots & 1 \\ 1 & \omega_N & \omega_N^2 & \omega_N^3 & \cdots & \omega_N^{N-1} \\ 1 & \omega_N^2 & \omega_N^4 & \omega_N^6 & \cdots & \omega_N^{2(N-1)} \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 1 & \omega_N^{(N-1)} & \omega_N^{2(N-1)} & \omega_N^{3(N-1)} & \cdots & \omega_N^{(N-1)(N-1)} \end{pmatrix}$$

where $\omega_N := \exp\left(-i\frac{2\pi}{N}\right)$. Then the discrete Fourier transform $\mathfrak{F}x = \mathbb{F}_N x$ i.e. the Fourier operator in $\ell^2(\mathbb{Z}_N)$ can be represented in terms of the matrix $\mathbb{F}_N$. It is easy to show that the matrix $\mathbb{F}_N$ is invertible and

The discrete Fourier transform satisfies a discrete Plancherel formula (and related Parseval identity) according to which if $\hat{x} = \mathfrak{F}x$, $\hat{y} = \mathfrak{F}y$ then

$$\sum_{n=0}^{N-1} x_n y_n^* = \frac{1}{N} \sum_{k=0}^{N-1} \hat{x}_k \hat{y}_k^*$$

where $*$ denotes the complex conjugate. In the special case where $x = y$ the Plancherel formula can be used to show that the discrete Fourier transform defines an isometry on $\ell^2(\mathbb{Z}_N)$.

It is very useful in the study of the properties of the discrete Fourier transform to extend the definition of a vector $x \in \ell^2(\mathbb{Z}_N)$ by using its periodic extension $x(n+N) = x(n)$, $n \in \mathbb{Z}$. Then we may consider the elements of $\ell^2(\mathbb{Z}_N)$ as elements of $\ell^2(\mathbb{Z}$ i.e. as periodic sequences with period N. Adopting this convention, we will from now on denote by $\ell^2(\mathbb{Z}_N)$ the space of periodic sequences of period N, where of course square summability is required only in the fundamental domain.

We may also give the following important definition:

Definition 4. The translation operator $T_m : \ell^2(\mathbb{Z}_N) \to \ell^2(\mathbb{Z}_N)$, $m \in \mathbb{Z}$ is defined by $(T_m x)_n = x_{n-m}$.

Definition 5. A linear operator $\mathsf{A} : \ell^2(\mathbb{Z}_N \to \ell^2(\mathbb{Z}_N)$ is called translation invariant if it commutes with the translation operator, i.e. if $T_m \mathsf{A} = \mathsf{A}\, T_m$ for all $m \in \mathbb{Z}$.

An important example of a translation invariant operator is the discrete Laplacian $(\mathsf{A}x)_n = -\frac{1}{2}(x_n - x_{n-1}) - \frac{1}{2}(x_n - x_{n+1})$.

Translation invariant linear operators have a number of interesting properties. For example they have as eigenfunctions the vectors $\mathfrak{E}^{(m)} = (\mathfrak{e}_0^{(m)}, \cdots, \mathfrak{e}_{N-1}^{(m)})^{tr}$ where $\mathfrak{e}_r^{(m)} = \frac{1}{N}\exp\left(2\pi i \frac{m}{N} r\right)$. That means that for all $m \in \mathbb{Z}$ there exist $\lambda_m \in \mathbb{C}$ such that $\mathsf{A}\, \mathfrak{E}^{(m)} = \lambda_m\, \mathfrak{E}^{(m)}$. The sequence $\{\lambda_m\}$ is the sequence of eigenvalues of the operator A, which depends on the particular form of the operator. However, the eigenfunctions are independent of the particular form of the operator and are only determined by the translation invariant property of A! The set of eigenfunctions $\{\mathfrak{E}^{(m)}\}$, $m = 0, \cdots, N-1$ forms a basis for $\ell^2(\mathbb{Z}_N$, called the Fourier basis, and in this basis all translation invariant operators achieve a diagonal form. This simplifies considerably the study of dynamical systems involving translation invariant operators. Furthermore, the translation invariant property of operators implies an important structure when these are represented in the "standard" basis of $\ell^2(\mathbb{Z}_N)$, $\{e^{(m)}\}$, $e_r^{(m)} = \delta_{rm}$; in this basis the operator A is represented by a circulant matrix i.e., in terms of matrix $\{a_{nm}\}$ such that $a_{n+1,m+1} = a_{n,m}$, $n, m \in \mathbb{Z}$ (of course the matrix elements are "periodicized" in the obvious manner).

The properties of translation invariant operators can best be expressed in terms of the concept of the convolution:

Definition 6. Let $x, y \in \ell^2(\mathbb{Z}_N)$. By $\star$ we denote the convolution of x and y defined by

$$(x \star y)_n = \sum_{m=0}^{N} x_{n-m} y_m, \quad n \in \mathbb{Z}.$$

We may then define the linear operator $\mathcal{C}_x : \ell^2(\mathbb{Z}_N) \to \ell^2(\mathbb{Z}_N)$ by $\mathcal{C}_x y := x \star y$, called the convolution operator with kernel $x \in \ell^2(\mathbb{Z}_N)$.

Circulant matrices generate convolution operators in $\ell^2(\mathbb{Z}_N)$ and thus convolution operators are translation invariant operators. In fact the converse statement holds as well.

Proposition 13. *An operator $\mathsf{A} : \ell^2(\mathbb{Z}_N) \to \ell^2(\mathbb{Z}_N)$ is translation invariant if and only if it is a convolution operator.*

Convolution behaves very well with respect to the discrete Fourier transform:

Proposition 14. *The discrete Fourier transform turns convolutions into products:*

$$\mathfrak{F}(x \star y) = \mathfrak{F}x\, \mathfrak{F}y,$$

meaning that $(\mathfrak{F}(x \star y))_k = (\mathfrak{F}x)_k\, (\mathfrak{F}y)_k$, $k \in \mathbb{Z}$.

Under certain symmetry constraints on the sequence x the discrete Fourier transform $\hat{x}$ can be real valued. Since $(\mathfrak{F}x^*)_k = (\mathfrak{F}x)^*_{-k}$, it can be seen that x is real if and only if $(\mathfrak{F}x)_k = ((\mathfrak{F}x)_{-k})^*$. So if the discrete Fourier transform is applied on a real valued vector x then $\hat{x} = \mathfrak{F}x$ obeys the symmetry $\hat{x}_k = \hat{x}^*_{N-k}$, where all quantities are interpreted modulo N. That means that if we apply the discrete Fourier transform to a real vector then only "half" the Fourier amplitudes are needed.

Example 9. Examples of real vectors whose Fourier transform is real are vectors which are (real) linear combinations of the vectors

$$\mathfrak{C}^{(m)} := \Re(\mathfrak{E}^{(m)}) = \left(1, \cos\left(2\pi\frac{m}{N}\right), \cdots, \cos\left(2\pi\frac{nm}{N}\right), \cdots, \cos\left(2\pi\frac{(N-1)m}{N}\right)\right),$$

for $m = 0, \cdots, N-1$. Therefore, if $x \in span(\mathfrak{C}^{(m)}; m = 0, \cdots, N-1)$ then $\hat{x}$ is a real valued vector (of course by span we mean the space of real linear combinations).

Example 10. Examples of real vectors whose Fourier transform is imaginary are vectors which are (real) linear combinations of the vectors

$$\mathfrak{S}^{(m)} := \Im(\mathfrak{E}^{(m)}) = \left(0, \sin\left(2\pi\frac{m}{N}\right), \cdots, \sin\left(2\pi\frac{nm}{N}\right), \cdots, \sin\left(2\pi\frac{(N-1)m}{N}\right)\right),$$

for $m = 0, \cdots, N-1$. Therefore, if $x \in span(\mathfrak{S}^{(m)}; m = 0, \cdots, N-1)$ then $\hat{x}$ is an imaginary valued vector (of course by span we mean the space of real linear combinations).

8.3 Discrete Fourier transform in $\mathbb{Z}$

Definition 7 (Convolution)**.** Let $\{a_n\}$ and $\{b_n\}$, $n \in \mathbb{Z}$ be two sequences. The convolution of $a \star b$ is defined as the sequence

$$(a \star b)_n = \sum_{m\in\mathbb{Z}} a_{n-m} b_m,$$

whenever the infinite sum is well defined.

The following discrete version of Young's inequality, provides conditions under which the discrete convolution is well defined:

Proposition 15 (Young inequality)**.** *Let $a \in \ell^1(\mathbb{Z})$ and $b \in \ell^p(\mathbb{Z})$, $1 \le p \le infty$. Then the discrete convolution $a \star b$ is well defined as an element of $\ell^p(\mathbb{Z})$ and satisfies the estimate*

$$||a \star b||_{\ell^p(\mathbb{Z})} \le ||a||_{\ell^1(\mathbb{Z})} \, ||b||_{\ell^p(\mathbb{Z})}.$$

Definition 8 (Fourier transform in $\ell^2(\mathbb{Z})$)**.** Let $x \in \ell^2(\mathbb{Z})$. The Fourier transform of this sequence is a function on the unit circle $\mathbb{S}^1$ defined by

$$(\mathfrak{F}x)(\mathcal{K}) := \hat{x}(\mathcal{K}) := \sum_{n\in\mathbb{Z}} x_n e^{in\mathcal{K}}, \ \ \mathcal{K} \in [-\pi, \pi]$$

or equivalently

$$(\mathfrak{F}x)(z) := \hat{x}(z) := \sum_{n\in\mathbb{Z}} x_n z^n, \ \ z \in \Gamma = \partial\mathbb{D},$$

where $\mathbb{D} := \{z \in \mathbb{C} : |z| \le 1$ and $\Gamma = \{z \in \mathbb{C} : |z| = 1$.

The Fourier transform in $\ell^2(\mathbb{Z})$ can be considered as a map $mathfrakF : \ell^2(\mathbb{Z}) \to L^2([-\pi, \pi])$ (equivalently $L2(\Gamma)$) which is an invertible isometry. The inverse is the inverse Fourier transform, defined by

$$x_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-i\mathcal{K}n}\,(\mathfrak{F}x)(\mathcal{K})\,d\mathcal{K} = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-i\mathcal{K}n}\,\hat{x}(\mathcal{K})\,d\mathcal{K}$$

while the isometry is the Plancherel formula

$$||x||^2_{\ell^2(\mathbb{Z})} = \sum_{n\in\mathbb{Z}} |x_n|^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} |\hat{x}(\mathcal{K})|^2 d\mathcal{K} = ||\hat{x}||^2_{L^2([-\pi,\pi])}.$$

Remark 23. The Fourier transform and its inverse can be defined under weaker assumptions on the sequence x, i.e. they can be defined if $x \in \ell^1(\mathbb{Z})$ (note of course that $\ell^2(\mathbb{Z}) \subset \ell^1(\mathbb{Z})$), the definitions being formally the same.

The Fourier transform in $\ell^2(\mathbb{Z})$ has the same convenient properties with respect to convolutions as its finite counterpart.

Proposition 16. *Assume that $a \star b$ is well defined (see Proposition 15) as well as the Fourier transforms of a, b, $a \star b$. Then,*

$$(\mathfrak{F}(a \star b))(\mathcal{K}) = (\mathfrak{F}a)(\mathcal{K})(\mathfrak{F}b)(\mathcal{K}), \;\; \mathcal{K} \in [-\pi, \pi],$$

or

$$\widehat{a \star b}(\mathcal{K}) = \hat{a}(\mathcal{K})\hat{b}(\mathcal{K}), \;\; \mathcal{K} \in [-\pi, \pi].$$

The proof of the above essentially relies on the application of the Cauchy product formula for series.

8.4 Diagonalizable operators

Definition 9 (Riesz basis)**.** Let $\mathbb{H}$ be a Hilbert space. A sequence $\{\phi_n\} \subset \mathbb{H}$ is called a Riesz basis if there exists a invertible and bounded operator $\mathsf{T} : \mathbb{H} \to \ell^2$ such that $\mathsf{T}\phi_n = e_n$, $n \in \mathbb{N}$, where $\{e_n\}$ is the standard orthonormal basis in ℓ^2.

If $\{\phi_n\}$ is a Riesz basis of $\mathbb{H}$ the sequence $\{\tilde{\phi}_n\}$ defined by $\tilde{\phi}_n := \mathsf{T}^*\mathsf{T}\phi_n$, $n \in \mathbb{N}$ is also a Riesz basis called the biorthogonal sequence to $\{\phi_n\}$.

Definition 10. An operator (possibly unbounded) $\mathsf{A} : D(\mathsf{A}) \to \mathbb{H}$ is called diagonalizable if $\rho(\mathsf{A}) \neq \emptyset$ and there exists a Riesz basis $\{\phi_n\}$ in $\mathbb{H}$ that consists of the eigenvectors of A, i.e. these exists a sequence $\{\lambda_n\} \subset \mathbb{C}$ such that $\mathsf{A}\phi_n = \lambda_n\phi_n$.

For a diagonalizable operator $D(\mathsf{A})$ is dense in $\mathbb{H}$. A diagonalizable operator can be expressed as a infinite series in terms of the spectrum (i.e. the sequence of eigenvalues $\{\lambda n\}$), the Riesz basis $\{\phi_n\}$ and the biorthogonal basis $\{\tilde{\phi}_n\}$:

$$\mathsf{A}x := \sum_n \lambda_n \langle x, \tilde{\phi}_n\rangle \, \phi_n,$$

and this holds for every x for which this expression makes sense, i.e., for every $x \in D(\mathsf{A})$ with

$$D(\mathsf{A}) := \{x \in \mathbb{H} \, : \, \sum_n (1 + |\lambda_n|^2)|\langle x, \tilde{\phi}_n\rangle|^2 < \infty \}.$$

Proposition 17. *The diagonalizable operator A defined as above is bounded if and only if the sequence $\{\lambda_n\}$ is bounded. In this case the operator A is defined on the whole of $\mathbb{H}$.*

However, results may be provided for the unbounded case as well.

Proposition 18. *The diagonalizable operator A is the generator of a strongly continuous (C_0) semigroup on $\mathbb{H}$ if and only if $\sup_n Re(\lambda_n) < \infty$. In this case, the semigroup $S(t)$ generated by A is also diagonalizable and can be expressed in terms of the infinite series*

$$S(t) = \sum_n e^{\lambda_n t}\langle x, \tilde{\phi}_n\rangle \, \phi_n.$$

Proposition 19. *A diagonalizable operator A is compact if and only if $\lim_n \lambda_n = 0$.*

8.5 Connections with eigenfunction expansions

The general case when there is no translation invariance can be treated using expansions in terms of eigenfunctions of the operator A.

To make this point more clear, we present here the equivalence of this approach with the Fourier transform approach, in the special case of translation invariant systems in the finite lattice case $\mathbb{Z}_N$ with an observation that will prove very useful in treating the case where the translation invariance property does not hold. If A, B and C are translation invariant then they have the eigenfunctions $\mathfrak{E}_k$, $k = 0, 1, \cdots, N-1$ which form a complete set for $\ell^2(\mathbb{Z}_N)$. Note that all the operators have the same eigenfunctions (as a result of the translation invariance property) and the only quantities that depend on the actual form of the operator are the eigenvalues, which by the general theory of translation invariant operators are the coordinates of the vectors $\hat{a}$, $\hat{b}$ and $\hat{c}$ respectively. For example, the eigenvalue of A corresponding to the eigenfunction $\mathfrak{E}_k$ for some $k \in \{0, 1, \cdots, N-1)\}$ is $\hat{a}_k$. This information allows us to view the Fourier transform method of solving the problem in the case of translation invariant systems in a different manner that allows for generalization even in the case where the operators are no longer translation invariant.

Since the set $\mathcal{E} := \{\mathfrak{E}_k,\ k = 0, \cdots, N-1\}$ forms an orthogonal basis of $\ell^2(\mathbb{Z}_N)$ then any element of $\ell^2(\mathbb{Z}_N)$ can be expanded as a linear combination of the elements of the set $\mathcal{E}$. Therefore, for any t, the solution of (2), $x(t)$ can be expressed as the linear combination $x(t) = \sum_{k=0}^{N-1} \bar{x}_k(t)\mathfrak{E}_k$ where $\bar{x}_k$ is a scalar and $\mathfrak{E}_k$ is a vector in $\ell^2(\mathbb{Z}_N)$, so that $x(t) \in \ell^2(t)$, with similar expansions for $u(t)$ and $v(t)$. Since $\mathfrak{E}_k$ is an eigenfunction of the operators A, B, C, with eigenvalues $\hat{a}_k$, $\hat{b}_k$, $\hat{c}_k$ respectively, we can see straightaway that the action of these operators on $x(t)$ is very simple and in particular

$$\begin{aligned}
\mathsf{A}x(t) &= \sum_{k=0}^{N-1} \hat{a}_k \bar{x}_k(t)\mathfrak{E}_k, \\
\mathsf{B}u(t) &= \sum_{k=0}^{N-1} \hat{b}_k \bar{u}_k(t)\mathfrak{E}_k, \\
\mathsf{C}v(t) &= \sum_{k=0}^{N-1} \hat{c}_k \bar{v}_k(t)\mathfrak{E}_k,
\end{aligned}$$

therefore these operators are diagonal in this basis. In particular all three operators are diagonalizable in the same basis (as a consequence of the translation invariance). In a similar fashion we may assume that $w(t)$ is expressed in this basis as

$$w(t) = \sum_{k=0}^{N-1} \hat{w}_k \mathfrak{E}_k,$$

so that

$$\mathsf{C}w(t) = \sum_{k=0}^{N-1} \hat{c}_k \hat{w}_k \mathfrak{E}_k.$$

We now substitute the above expansions in the Ornstein-Uhlenbeck equation (2), and then project on the various $\mathfrak{E}_k$ for $k = 0, 1, \cdots, N-1$, using the orthogonality to obtain the set of equations for the scalar amplitudes $\bar{x}(k)$:

$$d\bar{x}(k) = (\hat{a}_k \bar{x}_k(t) + \hat{b}_k \bar{u}_k(t) + \hat{c}_k \bar{v}_k(t))dt + \hat{c}_k d\bar{w}_k(t), \quad k \in \mathbb{Z}_N$$

which of course is nothing else than the Fourier transform of (2). However, we have arrived to this equation via a different procedure, which was expansion of the solution in terms of a complete orthogonal set of $\ell^2(\mathbb{Z}_N)$ and then projection on the elements of the basis. This approach, known as the

Galerkin approximation is very general and does not require at all translation invariance, in contrast with the Fourier transform. The fact that this basis may be chosen as consisting of the common eigenfunctions of all the operators involved is immaterial, apart from the fact that renders a projected system which is diagonal (fully decoupled) and therefore very easy to treat analytically. Of course the possibility of choosing the basis like that stems from the translation invariance of the operators involved. However, even if the projection leads to coupled systems the Galerkin expansion is a very useful tool both for showing well posedeness of the system as well as for proving qualitative results through perturbation theory arguments. We will use this interpretation of the Fourier transform later on in this paper, when we treat the general linear quadratic problem in the absence of translation invariance.

8.6 Linear quadratic problem for the infinite lattice $\mathbb{Z}$

The results of the previous section on translation invariant systems in the finite lattice $\mathbb{Z}_N$ can be generalized to the case of translation invariant systems in the infinite lattice $\mathbb{Z}$. The translation operator $T : \ell^2(\mathbb{Z}) \to \ell^2(\mathbb{Z})$ is now defined as $T_m x_n = x_{n-m}$, for all $m \in \mathbb{Z}$ and the invariance condition now becomes $T_m \mathsf{A} = \mathsf{A} T_m$ for all $m \in \mathbb{Z}$. Now the periodicity assumption on the lattice is waived and the dynamical system is a genuine infinite dimensional system. The choice of $\ell^2(\mathbb{Z})$ as state space imposes restrictions on the behavior of the dynamical systems at infinity (i.e. at $n \to \pm\infty$) and in particular implies that the dynamics decay fast enough for large n so that the sequence $\{x_n\}$ is square summable[14].

On the infinite lattice $\mathbb{Z}$ we may define a convolution operator as

$$(\mathsf{A}x)_n = \sum_{m \in \mathbb{Z}} a_{n-m} x_n$$

where a is a sequence in ℓ^∞. This defines an operator $\mathsf{A} : \ell^2 \to \ell^2$ (from now on for simplicity of notation we use $\ell^2 = \ell^2(\mathbb{Z})$). Note that $a \in \ell^\infty$ is the weaker condition we may impose on the sequence $\{a_n\}$. Clearly, under the strongest condition $a \in \ell^2$ (strongest since $\ell^2 \subset \ell^\infty$) the above statements are still true.

We now need to use the Fourier transform on $\mathbb{Z}$, rather than on $\mathbb{Z}_N$. This is defined as an infinite sum

$$(\mathfrak{F}x)(\mathcal{K}) = \hat{x}(\mathcal{K}) = \sum_{n \in \mathbb{Z}} x_n e^{-in\mathcal{K}}, \;\; \mathcal{K} \in [-\pi, \pi].$$

Note that now the Fourier transform $\hat{x}$ depends on a continuous variable $\mathcal{K} \in [-\pi, \pi]$ and not on a discrete variable $k \in \mathbb{Z}_N$ as in the case of the finite lattice. Therefore, the Fourier transform can no longer be considered as a function on the same set as the original vector, but as a function on a "dual" set[15]. For a further discussion of this, in more general settings we refer to Section 4.4.

Using general results of abstract harmonic analysis it is possible to show that in the infinite dimensional case as well the Fourier transform turns the convolution into a product, therefore under

[14] Other choices can be possible (such as e.g. weighted ℓ^2 spaces) which will keep the theory within the context of Hilbert space. However these impose subtleties on the concept of translation invariance which we prefer not to address here as they require abstract mathematical notions such as the concept of operator algebras that are clearly beyond the scope of the present work. Furthermore, we may relax the square summability assumption and use other sequence spaces as phase space for the infinite dimensional dynamical system, but that will require us to consider the dynamics in Banach space rather than Hilbert space. For instance we might relax square summability and assume only the state is bounded at infinity, thus leading to the choice of $\ell^\infty(\mathbb{Z})$ as phase space for the dynamics. Working in Banach space instead of Hilbert space introduces new subtleties into the problem, and requires more involved analysis, which again is outside the scope of the present work.

[15] $\mathbb{Z}$ is the original set while $[-\pi, \pi] \simeq \Gamma$ where $\simeq$ means homeomorphic and Γ is the unit circle in the complex domain is the dual set. This phenomenon holds in general frameworks and is called Pontryagin duality (see Section 4.4).

the assumption that the operators A, B, C are translation invariant (convolution operators) we may apply the Fourier transform to the state equation (2) and obtain the Fourier transformed problem

$$d\hat{x}(t;\mathcal{K}) = (\hat{a}(\mathcal{K})\hat{x}(t;\mathcal{K}) + \hat{b}(\mathcal{K})\hat{u}(t;\mathcal{K}) + \hat{c}(\mathcal{K})\hat{v}(t;\mathcal{K}))dt + \hat{c}(\mathcal{K})d\hat{w}(t;\mathcal{K}), \;\; \mathcal{K} \in [-\pi,\pi]. \tag{41}$$

Note that this is an infinite family of stochastic differential equations, indexed by $\mathcal{K} \in [-\pi,\pi]$, which is fully decoupled in $\mathcal{K}$. For each $\mathcal{K}$ we may solve the above system and obtain $\hat{x}(t;\mathcal{K})$ and then invert the Fourier transform and regain the solution in physical space by performing the integration

$$x_n(t) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \hat{x}(t;\mathcal{K})e^{-in\mathcal{K}}d\mathcal{K}, \;\; n \in \mathbb{Z}.$$

Furthermore using the Plancherel formula, in the special case where $\mathsf{P} = pI$, $\mathsf{Q} = qI$ and $\mathsf{R} = \theta I$ we may rewrite the control functional as

$$\mathcal{J} := \frac{1}{2\pi}\,\mathbb{E}_Q\left[\int_{-\pi}^{\pi}\int_0^{\infty} e^{-rt}p(\hat{x}(t;\mathcal{K}))^2 + q(\hat{u}(t;\mathcal{K}))^2 - \theta(\hat{v}(t;\mathcal{K}))^2)dt\, d\mathcal{K}\right]$$

subject to the (decoupled) Fourier space dynamics (41).

The very simple form of the system now allows a full treatment. In particular if for each $\mathcal{K} \in [-\pi,\pi]$ we may solve the optimal control problem of

$$\min_{\hat{u}(\cdot;\mathcal{K})}\max_{\hat{v}(\cdot;\mathcal{K})}\mathbb{E}_Q\left[\int_0^{\infty} e^{-rt}p(\hat{x}(t;\mathcal{K}))^2 + q(\hat{u}(t;\mathcal{K}))^2 - \theta(\hat{v}(t;\mathcal{K}))^2)dt\right]$$

subject to the (decoupled) Fourier space dynamics (41) then this solution will also be a solution of the original problem. This is a family (of decoupled) optimal control problems, each member of the family indexed by $\mathcal{K} \in [-\pi,\pi]$. Fixing $\mathcal{K} \in [-\pi,\pi]$ we may solve the problem for each particular $\mathcal{K}$ by solution of the HJBI equations. In general the Fourier space dynamics (41) will be complex valued, thus leading to representation of the optimal control in terms of the complex valued Riccati equation. Though a bit messier in terms of algebraic manipulations, this does not constitute any major conceptual difficulties.

Since we later on treat the problem in its full generality, to facilitate our analysis and make the results more transparent to the reader, we stick to the particular case that the Fourier transformed system is real and admits real solutions. Clearly if x is a real valued sequence, such that $x_n = x_{-n}$, for all $n \in \mathbb{Z}$ then $\hat{x}(\mathcal{K}) \in \mathbb{R}$ for every $\mathcal{K} \in [-\pi,\pi]$. This condition means that the spatial patterns we are interested in are symmetric about the origin $n = 0$.

Let us assume that the Wiener process w is chosen so that it admits the same spatial symmetry $w_n(t) = w_{-n}(t)$. That means that the noise factors affecting the economy involve $n = 0, 1, \cdots$ independent Wiener processes, that correspond to uncertainty factors concerning the sites $n \in \mathbb{Z}_+ \cup \{0\}$, while the uncertainty factors concerning the remaining sites $n \in \mathbb{Z}_-$ are correlated with the previously mentioned factors, and in particular the uncertainty corresponding to site $-n$ coincides with the uncertainty corresponding to site $-n$. This symmetry condition guarantees that the Fourier transform of the Wiener process is real for all t, $\hat{w}(t;\mathcal{K}) \in \mathbb{R}$ for all $\mathcal{K} \in [-\pi,\pi]$. Furthermore, since on account of symmetry

$$\hat{w}(t;\mathcal{K}) := \sum_{n\in Z} w_n(t)e^{in\mathcal{K}} = w_0(t) + 2\sum_{n\in\mathbb{Z}_+} w_n(t)cos(n\mathcal{K}), \;\; \mathcal{K} \in [-\pi,\pi]$$

an application of Levy theorem implies that $\hat{w}(t;\mathcal{K})$ is a family of Wiener processes (indexed by $\mathcal{K} \in [-\pi,\pi]$ with mean 0 and variance

$$\sigma^2(t;\mathcal{K}) = t\left(\sum_{n\in Z} e^{i2n\mathcal{K}}\right) = t\left(1 + 4\sum_{n\in\mathbb{Z}_+} cos^2(n\mathcal{K})\right), \;\; \mathcal{K} \in [-\pi,\pi].$$

Recall the Poisson summation formula

$$\sum_{n\in\mathbb{Z}} \delta_{2n\pi}(\mathcal{K}) = \frac{1}{2\pi}\sum_{n\in\mathbb{Z}} e^{in\mathcal{K}}$$

an application of which yields

$$\sigma^2(t;\mathcal{K}) = 2\pi t \sum_{n\in\mathbb{Z}} \delta_{2n\pi}(2\theta) = 2\pi t \sum_{n\in\mathbb{Z}} \delta_{n\pi}(\mathcal{K})$$

meaning that $\hat{w}(t;\mathcal{K})$ has zero variance unless $\mathcal{K} = n\pi$, $n \in Z$ in which case the variance is $2\pi t$. But since $\mathcal{K} \in [-\pi,\pi]$ the only members of the family $\hat{w}(t;\mathcal{K})$ of non zero variance are the ones corresponding to $\mathcal{K} = -\pi$, $\mathcal{K} = 0$ and $\mathcal{K} = \pi$. The meaning of that is that out of the family of stochastic differential equation (41), the members of the family corresponding to $\mathcal{K} = \pi, 0, \pi$ are stochastic differential equations while the rest are deterministic differential equations. This is because of symmetry, in the Fourier space $\mathcal{K} \in \Gamma$ (the dual space) the combined effect of all uncertainty vanishes.

Therefore, in this symmetric case we have in Fourier space

$$d\hat{x}(t;\mathcal{K}) = \begin{cases} (\hat{a}(\mathcal{K})\hat{x}(t;\mathcal{K}) + \hat{b}(\mathcal{K})\hat{u}(t;\mathcal{K}) + \hat{c}(\mathcal{K})\hat{v}(t;\mathcal{K}))dt & \mathcal{K} \neq -\pi, 0, \pi, \\ (\hat{a}(\mathcal{K})\hat{x}(t;\mathcal{K}) + \hat{b}(\mathcal{K})\hat{u}(t;\mathcal{K}) + \hat{c}(\mathcal{K})\hat{v}(t;\mathcal{K}))dt + \sqrt{2\pi}\hat{c}(\mathcal{K})d\mathfrak{w}(t;\mathcal{K}), & \mathcal{K} = -\pi, 0, \pi, \end{cases} \tag{42}$$

where $\mathfrak{w}(\cdot;\mathcal{K})$ is a standard Wiener process.

Therefore, as long as the convolution kernels a, b, c of the translation invariant operators A, B and C, the Wiener process w, the initial condition $x(0)$ share this symmetry property and if we restrict the control processes u and v to the space of processes with this symmetry, then the solution of the decoupled system (41) (or equivalently (42)) will have the same symmetry for all times.

Remark 24. Note, that this "disappearance" of noise and uncertainty happens in Fourier space, and not in the real space; therefore, it does **not** mean that there is no uncertainty is certain sites in real space

The above discussion implies that an analogue of Propositions 4, 5 holds for the infinite lattice $\mathbb{Z}$, and repeating the analysis that led to Proposition 6 we obtain the following:

Proposition 20 (Solution of primal problem)**.** *For fixed $\mathcal{K} \in [-\pi,\pi]$ the solution of the primal problem*

$$\min_{\hat{u}(\cdot,\mathcal{K})}\max_{\hat{v}(\cdot,\mathcal{K})} \mathbb{E}_Q\left[\int_0^\infty e^{-rt}p(\hat{x}(t;\mathcal{K}))^2 + q(\hat{u}(t;\mathcal{K}))^2 - \mathcal{K}(\hat{v}(t;\mathcal{K}))^2)dt\right], \tag{43}$$

subject to (42) is given by the optimal state equation

$$d\hat{x}^*(t;\mathcal{K}) = R(\mathcal{K})\hat{x}^*(t,\mathcal{K})dt + \hat{c}(t,\mathcal{K})\sigma(t;\mathcal{K})d\mathfrak{w}(t) \tag{44}$$

where

$$R(\mathcal{K}) := \hat{a}(\mathcal{K}) - \frac{\hat{b}^2(\mathcal{K})M_2(\mathcal{K})}{2q} + \frac{\hat{c}^2(\mathcal{K})\,M_2(\mathcal{K})}{2\theta}.$$

and $M_{2,k}$ is the solution of

$$\left(\frac{\hat{c}^2(\mathcal{K})}{2\theta} - \frac{\hat{b}^2(\mathcal{K})}{2q}\right) M_2^2(\mathcal{K}) + (2\,\hat{a}(\mathcal{K}) - r)\,M_2(\mathcal{K}) + 2p = 0. \tag{45}$$

The optimal controls are given by the feedback laws

$$\hat{u}^*(\mathcal{K}) = -\frac{\hat{b}(\mathcal{K})M_2(\mathcal{K})}{2q}\hat{x}^*(\mathcal{K}), \qquad \hat{v}^*(\mathcal{K}) = \frac{\hat{c}(\mathcal{K})\,M_2(\mathcal{K})}{2\theta}\hat{x}^*(\mathcal{K}).$$

The infinite lattice problem has certain complications not encountered in the finite lattice. For example, the solution we have obtained in Proposition 20 for fixed $\mathcal{K}$ must be such that it is a valid Fourier transform of an $x \in \ell^2(\mathbb{Z})$. That means that $\hat{x}(t;\mathcal{K})$ must have such regularity so that the integral representation

$$x_n(t) = \frac{1}{2\pi}\int_{\pi}^{\pi} \hat{x}(t;\mathcal{K})e^{-in\mathcal{K}}d\mathcal{K}, \ \ n \in \mathbb{Z}$$

is well defined. A quick check comes through Parceval's identity, which yields (formally) that for all t, it should hold that

$$\mathbb{E}_Q[\sum_{n\in\mathbb{Z}} x_n^2(t)] = \frac{1}{2\pi}\mathbb{E}_Q\left[\int_{-\pi}^{\pi} \hat{x}^2(t;\mathcal{K})d\mathcal{K}\right] \tag{46}$$

(since we have restricted to the case where both x and $\hat{x}$ are real). Therefore, the optimal path is well defined in $\ell^2(\mathbb{Z})$ for all t as long as the integral in the right hand side is well defined for all t. This imposes integrability conditions on $\hat{x}(t;\mathcal{K})$ as a function of $\mathcal{K}$, and in particular it imposes the condition $\hat{x}(t;\cdot) \in L^2([-\pi,\pi])$ (as a function of $\mathcal{K}$) for all t. As one may see this imposes constraints on $M_2(\mathcal{K})$, which through the quadratic equation (45) leads to constraints on $\hat{a}(\mathcal{K})$, $\hat{b}(\mathcal{K})$ and $\hat{c}(\mathcal{K})$, which in turn correspond to constraints on the operators A B and C. These constraints can be best understood in the next section, where the concept of a hot spot is introduced, which is related to possible breakdowns of the robust control procedure, or spatial pattern formation.

8.7 The simplest possible coupled system: Two sites

We present here a completely worked example of the simplest possible coupled system. This consists of two sites, north and west, coupled through the state equation, where now the operators A, B, C are 2×2 matrices of the form

$$\mathsf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \ \ \mathsf{B} = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}, \ \ \mathsf{C} = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}.$$

We have assumed, without loss of generality, that there are as many risk factors as sites.

We furthermore assume that

$$\mathsf{P} = \begin{pmatrix} p_1 & 0 \\ 0 & p_2 \end{pmatrix}, \ \ \mathsf{Q} = \begin{pmatrix} q_1 & 0 \\ 0 & q_2 \end{pmatrix}, \ \ \mathsf{R} = \begin{pmatrix} \theta_1 & 0 \\ 0 & \theta_2 \end{pmatrix}.$$

This choice corresponds to the localized entropic robust control problem, in the case of a 2 site lattice.

We may look for a solution of the HJBI equation of the form

$$V(x_1,x_2) = \frac{1}{2}\sum_{i=1}^{2}\sum_{j=1}^{2}\bar{H}_{ij}x_ix_j + \sum_{i=1}^{2} j_ix_i + K.$$

A quick calculation shows that $j_1 = j_2 = 0$. Furthermore the value function may be expressed in the symmetric form

$$V(x_1,x_2) = \frac{1}{2}H_{11}x_1^2 + H_{12}x_1x_2 + H_{22}x_2^2 + K$$

in terms of the symmetric 2×2 matrix

$$\mathsf{H} = \begin{pmatrix} H_{11} & H_{12} \\ H_{12} & H_{22} \end{pmatrix},$$

where

$$H_{11} = \bar{H}_{11}, \;\; H_{22} = \bar{H}_{22}, \;\; H_{12} = \frac{1}{2}(\bar{H}_{12} + \bar{H}_{21}).$$

The symmetric matrix H is the solution of the matrix Riccati equation

$$\mathsf{H}A + A^{tr}\mathsf{H} - \mathsf{HEH} - r\mathsf{H} + \mathsf{P} = 0$$

(the adjoint operator now corresponds to the transpose matrix) where E is the matrix

$$\mathsf{E} = \frac{1}{2}(\mathsf{B}^{tr}\mathsf{Q}^{-1}\mathsf{B} + \mathsf{B}\mathsf{Q}^{-1}\mathsf{B}^{tr} - \mathsf{C}\mathsf{R}^{-1}\mathsf{C}^{tr} - \mathsf{C}^{tr}\mathsf{R}^{-1}\mathsf{C})$$

This is a symmetric 2×2 matrix that can be explicitly calculated. Its elements are given by

$$\begin{aligned}
e_{11} &= \frac{b_{11}^2}{q_1} + \frac{b_{12}^2 + b_{21}^2}{2q_2} - \frac{c_{11}^2}{\theta_1} - \frac{c_{12}^2 + c_{21}^2}{2\theta_2}, \\
e_{12} &= \frac{b_{11}(b_{12} + b_{21})}{2q_1} + \frac{b_{22}(b_{12} + b_{21})}{2q_2} - \frac{c_{11}(c_{12} + c_{21})}{2\theta_1} - \frac{c_{22}(c_{12} + c_{21})}{2\theta_2}, \\
e_{22} &= \frac{b_{12}^2 + b_{21}^2}{2q_1} + \frac{b_{22}^2}{q_2} - \frac{c_{12}^2 + c_{21}^2}{2\theta_1} - \frac{c_{22}^2}{\theta_2}
\end{aligned}$$

The feedback laws are of the form

$$\begin{aligned}
u_1 &= -\frac{1}{4q_1}\left((2H_{11}b_{11} + H_{12}b_{21} + H_{21}b_{21})x_1 + (H_{12}b_{11} + H_{21}b_{11} + 2H_{22}b_{21})x_2\right), \\
u_2 &= -\frac{1}{4q_2}\left((2H_{11}b_{12} + H_{12}b_{22} + H_{21}b_{22})x_1 + (H_{12}b_{12} + H_{21}b_{12} + 2H_{22}b_{22})x_2\right), \\
v_1 &= \frac{1}{4\theta_1}\left((2H_{11}c_{11} + H_{12}c_{21} + H_{21}c_{21})x_1 + (H_{12}c_{11} + H_{21}c_{11} + 2H_{22}c_{21})x_2\right), \\
v_2 &= \frac{1}{4\theta_2}\left((2H_{11}c_{12} + H_{12}c_{22} + H_{21}c_{22})x_1 + (H_{12}c_{12} + H_{21}c_{12} + 2H_{22}c_{22})x_2\right)
\end{aligned}$$

and the optimal path is given by the solution of the system

$$\begin{aligned}
dx_1 &= (R_{11}x_1 + R_{12}x_2)dt + c_{11}dw_1 + c_{12}dw_2, \\
dx_2 &= (R_{21}x_1 + R_{22}x_2)dt + c_{21}dw_1 + c_{22}dw_2
\end{aligned}$$

where

$$\begin{aligned}
R_{11} = a_{11} &- \frac{b_{11}}{4q_1}(2H_{11}b_{11} + H_{12}b_{21} + H_{21}b_{21}) - \frac{b_{12}}{4q_2}(2H_{11}b_{12} + H_{12}b_{22} + H_{21}b_{22}) \\
&+ \frac{c_{11}}{4\theta_1}(2H_{11}c_{11} + H_{12}c_{21} + H_{21}c_{21}) + \frac{c_{12}}{4\theta_2}(2H_{11}c_{12} + H_{12}c_{22} + H_{21}c_{22}),
\end{aligned}$$

$$\begin{aligned}
R_{12} = a_{12} &- \frac{b_{11}}{4q_1}(H_{12}b_{11} + H_{21}b_{11} + 2H_{22}b_{21}) - \frac{b_{12}}{4q_2}(H_{12}b_{12} + H_{21}b_{12} + 2H_{22}b_{22}) \\
&+ \frac{c_{11}}{4\theta_1}(H_{12}c_{11} + H_{21}c_{11} + 2H_{22}c_{21}) + \frac{c_{12}}{4\theta_2}(H_{12}c_{12} + H_{21}c_{12} + 2H_{22}c_{22}),
\end{aligned}$$

$$\begin{aligned}
R_{21} = a_{21} &- \frac{b_{21}}{4q_1}(2H_{11}b_{11} + H_{12}b_{21} + H_{21}b_{21}) - \frac{b_{22}}{4q_2}(2H_{11}b_{12} + H_{12}b_{22} + H_{21}b_{22}) \\
&+ \frac{c_{21}}{4\theta_1}(2H_{11}c_{11} + H_{12}c_{21} + H_{21}c_{21}) + \frac{c_{22}}{4\theta_2}(2H_{11}c_{12} + H_{12}c_{22} + H_{21}c_{22}),
\end{aligned}$$

$$R_{22} = a_{22} - \frac{b_{21}}{4q_1}(2H_{11}b_{11} + H_{12}b_{21} + H_{21}b_{21}) - \frac{b_{22}}{4q_2}(2H_{11}b_{12} + H_{12}b_{22} + H_{21}b_{22})$$
$$+ \frac{c_{21}}{4\theta_1}(2H_{11}c_{11} + H_{12}c_{21} + H_{21}c_{21}) + \frac{c_{22}}{4\theta_2}(2H_{11}c_{12} + H_{12}c_{22} + H_{21}c_{22}).$$

The long run behavior of the optimal system, in average is given by the spectrum of the 2×2 matrix $\mathcal{R} = (R_{ij})$. A detailed parametric study will give us the values of the parameters for which hot spot formation is possible.

It is interesting to note that even in this simple case it is not possible to obtain the solution of the Riccati equation in closed form, as we did in the special case where the operators are translation invariant (and therefore the matrix Riccati equation degenerated to a scalar quadratic equation). Of course numerical solution is feasible, and may reveal interesting features, however this is clearly beyond the scope of the present study. Furthermore approximate forms of the solution may be found in various asymptotic regimes, providing us with intuition concerning the qualitative behavior of the controlled system however, this is again beyond the scope if the present study.

8.8 Linear quadratic approximation of nonlinear robust control problems

The approach in sections 2-4 which provided the most tractable results, was based on linear quadratic problems. However, by using the linear quadratic approximation of general nonlinear control problems (see, e.g., Magill (1977a)) we may adapt our results for the linear quadratic problem to obtain an approximation to a nonlinear robust optimal control problem. We sketch this approach.

Assume that we have a nonlinear problem, subject to weak additive noise, of the general form,

$$dx_n = f(\sum_m a_{nm}x_m, \sum_m b_{nm}u_m)dt + \sum_m c_{nm}dw_m. \tag{47}$$

The problem is subject to model uncertainty which may be modelled in terms of a drift $\{v\}$ so that applying Girsanov's theorem (see, e.g., Karatzas and Shreve (1991) for the finite dimensional theory or Carmona and Tehranchi (2006) for the case of infinite dimensions) in the same fashion as for the linear case we obtain the family of models

$$dx_n = \left(f(\sum_m a_{nm}x_m, \sum_m b_{nm}u_m) + \epsilon \sum_m c_{nm}v_m \right) dt + \epsilon^{1/2} \sum_m c_{nm}dw_m. \tag{48}$$

Assume now that the control $\{u\}$ has to be chosen so as to maximize a cost functional of the form

$$J(u,v) = \max_{\{u\}} \min_{\{v\}} \mathbb{E}_P[\int_0^\infty e^{-rt} \sum_n \big(U(x_n(t), u_n(t)) + \theta v_n^2(t)\big) \tag{49}$$

which is a robust control problem of maximizing the utility function for the system for the worst possible model.

Since the noise is assumed to be weak we may consider as a zeroth order approximation a deterministic optimal path which is uniform in space, i.e. a solution $\{x_n^0(t)\}$ such that $x_n^0(t) = x^0(t)$ for all $n \in \mathbb{Z}$. This is the solution of a deterministic optimal control problem, which corresponds to the minimization of $J(u,v)$ for the unperturbed (deterministic) state equation (48) with $\epsilon = 0$ and is supported by a uniform in space control $\{u^0\}$ and uncertainty drift $\{v^0\}$. Let us consider perturbations of $\{x,u,v\}$ around this reference solution, i.e. let us consider solutions of the above problem of the form

$$\{x,u,v\} = \{x^0,u^0,v^0\} + \epsilon\{x^1,u^1,v^1\}$$

where now $\{x,u,v\}$ are subject to uncertainty and are solutions of the stochastic state equation (48) with ϵ a small parameter. The perturbation is assumed to be spatially dependent.

We linearize the state equation around the state $\{x^0, u^0, v^0\}$ to obtain to first order in ϵ that

$$dx_n^1 = \left(A \sum_m a_{nm} x_m^1 + B \sum_m b_{nm} u_m^1 + \sum_m c_{nm} v_m^1 \right) dt + \sum_m c_{nm} dw_m \tag{50}$$

where

$$A := f_1(\sum_m a_{nm} x_m^0(t), \sum_m b_{nm} u_m^0(t)), \quad B := f_2(\sum_m a_{nm} x_m^0(t), \sum_m b_{nm} u_m^0(t))$$

where f_1, f_2 are the partial derivatives of f with respect to the first and the second variable respectively. In general A and B are functions of time but not of space. In the special cases where either $\{x^0, u^0\}$ are steady states or the operators generated by the matrices $\{a_{nm}\}$, $\{b_{nm}\}$ are such that $\sum_m a_{nm} = 0$, $\sum_m b_{nm} = 0$ (diffusive coupling) the functions A and B are constant.

We furthermore look at the local behavior of the cost functional $J(u, v)$ around the state $\{x^0, u^0\}$. To be more precise, we calculate $J(u^0 + \epsilon u^1, v^0 + \epsilon v^1)$ and Taylor expand in ϵ. The first order term in the expansion is effectively the Gâteaux derivative of the functional J calculated at $\{u^0, v^0\}$, and by the extremality properties of $\{x^0, u^0, v^0\}$ this vanishes. We are thus left with the second order terms in this expansion which are

$$J^1 = \int_0^\infty e^{-rt} \sum_n (U_{11}(x_n^1(t))^2 + U_{22}(u_n^1(t))^2 - |\theta|(v_n^1(t))^2 dt$$

where U_{ij}, $i, j = 1, 2$ are the second derivatives of the utility function U with respect to the first and second variable calculated at $(\sum_m a_{nm})x^0(t)$ and $(\sum_m b_{nm})u^0(t)$. Therefore, $U_{ij} = U_{ij}(t)$ are deterministic functions of time (but not of space). If the utility function U is separable then $U_{12} = 0$. We may assume this without loss of generality.

The above discussion shows (rather informally) that the problem of minimizing $J(u, v)$ subject to the stochastic state equation (48) may be approximated by the problem of minimizing $J^1(u^1, v^1)$ subject to the linear state equation (50). This is a linear quadratic control problem similar to the one studied here. If x^0, u^0, v^0 are time independent (steady states) then this linear quadratic control problem is of the exact form studied here and the results of this paper may be used in their exact form to study the approximation of the nonlinear problem. If x^0, u^0, v^0 are time dependent then the linear quadratic control problem is one with time varying deterministic coefficients, which is still manageable by a slight modification of the results of this paper.

Remark 25. The above linear quadratic approximation to a nonlinear problem may also have an alternative interpretation as follows: Consider any desired path $\{x^0(t)\}$ and study small deviations from that. Linearizing the state equation we obtain a linear system similar to (50). Then the problem is to pick the controls u so as to keep the perturbed problem (50) as close as possible to the desired target $\{x^0(t)\}$,for the worst case scenario in terms of a whole family of models (specified by $\{v\}$). If the distance from the target is given by a quadratic distance functional, and the model misspecification is given by an entropic measure then it is easy to see that the above mentioned "stabilization" problem is equivalent to a linear quadratic robust control problem of the form treated in this paper.

An interesting question, which is clearly beyond the scope of the present work, is the detailed study of nonlinear systems beyond the level of section 6.

References

Albeverio, S., K. Makarov, and A. Motovilov (2003). Graph subspaces and the spectral shift function. *Canad. J. Math 55*(3), 449–503.

Anderson, E., L. Hansen, and T. Sargent (2003). A quartet of semigroups for model specification, robustness, prices of risk, and model detection. *Journal of the European Economic Association 1*(1), 68–123.

Athanassoglou, S. and A. Xepapadeas (2011). Pollution control: when, and how, to be precautious. *Fondazione Eni Enrico Mattei Working Papers*, 569.

Aubin, J. and I. Ekeland (1984). *Applied nonlinear analysis.* Wiley New York.

Bamieh, B., F. Paganini, and M. Dahleh (2002). Distributed control of spatially invariant systems. *Automatic Control, IEEE Transactions on 47*(7), 1091–1107.

Bensoussan, A., G. Da Prato, M. Delfour, and S. Mitter (1992). *Representation and Control of Infinite Dimensional Systems: Vol 2 (Systems & Control: Foundations & Applications).* Birkhäuser Boston Inc., Boston, MA.

Brock, W., A. Xepapadeas, and A. Yannacopoulos. The emergence of optimal agglomeration in dynamic economics. *submitted*.

Brock, W., A. Xepapadeas, and A. Yannacopoulos (2014). Robust control and hot spots in spatiotemporal economic systems. *Dynamic Games and Applications, in press*.

Carmona, R. and M. Tehranchi (2006). *Interest rate models: an infinite dimensional stochastic analysis perspective*, Volume 13. Springer Verlag.

Cerrai, S. (2001). *Second order PDE's in finite and infinite dimension: a probabilistic approach.* Number 1762. Springer Verlag.

Curtain, R., O. Iftime, and H. Zwart (2008). System theoretic properties of platoon-type systems. In *Decision and Control, 2008. CDC 2008. 47th IEEE Conference on*, pp. 1442–1447. IEEE.

Da Prato, G. (2002). Linear quadratic control theory for infinite dimensional systems. *Mathematical Control Theory. ICTP Lecture Notes 8*, 59–105.

Da Prato, G. and J. Zabczyk (1996). *Ergodicity for infinite dimensional systems*, Volume 229. Cambridge Univ Pr.

Da Prato, G. and J. Zabczyk (2002). *Second order partial differential equations in Hilbert spaces*, Volume 293. Cambridge Univ Pr.

Dautray, R., J. Lions, M. Artola, and M. Cessenat (2000). *Mathematical analysis and numerical methods for science and technology: Spectral theory and applications*, Volume 5. Springer Verlag.

Hansen, L. and T. Sargent (2001). Robust control and model uncertainty. *American Economic Review*, 60–66.

Isaacs, R. (1999). *Differential games: a mathematical theory with applications to warfare and pursuit, control and optimization.* Dover Pubns.

Karachalios, N. (2006). A remark on the existence of breather solutions for the discrete nonlinear schrodinger equation in infinite lattices: the case of site-dependent anharmonic parameters. *Proceedings of the Edinburgh Mathematical Society 49*(1), 115–130.

Karachalios, N. and A. Yannacopoulos (2005). Global existence and compact attractors for the discrete nonlinear schrödinger equation. *Journal of Differential Equations 217*(1), 88–123.

Karachalios, N. and A. Yannacopoulos (2007). The existence of a global attractor for the discrete nonlinear schrodinger equation. ii. compactness without tail estimates in F^N, $n \geq 1$, lattices. *PROCEEDINGS-ROYAL SOCIETY OF EDINBURGH A 137*(1), 63.

Karatzas, I. and S. Shreve (1991). *Brownian motion and stochastic calculus.* Springer.

Ljungqvist, L. and T. Sargent (2004). *Recursive macroeconomic theory.* The MIT press.

Magill, M. (1977a). A local analysis of n-sector capital accumulation under uncertainty. *Journal of Economic Theory 15*(1), 211–219.

Magill, M. (1977b). Some new results on the local stability of the process of capital accumulation. *Journal of Economic Theory 15*(1), 174–210.

Prato, G. and J. Zabczyk (1995). Convergence to equilibrium for classical and quantum spin systems. *Probability theory and related fields 103*(4), 529–552.

Rudin, W. (1990). *Fourier analysis on groups*, Volume 12. Wiley-Interscience.