DISCRETE AND CONTINUOUS
DYNAMICAL SYSTEMS
Volume **23**, Number **1&2**, January & February **2009**

doi:10.3934/dcds.2009.23.521

pp. **521–540**

UPPER SEMI-CONTINUITY OF STATIONARY STATISTICAL PROPERTIES OF DISSIPATIVE SYSTEMS

Xiaoming Wang

Department of Mathematics
Florida State University, Tallahassee, FL32306, USA

Abstract. We show that stationary statistical properties for uniformly dissipative dynamical systems are upper semi-continuous under regular perturbation and a special type of singular perturbation in time of relaxation type. The results presented are applicable to many physical systems such as the singular limit of infinite Prandtl-Darcy number in the Darcy-Boussinesq system for convection in porous media, or the large Prandtl asymptotics for the Boussinesq system.

1. **Introduction.** Turbulence is ubiquitous in fluid phenomena. Many (generalized) dynamical systems arising in fluid applications are dissipative complex systems in the sense that they possess a compact global attractor and the dynamics are turbulent/chaotic [32, 20]. For instance the Navier-Stokes equations at large Reynolds number or Grashoff number, the Boussinesq system for convection at large Rayleigh number, the Darcy-Boussinesq system for convection in fluid saturated porous media at large Darcy-Rayleigh number, the Lorenz 63 and 96 model are all such systems. The complex behavior is not necessarily related to the well-posedness of the system. For instance, the simple discrete dynamical system on the unit interval induced by the tent map $T(x) = 2x, x \in [0, \frac{1}{2}]; \quad T(x) = 2 - 2x, x \in [\frac{1}{2}, 1]$, is well-posed but has chaotic behavior with generic sensitive dependence on initial data. The generic sensitive dependence on initial data renders it impossible to make precise long time prediction (loss of predictability). It has been long understood that statistical properties of this kind of complex systems are much more relevant and coherent than single trajectories[25, 14, 18, 22]. If the system reaches some kind of statistical stationary state, the objects that characterize the stationary statistical properties are the invariant measures or stationary statistical solutions of the system.

Consider now an abstract continuous dynamical system $\{S(t), t \geq 0\}$ on a phase space H where H is a separable Hilbert space in general, but could be a metric space as well. Uncertainty in initial data characterized by an initial probability measure μ_0 on the phase space will propagate to future time t characterized by (time dependent/non-stationary) statistical solution μ_t which is defined through (strong formulation) pull-back as

$$\mu_t(E) = \mu_0(S^{-1}(t)(E)), \forall t \geq 0$$

2000 *Mathematics Subject Classification.* Primary: 37L40, 37L15; Secondary: 35Q35, 76D06, 76M45.

Key words and phrases. Stationary statistical solution, invariant measure, dissipative system, upper semi-continuity, Rayleigh-Bénard convection.

This work is supported in part by NSF grant DMS0606671.

for all Borel measurable sets E, or through (weak formulation) push-forward as

$$\int_H \varphi(\mathbf{u})\, d\mu_t(\mathbf{u}) = \int_H \varphi(S(t)\mathbf{u})\, d\mu_0(\mathbf{u})$$

for all suitable test functionals φ.

Now if the system reaches a statistical equilibrium in the sense that the statistics are time independent (stationary), the probability measure μ that describes the stationary uncertainty can be characterized via either the strong (pull-back) or weak (push-forward) formulation. In the *pull-back* case, we have the notion of **Invariant Measure** or *stationary statistical solution*

$$\mu(E) = \mu(S^{-1}(t)(E)),\ \forall t \geq 0 \tag{1}$$

and in the *push-forward* case we have the weak invariance

$$\int_H \varphi(\mathbf{u})\, d\mu(\mathbf{u}) = \int_H \varphi(S(t)\mathbf{u})\, d\mu(\mathbf{u}) \tag{2}$$

for all suitable test functionals φ.

Once we have a physically relevant invariant measure (stationary statistical solution), various statistical quantities can be computed/approximated by evaluating the left-hand-side in the push-forward formulation. Hence all stationary statistical properties are encoded in the invariant measures.

Besides uncertainty in initial data, we also have uncertainty in many physical parameters. Therefore we also need to consider the influence of these uncertainties in physical parameters on the statistical properties. This is the question of dependence on parameters of the stationary statistical properties which is the *main concern* of this manuscript.

Few works on dependence of stationary statistical properties on parameters exist. See [8] for the case of vanishing viscosity limit of the stationary statistical properties for the damped driven two dimensional Navier-Stokes system, and [40] for the case of the singular limit of infinite Prandtl number in the Rayleigh-Bénard system. The results presented here are in a general setting that are relatively easily applicable to other systems. In particular, our results cover the case of partially/weakly dissipative systems such as the weakly damped-driven nonlinear Schrödinger equation [32, 36], the damped and driven KdV equation [32], fluid of second grade [24], Darcy-Boussinesq system for convection in fluid saturated porous media [13, 27] etc.

The rest of the manuscript is organized as follows: In section 2 we present a few general results on invariant measures for dissipative dynamical systems. The results presented in this section are mostly well-known in literatures [1, 14, 35] albeit under different and/or additional assumptions. We provide the details here since the versions that we need for some of our applications, say convection in fluid saturated porous media, are not available in the literatures to the best of our knowledge. In section 3 we discuss perturbation of stationary statistical properties of uniformly dissipative systems. We show that the stationary statistical properties are in fact upper semi-continuous in the case of regular perturbation and a special type of singular perturbation of two time scale of relaxation type. The results presented in this section are new to the best of our knowledge, and can be applied to many physical situations such as Rayleigh-Bénard convection at large Prandtl number

(an example of strongly dissipative generalized dynamical system), and the Darcy-Boussinesq system for convection in fluid saturated porous media at large Darcy-Prandtl number (an example of weakly/partial dissipative dynamical system). In the last section we offer comments on applications and extensions.

2. **Some general results on stationary statistical properties for dissipative systems.** Here we present a few general results on stationary statistical properties for dissipative dynamical systems relating invariant measures and the global attractors, time and spatial averages, ergodicity and extremal points of the set of invariant measures. Very similar results are well-known in the literatures under various different or additional assumptions although the current forms that are applicable to the case of weakly/partially dissipative systems are new to the best of our knowledge.

Recall a dynamical system is called **dissipative** if it possesses a **global attractor** in the phase space that is compact, invariant and it attracts all bounded sets [32, 15, 28].

It is easy to see that if the dynamical system possesses a steady state solution, the delta measure concentrated at the steady state is a (singular) invariant measure of the system. Invariant measure is in general not unique due to possible existence of multiple steady states and/or periodic orbits etc. Even different generalized limits and different trajectories in the second theorem (2) may induce different invariant measures. For dissipative systems we naturally anticipate that all invariant measures are supported on the global attractor since the global attractor attracts all bounded sets. Since the global attractor is compact, we also anticipate some kind of compactness of the set of all invariant measures which we denote $\mathcal{IM}$. Indeed, we have

Theorem 1. *[$\mathcal{IM}$ and the global attractors] The support of any invariant measure μ of a given continuous dissipative dynamical system is included in the global attractor. Moreover, the set of all invariant measures, $\mathcal{IM}$, is a convex compact set (with respect to the weak topology) in the space of Borel measures on the phase space.*

Proof. Since the invariant measure μ is a probability measure, for any $\varepsilon > 0$, there exists $R > 0$ such that

$$\mu(B_R) \geq 1 - \varepsilon$$

where B_R is a ball of radius R.

Since $S^{-1}(t)S(t)B_R \supset B_R, \forall t \geq 0$, we have together with the invariance of μ,

$$\mu(S^{-1}(t)S(t)B_R) = \mu(S(t)B_R) \geq \mu(B_R) \geq 1 - \varepsilon$$

Since $\mathcal{A}$ is the global attractor and hence it attracts B_R, we see that for any $\delta > 0$, there exists a $T_\delta > 0$ such that

$$dist_H(S(t)B_R, \mathcal{A}) < \delta, \text{ or } S(t)B_R \subset \mathcal{A}_\delta$$

i.e., $S(t)B_R$ is within the open delta neighborhood of $\mathcal{A}$, denoted $\mathcal{A}_\delta$.

On the other hand, μ is regular (since it is a Borel measure [29]) and hence for the given $\varepsilon > 0$, there exists a $\delta_0 > 0$ such that

$$\mu(\mathcal{A}) \geq \mu(\mathcal{A}_{\delta_0}) - \varepsilon.$$

Combining the above inequalities we have

$$\begin{aligned} \mu(\mathcal{A}) &\geq \mu(\mathcal{A}_{\delta_0}) - \varepsilon \\ &\geq \mu(S(t)B_R) - \varepsilon, \quad \forall t \geq T_{\delta_0} \\ &\geq 1 - 2\varepsilon. \end{aligned}$$

This finishes the proof that any invariant measure is supported on the global attractor.

It is easy to see that the set of all invariant measures, $\mathcal{IM}$, is a convex and closed (under the weak convergence topology) set utilizing the push-forward (weak) formulation for instance. The compactness follows from Prokharov's tightness theorem, the compactness of the global attractor, and the fact that all invariant measures are supported on the global attractor.

This ends the proof of the theorem. □

A proof of the first part of this result which relies on the existence of a compact absorbing set is essentially included in [14]. However, not all dissipative systems possess compact absorbing set [32, 15].

For physically interesting complex systems, steady state may not exist even for dissipative system. Moreover, even if steady states exist, they may not be physically relevant for typical statistical behavior. For instance, it is hard to imagine any steady state is physically relevant at large time for turbulent flow. On the other hand, practitioners have been using long time average

$$\frac{1}{T}\int_0^T \varphi(S(t)\mathbf{u}_0)\,dt, \quad T \gg 1$$

to compute/approximate stationary statistics based on the ergodicity idea that generic trajectory will traverse almost all parts of the phase space and hence temporal and spatial averages are equivalent. The most well-know result is probably **Birkhoff's ergodicity theorem** [18, 35, 30, 42] which dictates that for any given invariant measure μ, limit of long time average exists almost surely with respect to μ. The limit of the long time average equals to spatial average if μ is an **ergodic invariant measure** in the sense that the only invariant sets are trivial (having probability 1 or 0). Such an approach is very appealing since only one trajectory/simulation/experiment/observation is needed. This is particularly useful if the dynamical system is not known explicitly and we only have historical data (one trajectory observation). However, we usually do not have an invariant measure to start with in applications. This difficulty may be circumvented via long time average.

We recall that the long time average may not have limit. This difficulty can be overcome by considering generalized (Banach) limit [19]. This together with the dissipative nature of the system and the Kakutani-Riesz representation theorem [19, 14] enables us to establish the following equivalence of spatial and temporal averages.

Theorem 2. *[$\mathcal{IM}$ generated by time averages] Let $\{S(t), t \geq 0\}$ be a continuous dissipative dynamical system on a reflexive Banach space H. Then for any given initial data $\mathbf{u}_0$ and any choice of generalized limit LIM, there exists a unique invariant measure μ of the dynamical system so that the time average defined via LIM is equivalent to spatial average with respect to μ, i.e.*

$$LIM_{T\to\infty}\frac{1}{T}\int_0^T \varphi(S(t)\mathbf{u}_0)\,dt = \int_H \varphi(\mathbf{u})\,d\mu(\mathbf{u}), \forall \varphi \in C(H).$$

Proof. The first half of the proof, i.e., the proof of existence of an invariant measure so that the spatial and temporal averages are the same for *weakly* continuous test

functionals is classical [14]. We reproduce the proof here for the sake of completeness.

Since the system possesses a compact global attractor $\mathcal{A}$, there exists a closed absorbing ball $\mathcal{B}_a$ in H for each fixed initial data $\mathbf{u}_0$ so that

$$S(t)\mathbf{u}_0 \in \mathcal{B}_a, \forall t \geq 0.$$

$\mathcal{B}_a$ is weakly compact since H is reflexive thanks to Banach-Alaoglu theorem [19] (generalized Heine-Borel theorem). Therefore, thanks to the Kakutani-Riesz representation theorem [19], for a fixed initial data $\mathbf{u}_0$ and generalized limit LIM, there exists a Borel probability measure μ on $\mathcal{B}_a$ such that

$$\mathrm{LIM}_{T\to\infty}\frac{1}{T}\int_0^T \varphi(S(t)\mathbf{u}_0)\,dt = \int_{\mathcal{B}_a} \varphi(\mathbf{u})\,d\mu(\mathbf{u}), \forall \varphi \in C_w(\mathcal{B}_a)$$

since the long time average defined through the generalized limit on the left hand side defines a continuous linear functional on $C_w(\mathcal{B}_a)$, the space of all weakly continuous functionals on $\mathcal{B}_a$.

Thanks to the Tietze extension theorem [19] and the fact that $\mathcal{B}_a$ is weakly closed, any weakly continuous functionals on $\mathcal{B}_a$ can be extended to a weakly continuous functional on H. The Borel probability measure μ on $\mathcal{B}_a$ can be extended to a Borel probability measure on the whole space in a trivial manner (assigning zero probability to $H\backslash\mathcal{B}_a$) and thus the same equivalence of spatial and temporal average relation holds for the extended functional. Moreover, we see that if two weakly continuous functionals agree on $\mathcal{B}_a$, then the long time averages defined through the generalized limit are the same since the whole trajectory belongs to $\mathcal{B}_a$. Also, the restriction of any weakly continuous functional on H onto $\mathcal{B}_a$ is weakly continuous. Therefore, we have the desired equivalence of the spatial and temporal averages for all $\varphi \in C_w(H)$, i.e.,

$$\mathrm{LIM}_{T\to\infty}\frac{1}{T}\int_0^T \varphi(S(t)\mathbf{u}_0)\,dt = \int_H \varphi(\mathbf{u})\,d\mu(\mathbf{u}), \forall \varphi \in C_w(H).$$

We need to show that μ is invariant under the flow and that the invariance is in fact valid for all continuous functionals on H (weakly continuous functionals are automatically continuous functionals but not vice versa).

We now verify that μ is invariant under the flow. For this purpose we fix $\tau > 0$ and we have

$$\begin{aligned}
&\int_H \varphi(S(\tau)\mathbf{u})\,d\mu(\mathbf{u}) \\
=\ & \mathrm{LIM}_{T\to\infty}\frac{1}{T}\int_0^T \varphi(S(\tau)S(t)\mathbf{u}_0)\,dt \\
=\ & \mathrm{LIM}_{T\to\infty}\frac{1}{T}\int_\tau^{T+\tau} \varphi(S(t)\mathbf{u}_0)\,dt \\
=\ & \mathrm{LIM}_{T\to\infty}\frac{1}{T}\{\int_0^T + \int_T^{T+\tau} - \int_0^\tau\}\varphi(S(t)\mathbf{u}_0)\,dt \\
=\ & \mathrm{LIM}_{T\to\infty}\frac{1}{T}\int_0^T \varphi(S(t)\mathbf{u}_0)\,dt + \mathrm{LIM}_{T\to\infty}\frac{1}{T}\{\int_T^{T+\tau} - \int_0^\tau\}\varphi(S(t)\mathbf{u}_0)\,dt \\
=\ & \int_H \varphi(\mathbf{u})\,d\mu(\mathbf{u})
\end{aligned}$$

where we have used the boundedness of the weakly continuous functional φ on the weakly compact set $\mathcal{B}_a$ which contains the whole trajectory. This ends the proof of the first half of the theorem.

Next, we need to show that the equivalence between spatial and temporal averages are in fact valid for any continuous functionals on H.

Since the weak Borel sets and strong Borel sets are the same, μ is also strongly invariant and hence its support contained in the global attractor $\mathcal{A}$ by the previous theorem.

Now let $\varphi \in C(H)$ (continuous but not necessarily weakly continuous), the restriction of φ on $\mathcal{A}$ is weakly continuous due to the compactness of $\mathcal{A}$. We also notice that $\mathcal{A}$ is weakly closed since any weak limit would also be a strong limit thanks to the compactness. Now let $\tilde{\varphi}$ be a weakly continuous extension of $\varphi|_{\mathcal{A}}$ to H. Such an extension exists due to Tietze theorem. We also notice that the two functionals must be asymptotically the same along the trajectory in the sense that

$$\varphi(S(t)\mathbf{u}_0) - \tilde{\varphi}(S(t)\mathbf{u}_0) \to 0, t \to \infty.$$

Indeed, if it were not true, we must have a $\delta > 0$ and a time sequence $\{t_n, n = 1, 2, \cdots\}$ with $t_n \to \infty$, as $n \to \infty$ so that

$$|\varphi(S(t_n)\mathbf{u}_0) - \tilde{\varphi}(S(t_n)\mathbf{u}_0)| \geq \delta.$$

Due to the attracting nature of the global attractor $\mathcal{A}$, there exists a sub time sequence, still denoted $\{t_n, n = 1, 2, \cdots\}$, and $\mathbf{u}_\infty \in \mathcal{A}$, so that

$$S(t_n)\mathbf{u}_0 \to \mathbf{u}_\infty \in \mathcal{A}, n \to \infty.$$

Therefore, thanks to the continuity of φ and weak continuity of $\tilde{\varphi}$, and the fact that $\varphi(\mathbf{u}_\infty)$ and $\tilde{\varphi}(\mathbf{u}_\infty)$ are the same since $\mathbf{u}_\infty \in \mathcal{A}$, we have

$$\lim_{n\to\infty} \varphi(S(t_n)\mathbf{u}_0) = \varphi(\mathbf{u}_\infty) = \tilde{\varphi}(\mathbf{u}_\infty) = \lim_{n\to\infty} \tilde{\varphi}(S(t_n)\mathbf{u}_0)$$

which contradicts the choice of the time sequence.

With the asymptotic equivalence between $\varphi(S(t)\mathbf{u}_0)$ and $\tilde{\varphi}(S(t)\mathbf{u}_0)$, we have, for any $\varphi \in C(H)$,

$$\begin{aligned}
& \int_H \varphi(\mathbf{u})\, d\mu(\mathbf{u}) \\
= & \int_{\mathcal{A}} \varphi(\mathbf{u})\, d\mu(\mathbf{u}) \quad \text{(support of } \mu) \\
= & \int_{\mathcal{A}} \tilde{\varphi}(\mathbf{u})\, d\mu(\mathbf{u}) \quad \text{(equivalence of } \varphi \text{ and } \tilde{\varphi} \text{ on } \mathcal{A}) \\
= & \int_H \tilde{\varphi}(\mathbf{u})\, d\mu(\mathbf{u}) \quad \text{(support of } \mu) \\
= & \ \mathrm{LIM}_{T\to\infty} \frac{1}{T} \int_0^T \tilde{\varphi}(S(t)\mathbf{u}_0)\, dt \quad (\tilde{\varphi} \in C_w(H) \text{ and the weak equivalence}) \\
= & \ \mathrm{LIM}_{T\to\infty} \frac{1}{T} \int_0^T \varphi(S(t)\mathbf{u}_0)\, dt. \quad \text{(asymptotic equivalence of } \varphi \text{ and } \tilde{\varphi})
\end{aligned}$$

This ends the proof of the theorem. □

A proof that relies on the existence of a compact absorbing set, or a weaker result that the equivalence is valid for only weakly continuous functionals on H, i.e. $C_w(H)$ is essentially included in [14, 34].

We recall from Birkhoff's ergodic theorem that spatial and temporal averages are equivalent if the underlying invariant measure is ergodic. Not all invariant measures are ergodic of course. However, ergodic measures are the building blocks of the set of all invariant measures of a dissipative system. Indeed, since $\mathcal{IM}$ is convex and compact, it is the closed convex hull of the extremal points of this set (Krein-Milman theorem [19, 14]). This together with the following result which states that extremal invariant measures are ergodic completes our argument.

Theorem 3. *[Ergodicity and extremal points] Let $\mathcal{IM}$ be the set of all invariant probability measures of a dissipative dynamical system $\{S(t), t \geq 0\}$. Then an invariant measure μ is ergodic if μ is an extreme point of $\mathcal{IM}$. Moreover, if the dynamical system is injective on the global attractor $\mathcal{A}$, then every ergodic invariant measure must be an extremal point of $\mathcal{IM}$.*

Proof. The proof of the sufficiency is classical [1, 18, 35]. We reproduce here for the sake of completeness.

Assume that μ is an extremal point of the set of invariant measures of the dynamical system.

Suppose that μ is not ergodic. Then there must exist an invariant set E_0 such that $0 < \mu(E_0) < 1$. Define two measures on the phase space as

$$\begin{aligned} \mu_1(E) &= \frac{\mu(E \bigcap E_0)}{\mu(E_0)}, \\ \mu_2(E) &= \frac{\mu(E \bigcap (H \setminus E_0))}{\mu(H \setminus E_0)}. \end{aligned}$$

It is then easy to see that $\mu_1, \mu_2 \in \mathcal{IM}$. Indeed,

$$\mu_1(S^{-1}(t)E) = \frac{\mu((S^{-1}(t)E) \bigcap E_0)}{\mu(E_0)} = \frac{\mu(S^{-1}(t)(E \bigcap E_0))}{\mu(E_0)} = \frac{\mu(E \bigcap E_0)}{\mu(E_0)} = \mu_1(E).$$

On the other hand

$$\mu = \mu(E_0)\mu_1 + (1 - \mu(E_0))\mu_2, \mu(E_0) \in (0, 1)$$

and

$$\mu \neq \mu_1.$$

This contradicts the assumption that μ is an extremal point of $\mathcal{IM}$.

Next we prove the necessity using by way of contradiction.

Recall that the support of all invariant measures are included in the global attractor. Since the dynamical system is injective and onto on the global attractor, and that the global attractor is compact by definition, $S(t)$ is continuously invertible on $\mathcal{A}$.

Suppose that μ is ergodic and μ is not an extremal point of $\mathcal{IM}$. Then there exists $\mu_1, \mu_2 \in \mathcal{IM}$ and $\lambda \in (0, 1)$ such that

$$\mu = \lambda\mu_1 + (1 - \lambda)\mu_2.$$

Let

$$f(\mathbf{u}) = \frac{d\mu_1}{d\mu} \in L^1(H, \mu)$$

be the Radon-Nikodym derivative of μ_1 with respect to μ. Then for any continuous test functional $\varphi(\mathbf{u})$ we have

$$\begin{aligned}\int_H \varphi(S(t)\mathbf{u}) f(S(t)\mathbf{u})\, d\mu(\mathbf{u}) &= \int_H \varphi(\mathbf{u}) f(\mathbf{u})\, d\mu(\mathbf{u}) \quad \text{invariance}\\ &= \int_H \varphi(\mathbf{u})\, d\mu_1(\mathbf{u}) \quad \text{definition}\\ &= \int_H \varphi(S(t)\mathbf{u})\, d\mu_1(\mathbf{u}) \quad \text{invariance}\\ &= \int_H \varphi(S(t)\mathbf{u}) f(\mathbf{u})\, d\mu(\mathbf{u}) \quad \text{definition}\end{aligned}$$

In particular, for the special choice of test functional $\varphi(\mathbf{u}) = \psi(S^{-1}(t)\mathbf{u})$ (this is allowed by the fact that we can restrict to the global attractor and $S^{-1}(t)$ exists and is continuous on $\mathcal{A}$), we have

$$f(S(t)\mathbf{u}) = f(\mathbf{u}), \quad a.s..$$

Hence f must be a constant by Koopman's theorem [18, 42]. The constant must be one since both μ_1 and μ are probability measures. This is a contradiction.

This ends the proof of the necessity and hence we end the proof of the theorem. □

Proof of this result without the assumption on dissipativity, but restricted to discrete dynamical system or when the solution semi-group is in fact a group can be found in [1, 35].

3. Perturbation of stationary statistical properties.

For a given dynamical system with parameters (such as the Reynolds number, Grashoff number, Rayleigh number, Prandtl number ...), there is another level of uncertainty, the parameters, since they are all estimated. Therefore we are interested in the question of how stationary statistical properties (invariant measures) depend on various parameters. This may be viewed as a stability issue.

3.1. Regular perturbation.

For an abstract dynamical system with parameters, there is no obvious reason that the statistical properties should depend in some nice way on the parameters, even if trajectories converge on any given finite time interval (see [42] for a counter-example). However, the situation is relatively simpler if we consider uniformly dissipative systems only. What we can show, for uniformly dissipative system, is that the set of invariant measures is upper semi-continuous with respect to regular perturbation of parameters. The general argument is that if we have enough a priori estimates for a given parameter α close to α_0, we have tightness/weak compactness of the set of invariant measures for all interested parameter region thanks to Prokhorov's theorem [2, 14]. The limit of invariant measures of the perturbed system must be an invariant measure of the limit system by taking the limit in the Liouville type equation or the weak invariance formulation.

More precisely, we have the following theorem:

Theorem 4. *[Upper semi-continuity of $\mathcal{IM}$, regular version] Suppose the family of continuous dynamical systems $\{S(t,\epsilon), t \geq 0\}$ on a separable Hilbert space H is uniformly dissipative in the sense that $K = \bigcup_{0<|\epsilon|\leq\epsilon_0} \mathcal{A}_\epsilon$ is pre-compact where $\mathcal{A}_\epsilon$ denotes the global attractor for the system with parameter ϵ. Moreover, we assume*

that the trajectories converge on any finite time interval uniformly on the attractors, i.e.,

$$\lim_{\epsilon\to 0}\sup_{\mathbf{u}\in\mathcal{A}_\varepsilon}\|S(t,\epsilon)\mathbf{u}-S(t,0)\mathbf{u}\|_H=0,\forall t\geq 0.$$

Then the set of invariant measures are upper semi-continuous in the sense that for any $\{\mu_\epsilon\in\mathcal{IM}_\epsilon, 0<|\epsilon|\leq\epsilon_0\}$*, there exists* $\mu_0\in\mathcal{IM}_0$ *and a subsequence (still denoted* μ_ϵ*) such that*

$$\lim_{\epsilon\to 0}\mu_\epsilon=\mu_0$$

(the convergence is in the weak sense).

Proof. Since the support of any invariant measure is contained in the global attractor $\mathcal{A}_\epsilon$ which is contained in the same pre-compact set K, we see that $\{\mu_\epsilon, 0<|\epsilon|\leq\epsilon_0\}$ is tight in $\mathcal{P}M(H)$, the space of Borel probability measures on H, thanks to Prokhorov's theorem. Without loss of generality we assume that μ_ϵ weakly converges to μ_0.

Our goal now is to show that $\mu_0\in\mathcal{IM}_0$.

Now for any $t>0$, since $\mu_\epsilon\in\mathcal{IM}_\epsilon$, for any continuous test functional φ we have

$$\int_H\varphi(S(t,\epsilon)\mathbf{u})\,d\mu_\epsilon(\mathbf{u})=\int_H\varphi(\mathbf{u})\,d\mu_\epsilon(\mathbf{u}).$$

Thanks to the weak convergence we also have

$$\begin{aligned}\lim_{\epsilon\to 0}\int_H\varphi(\mathbf{u})\,d\mu_\epsilon(\mathbf{u}) &= \int_H\varphi(\mathbf{u})\,d\mu_0(\mathbf{u}),\\ \lim_{\epsilon\to 0}\int_H\varphi(S(t,0)\mathbf{u})\,d\mu_\epsilon(\mathbf{u}) &= \int_H\varphi(S(t,0)\mathbf{u})\,d\mu_0(\mathbf{u}).\end{aligned}$$

Hence, for any smooth cylindrical test functional φ,

$$\begin{aligned}&|\int_H\varphi(S(t,0))\mathbf{u})\,d\mu_0(\mathbf{u})-\int_H\varphi(\mathbf{u})\,d\mu_0(\mathbf{u})|\\ =\;&\lim_{\varepsilon\to 0}|\int_H\varphi(S(t,0))\mathbf{u})\,d\mu_\epsilon(\mathbf{u})-\int_H\varphi(\mathbf{u})\,d\mu_\epsilon(\mathbf{u})|\\ \leq\;&\lim_{\varepsilon\to 0}|\int_H(\varphi(S(t,\epsilon)\mathbf{u})-\varphi(\mathbf{u}))\,d\mu_\epsilon(\mathbf{u})|\\ &+\lim_{\varepsilon\to 0}|\int_H(\varphi(S(t,\epsilon)\mathbf{u})-\varphi(S(t,0)\mathbf{u}))\,d\mu_\epsilon(\mathbf{u})|\\ \leq\;&\lim_{\varepsilon\to 0}\int_H\sup\|\varphi'(\mathbf{u})\|\|S(t,\epsilon)\mathbf{u}-S(t,0)\mathbf{u}\|\,d\mu_\epsilon(\mathbf{u})\\ =\;&0\end{aligned}$$

where we have used the push-forward invariance of μ_ε under $S(t,\varepsilon)$, mean value theorem, and the uniform convergence of trajectories starting on $\mathcal{A}_\varepsilon$.

Now for an arbitrary continuous test functional $\varphi(\mathbf{u})$, it can be approximated by cylindrical continuous test functional $\varphi(P_m\mathbf{u})$ where P_m is the orthogonal projection onto the subspace spanned by the first m elements of a given orthonormal basis of H. And we have

$$\int_H\varphi(\mathbf{u})\,d\mu_0(\mathbf{u})=\lim_{m\to\infty}\int_H\varphi(P_m\mathbf{u})\,d\mu_0(\mathbf{u})$$

by the Lebesgue dominated convergence theorem and the fact that μ_0 is supported on the compact global attractor $\mathcal{A}$ and hence φ is bounded on $\mathcal{A}$. The finite dimensional cylindrical test functionals $\varphi(P_m\mathbf{u})$ can be further approximated by smooth cylindrical test functionals utilizing standard finite dimensional smoothing (mollifier) techniques, i.e., there exists smooth cylindrical test functionals $\varphi_{m,k}$ such that

$$\begin{aligned} \varphi_{m,k}(\mathbf{u}) &\to \varphi(P_m\mathbf{u}), k\to\infty, \\ ||\varphi_{m,k}||_{L^\infty} &\leq \|\varphi\circ P_m\|_{L^\infty}. \end{aligned}$$

Therefore,

$$\int_H \varphi(P_m\mathbf{u})\,d\mu_0(\mathbf{u}) = \lim_{k\to\infty}\int_H \varphi_{m,k}(\mathbf{u})\,d\mu_0(\mathbf{u}).$$

Combining the above we end the proof of the theorem. □

A corollary of this upper semi-continuity is that the extremal statistics (defined though long time averages) are upper semi-continuous in the following sense.

Theorem 5. *[Upper semi-continuity of extremal time averaged statistics] Under the same assumption as in the previous theorem, i.e., uniform dissipativity plus finite time convergence, we have for any fixed continuous test functional φ_0, the extremal statistics are saturated by ergodic invariant measures, i.e., there exist ergodic invariant measures $\nu_\epsilon\in\mathcal{IM}_\epsilon, \nu_0\in\mathcal{IM}_0$ such that*

$$\begin{aligned} \sup_{\mathbf{u}\in H}\limsup_{T\to\infty}\frac{1}{T}\int_0^T \varphi_0(S(t,\epsilon)\mathbf{u})\,dt &= \int_H \varphi_0(\mathbf{u})\,d\nu_\epsilon(\mathbf{u}), \\ \sup_{\mathbf{u}\in H}\limsup_{T\to\infty}\frac{1}{T}\int_0^T \varphi_0(S(t,0)\mathbf{u})\,dt &= \int_H \varphi_0(\mathbf{u})\,d\nu_0(\mathbf{u}). \end{aligned}$$

Moreover, the extremal statistics are upper semi-continuous in the parameter ε, i.e.,

$$\limsup_{\epsilon\to 0}\sup_{\mathbf{u}\in H}\limsup_{T\to\infty}\frac{1}{T}\int_0^T \varphi_0(S(t,\epsilon)\mathbf{u})\,dt \leq \sup_{\mathbf{u}\in H}\limsup_{T\to\infty}\frac{1}{T}\int_0^T \varphi_0(S(t,0)\mathbf{u})\,dt.$$

Proof. Recall that for a fixed initial data $\mathbf{u}_0$ and a fixed continuous test functional φ_0, there exists a special Banach/generalized limit that agrees with the lim sup for long time average on φ_0. This implies, thanks to the equivalence between spatial and temporal averages (theorem 2), there exist $\mu_{\epsilon,\mathbf{u}_0}\in\mathcal{IM}_\epsilon$ and $\mu_{0,\mathbf{u}_0}\in\mathcal{IM}_0$ such that

$$\begin{aligned} \limsup_{T\to\infty}\frac{1}{T}\int_0^T \varphi_0(S(t,\epsilon)\mathbf{u}_0)\,dt &= \int_H \varphi_0(\mathbf{u})\,d\mu_{\epsilon,\mathbf{u}_0}, \\ \limsup_{T\to\infty}\frac{1}{T}\int_0^T \varphi_0(S(t,0)\mathbf{u}_0)\,dt &= \int_H \varphi_0(\mathbf{u})\,d\mu_{0,\mathbf{u}_0}. \end{aligned}$$

Now for fixed ϵ, the set $\mathcal{IM}_\epsilon$ is tight (theorem 1) and hence the set $\{\mu_{\epsilon,\mathbf{u}_0}, \mathbf{u}_0\in H\}$ must contain a subsequence that converges to some μ_ϵ so that

$$\begin{aligned} \sup_{\mathbf{u}_0\in H}\int_H \varphi_0(\mathbf{u})\,d\mu_{\epsilon,\mathbf{u}_0} &= \int_H \varphi_0(\mathbf{u})\,d\mu_\epsilon, \\ \sup_{\mathbf{u}_0\in H}\int_H \varphi_0(\mathbf{u})\,d\mu_{0,\mathbf{u}_0} &= \int_H \varphi_0(\mathbf{u})\,d\mu_0. \end{aligned}$$

It is easy to see that $\mu_\epsilon \in \mathcal{IM}_\epsilon, \mu_0 \in \mathcal{IM}_0$ since the sets of invariant measures are closed.

Next, we define

$$\begin{aligned} EIM_0 &= \{\tilde{\mu}_0 \in \mathcal{IM}_0, \sup_{\mu \in \mathcal{IM}_0} \int_H \varphi_0(\mathbf{u})\, d\mu = \int_H \varphi_0(\mathbf{u})\, d\tilde{\mu}_0\}, \\ EIM_\varepsilon &= \{\tilde{\mu}_\varepsilon \in \mathcal{IM}_\varepsilon, \sup_{\mu \in \mathcal{IM}_\varepsilon} \int_H \varphi_0(\mathbf{u})\, d\mu = \int_H \varphi_0(\mathbf{u})\, d\tilde{\mu}_\varepsilon\}. \end{aligned}$$

It is easy to see that both EIM_0 and EIM_ε are non-empty compact sets by the compactness of $\mathcal{IM}_0$ and $\mathcal{IM}\varepsilon$ (theorem 1). The set of extremals is non-empty by the Krein-Milman theorem and our assumption of uniform dissipativity. Let ν_0 be an extremal point of EIM_0 and ν_ε be an extremal point of EIM_ε

It is easy to see that ν_0 must be an extremal point of $\mathcal{IM}_0$ and ν_ϵ is an extremal point of $\mathcal{IM}_\epsilon$. Indeed, if for some $\lambda \in (0,1)$ and $\mu_1, \mu_2 \in \mathcal{IM}_0$ we have

$$\nu_0 = \lambda\mu_1 + (1-\lambda)\mu_2$$

then

$$\int_H \varphi_0(\mathbf{u})\, d\nu_0 = \lambda \int_H \varphi_0(\mathbf{u})\, d\mu_1 + (1-\lambda) \int_H \varphi_0(\mathbf{u})\, d\mu_2.$$

Since $\int_H \varphi_0(\mathbf{u})\, d\nu_0 = \sup_{\mu \in \mathcal{IM}_0} \int_H \varphi_0(\mathbf{u})\, d\mu$, we see that

$$\int_H \varphi_0(\mathbf{u})\, d\mu_1 = \int_H \varphi_0(\mathbf{u})\, d\mu_2 = \sup_{\mu \in IM_\epsilon} \int_H \varphi_0(\mathbf{u})\, d\mu$$

and therefore, both μ_1 and μ_2 are elements of EIM_0 which contradicts the assumption that ν_0 is an extremal of EIM_0. The same argument works for ν_ε.

Thanks to the previous theorem, the $\limsup_{\epsilon \to 0}$ on the right hand side of the last inequality in the theorem is attained and the limit is satisfied by an an element $\tilde{\nu}_0$ of $\mathcal{IM}_0$. Hence

$$\begin{aligned} & \limsup_{\epsilon \to 0} \sup_{\mathbf{u}_0 \in H} \limsup_{T \to \infty} \frac{1}{T} \int_0^T \varphi_0(S(t,\epsilon)\mathbf{u}_0)\, dt \\ = & \limsup_{\epsilon \to 0} \sup_{\mathbf{u}_0 \in H} \int_H \varphi_0(\mathbf{u})\, d\mu_{\epsilon,\mathbf{u}_0} \\ = & \limsup_{\epsilon \to 0} \int_H \varphi_0(\mathbf{u})\, d\mu_\epsilon \\ \leq & \limsup_{\epsilon \to 0} \int_H \varphi_0(\mathbf{u})\, d\nu_\epsilon \\ = & \int_H \varphi_0(\mathbf{u})\, d\tilde{\nu}_0 \\ \leq & \int_H \varphi_0(\mathbf{u})\, d\nu_0 \\ = & \lim_{T \to \infty} \frac{1}{T} \int_0^T \varphi_0(S(t,0)\mathbf{u})\, dt \quad (a.s.\ w.r.t.\ \nu_0) \\ \leq & \sup_{\mathbf{u}_0 \in H} \limsup_{T \to \infty} \frac{1}{T} \int_0^T \varphi_0(S(t,0)\mathbf{u}_0)\, dt \end{aligned}$$

where in the second to the last step we have used the fact that extremals of $\mathcal{IM}_0$ are necessarily ergodic (theorem 3), and hence spatial and temporal averages are equivalent.

This ends the proof of the theorem. □

3.2. **Singular perturbation.** The case of singular perturbation is much more difficult in general. However, for a singular perturbation problem of two time scales of relaxation type, upper semi-continuity of statistical properties are still valid in some appropriate sense (after projection or lifting, see [40] for the case of singular limit of infinite Prandtl number in the Boussinesq system for convection). The singularity usually involves an initial layer in time [37] and hence render the problem not that singular if one considers long time behavior (such as stationary statistical properties). In terms of long time statistics, we can imagine that the fast variable quickly relaxes and hence essentially slaved by the slow variable at large time. Therefore we have that the long time statistics is essentially given by the slow dynamics with the fast dynamics slaved by the slow variable, i.e., the limit dynamics. To be more precise, we consider the following type of two time scale problem of relaxation type which is basically the same as those we considered for the problem of global attractors [39]

$$\varepsilon(\frac{d\mathbf{u}_1}{dt} + g(\mathbf{u}_1, \mathbf{u}_2)) = f_1(\mathbf{u}_1, \mathbf{u}_2), \ \mathbf{u}_1(0) = \mathbf{u}_{10}, \tag{3}$$

$$\frac{d\mathbf{u}_2}{dt} = f_2(\mathbf{u}_1, \mathbf{u}_2), \ \mathbf{u}_2(0) = \mathbf{u}_{20}, \tag{4}$$

where X_1, X_2 are two separable Hilbert spaces. The limit problem for $\varepsilon = 0$ is given by

$$0 = f_1(\mathbf{u}_1^0, \mathbf{u}_2^0), \tag{5}$$

$$\frac{d\mathbf{u}_2^0}{dt} = f_2(\mathbf{u}_1^0, \mathbf{u}_2^0), \mathbf{u}_2^0(0) = \mathbf{u}_{20}. \tag{6}$$

This is a two time scale problem with $\mathbf{u}_1$ being the fast variable and $\mathbf{u}_2$ being the slow variable.

Theorem 6. *[Upper semi-continuity of $\mathcal{IM}$, singular version] Consider a generalized dynamical system on $X_1 \times X_2$ with two explicitly separated time scales given by (3, 4) with the limit system given by (5,6).*

We postulate the following assumptions:

- H1 *(uniform dissipativity of the perturbed system) The two-time-scale system (3, 4) possesses a global attractor $\mathcal{A}_\varepsilon$ for all small positive ε such that $K = \bigcup_{0<\varepsilon<\varepsilon_0} \mathcal{A}_\varepsilon$ is pre-compact in $X_1 \times X_2$.*
- H2 *(dissipativity of the limit system) The limit system is wellposed and possesses a global attractor $\mathcal{A}_0$ in X_2.*
- H3 *(convergence of the slow variable) The slow variable of the solutions of the two time scale system converge uniformly on $\mathcal{A}_\varepsilon$, i.e.*

$$\lim_{\varepsilon\to 0} \sup_{\mathbf{u}_2 \in \mathcal{P}_2\mathcal{A}_\varepsilon} \|\mathcal{P}_2 S(t,\varepsilon)(F_1(\mathbf{u}_2, 0), \mathbf{u}_2) - S(t,0)\mathbf{u}_2\|_{X_2} = 0, \ \forall\ t \geq 0.$$

  where $\mathcal{P}_2$ is the projection from $X_1 \times X_2$ to X_2 defined as $\mathcal{P}_2(\mathbf{u}_1, \mathbf{u}_2) = \mathbf{u}_2$, and $\mathbf{u}_1 = F_1(\mathbf{u}_2, 0)$ is the unique solution to the first part of the limit system, i.e. $0 = f_1(F_1(\mathbf{u}_2, 0), \mathbf{u}_2)$.
- H4 *(smallness of the perturbation) The two time scale problem (3, 4) is a formally uniformly small perturbation of the limit problem (5, 6) when confined to the*

global attractors, i.e.,

$$\lim_{\varepsilon\to 0}\sup_{(\mathbf{u}_1,\mathbf{u}_2)\in\mathcal{A}_\varepsilon}\|\varepsilon(\frac{d\mathbf{u}_1}{dt}+g(\mathbf{u}_1,\mathbf{u}_2))\|_{X_1}=0.$$

H5 *(continuity of the slave relation) The first equation in the limit system (5) can be solved continuously for x_1^0 with given x_2^0 and a nontrivial left hand side, i.e., there exists a continuous function $F_1:\ X_2\times X_1\to X_1$ such that*

$$y=f_1(F_1(\mathbf{u}_2,y),\mathbf{u}_2).$$

Moreover, we assume F_1 is uniformly continuous for $y=0$ and $\mathbf{u}_2\in\mathcal{P}_2K$.

Then the stationary statistical properties are upper semi-continuous after lifting in this singular limit, i.e., for $\{\mu_\varepsilon\in\mathcal{IM}_\varepsilon, 0<\varepsilon\le\varepsilon_0\}$, there exists a weakly convergent subsequence, still denoted $\{\mu_\varepsilon\}$, and $\mu_0\in\mathcal{IM}_0$ such that

$$\mu_\varepsilon\rightharpoonup\mathcal{L}\mu_0,$$

where $\mathcal{L}$ is the lift from $\mathcal{PM}(X_2)$ to $\mathcal{PM}(X_1\times X_2)$ defined by

$$\int_{X_1\times X_2}\varphi(\mathbf{u}_1,\mathbf{u}_2)\,d(\mathcal{L}\mu)(\mathbf{u}_1,\mathbf{u}_2)=\int_{X_2}\varphi(F_1(\mathbf{u}_2,0),\mathbf{u}_2)\,d\mu(\mathbf{u}_2).$$

Proof. We will approach the problem utilizing ideas developed in [39, 40].

We first show that the statistical properties converge in the projected sense, i.e., the statistical properties converge when restricted to the slow manifold (equivalent of taking marginal distribution in some appropriate sense).

We first perform the following change of variables

$$\mathbf{u}_1=y_1+F_1(y_2,0),\quad \mathbf{u}_2=y_2$$

where F_1 is the one defined through the slave relation.

Let μ_ε be an invariant measure of the perturbed system on $X_1\times X_2$, the change of variable induces another Borel probability measure $\tilde\mu_\varepsilon$ on $X_1\times X_2$ which is defined by

$$\int\varphi(\mathbf{u}_1,\mathbf{u}_2)\,d\mu_\varepsilon(\mathbf{u}_1,\mathbf{u}_2)=\int\varphi(y_1+F_1(y_2,0),y_2)\,d\tilde\mu_\varepsilon(y_1,y_2).$$

Thanks to the uniform dissipativity assumption, the set $\{\mu_\varepsilon\}$ is tight in the space of Borel probability measures. This also implies that the set $\{\tilde\mu_\varepsilon\}$ is tight in the space of Borel probability measures on $X_1\times X_2$ since F_1 is continuous and K is pre-compact. Therefore the marginal distribution of $\{\tilde\mu_\varepsilon\}$ in X_2, denoted $\{M\tilde\mu_\varepsilon\}$, is also tight in the space of all Borel probability measures on X_2. Hence there must exist a weakly convergent subsequence so that

$$M\tilde\mu_\varepsilon\rightharpoonup\mu_0,\varepsilon\to 0.$$

Our first goal is to show that $\mu_0\in\mathcal{IM}_0$, i.e., the upper semi-continuity of the stationary statistical properties in the projected sense.

For this purpose we take a smooth cylindrical test functional φ_0 on X_2 and show the invariance of the average of φ_0 under the flow. Indeed, for any $t > 0$,

$$
\begin{aligned}
& \int_{X_2} \varphi_0(S^0(t)y_2)\, d\mu_0(y_2) \\
= \; & \lim_{\varepsilon\to 0} \int_{X_2} \varphi_0(S^0(t)y_2)\, d(M\tilde{\mu}_\varepsilon)(y_2) \\
= \; & \lim_{\varepsilon\to 0} \int_{X_1\times X_2} \varphi_0(S^0(t)y_2)\, d\tilde{\mu}_\varepsilon(y_1, y_2) \\
= \; & \lim_{\varepsilon\to 0} \int_{X_1\times X_2} \varphi_0(S^0(t)\mathbf{u}_2)\, d\mu_\varepsilon(\mathbf{u}_1, \mathbf{u}_2) \\
= \; & \lim_{\varepsilon\to 0} \int_{X_1\times X_2} \varphi_0(\mathcal{P}_2 S^\varepsilon(t)(F_1(\mathbf{u}_2, 0), \mathbf{u}_2))\, d\mu_\varepsilon(\mathbf{u}_1, \mathbf{u}_2) \\
& + \lim_{\varepsilon\to 0} \int_{X_1\times X_2} (\varphi_0(S^0(t)\mathbf{u}_2) - \varphi_0(\mathcal{P}_2 S^\varepsilon(t)(F_1(\mathbf{u}_2, 0), \mathbf{u}_2)))\, d\mu_\varepsilon(\mathbf{u}_1, \mathbf{u}_2) \\
= \; & \lim_{\varepsilon\to 0} \int_{X_1\times X_2} \varphi_0(\mathcal{P}_2(F_1(\mathbf{u}_2, 0), \mathbf{u}_2))\, d\mu_\varepsilon(\mathbf{u}_1, \mathbf{u}_2) \\
& + \lim_{\varepsilon\to 0} \int_{X_1\times X_2} (\varphi_0(S^0(t)\mathbf{u}_2) - \varphi_0(\mathcal{P}_2 S^\varepsilon(t)(F_1(\mathbf{u}_2, 0), \mathbf{u}_2)))\, d\mu_\varepsilon(\mathbf{u}_1, \mathbf{u}_2) \\
= \; & \lim_{\varepsilon\to 0} \int_{X_1\times X_2} \varphi_0(\mathbf{u}_2)\, d\mu_\varepsilon(\mathbf{u}_1, \mathbf{u}_2) \\
= \; & \int_{X_2} \varphi_0(y_2)\, d\mu_0(y_2)
\end{aligned}
$$

where we have used the weak convergence of $M\tilde{\mu}_\varepsilon$, the change of variables/measures (definition of $\tilde{\mu}_\varepsilon$), the invariance of μ_ε under $S^\varepsilon(t)$, and the following straightforward estimate

$$
\begin{aligned}
& |\varphi_0(S^0(t)\mathbf{u}_2) - \varphi_0(\mathcal{P}_2 S^\varepsilon(t)(F_1(\mathbf{u}_2, 0), \mathbf{u}_2))| \\
\leq \; & \|\varphi_0'\| \|S^0(t)\mathbf{u}_2 - \mathcal{P}_2 S^\varepsilon(t)(F_1(\mathbf{u}_2, 0), \mathbf{u}_2)\| \\
\to \; & 0, \varepsilon \to 0
\end{aligned}
$$

uniformly for $\mathbf{u}_2 \in \mathcal{P}_2\mathcal{A}_\varepsilon$ by the uniform convergence of the slow variable assumption.

A general continuous test functional can be approximated by smooth finite dimensional cylindrical functional just as in the case of regular perturbation.

This ends the the proof of the convergence in the projected sense.

For the convergence in the lifted sense, we have, for any smooth cylindrical test functional φ on $X_1 \times X_2$

$$\begin{aligned}
& |\int_{X_1\times X_2} \varphi(\mathbf{u}_1,\mathbf{u}_2)\, d\mu_\varepsilon(\mathbf{u}_1,\mathbf{u}_2) - \int_{X_1\times X_2} \varphi(\mathbf{u}_1,\mathbf{u}_2)\, d(\mathcal{L}\mu_0)(\mathbf{u}_1,\mathbf{u}_2)| \\
= \quad & |\int_{X_1\times X_2} \varphi(\mathbf{u}_1,\mathbf{u}_2)\, d\mu_\varepsilon(\mathbf{u}_1,\mathbf{u}_2) - \int_{X_2} \varphi(F_1(\mathbf{u}_2,0),\mathbf{u}_2)\, d\mu_0(\mathbf{u}_2)| \\
\leq \quad & |\int_{X_1\times X_2} \varphi(F_1(\mathbf{u}_2,0),\mathbf{u}_2)\, d\mu_\varepsilon(\mathbf{u}_1,\mathbf{u}_2) - \int_{X_2} \varphi(F_1(\mathbf{u}_2,0),\mathbf{u}_2)\, d\mu_0(\mathbf{u}_2)| \\
& +|\int_{X_1\times X_2} (\varphi(\mathbf{u}_1,\mathbf{u}_2) - \varphi(F_1(\mathbf{u}_2,0),\mathbf{u}_2))\, d\mu_\varepsilon(\mathbf{u}_1,\mathbf{u}_2)| \\
\leq \quad & |\int_{X_1\times X_2} \varphi(F_1(\mathbf{u}_2,0),\mathbf{u}_2)\, dM\tilde{\mu}_\varepsilon(\mathbf{u}_2) - \int_{X_2} \varphi(F_1(\mathbf{u}_2,0),\mathbf{u}_2)\, d\mu_0(\mathbf{u}_2)| \\
& + \int_{X_1\times X_2} \|\varphi'\| \|\mathbf{u}_1 - F_1(\mathbf{u}_2,0)\|\, d\mu_\varepsilon(\mathbf{u}_1,\mathbf{u}_2) \\
\leq \quad & |\int_{X_1\times X_2} \varphi(F_1(\mathbf{u}_2,0),\mathbf{u}_2)\, dM\tilde{\mu}_\varepsilon(\mathbf{u}_2) - \int_{X_2} \varphi(F_1(\mathbf{u}_2,0),\mathbf{u}_2)\, d\mu_0(\mathbf{u}_2)| \\
& + \int_{X_1\times X_2} \|\varphi'\| \|F_1(\mathbf{u}_2,\varepsilon(\frac{d\mathbf{u}_1}{dt} + g(\mathbf{u}_1,\mathbf{u}_2))) - F_1(\mathbf{u}_2,0)\|\, d\mu_\varepsilon(\mathbf{u}_1,\mathbf{u}_2) \\
\to \quad & 0, \varepsilon \to 0
\end{aligned}$$

where we have utilized the definition of lift, marginal distribution, mean value theorem, the slave property and the smallness of the perturbation and the uniform continuity of the slave property.

Again, a general continuous functional can be approximated by smooth finite dimensional cylindrical functional just as the case of regular perturbation.

This ends the proof of the theorem. □

A corollary of this result together with the compactness of $\mathcal{IM}_0$ (theorem 1) and the fact that extremal points of $\mathcal{IM}$ are ergodic (theorem 3) leads us to the upper semi-continuity of extreme time averaged statistics.

Theorem 7. *[Upper semi-continuity of extremal time averaged statistics, singular version] Under the same assumption as in the previous theorem, we have for any fixed continuous test functionals $\varphi_0(\mathbf{u}_1,\mathbf{u}_2)$ and $\varphi_{02}(\mathbf{u}_2)$, the extremal statistics are saturated by ergodic invariant measures, i.e., there exist ergodic invariant measures $\nu_\epsilon \in \mathcal{IM}_\epsilon, \nu_0 \in \mathcal{IM}_0$ such that*

$$\begin{aligned}
& \sup_{(\mathbf{u}_1,\mathbf{u}_2)\in X_1\times X_2} \limsup_{T\to\infty} \frac{1}{T}\int_0^T \varphi_0(S(t,\epsilon)(\mathbf{u}_1,\mathbf{u}_2))\, dt \\
= & \int_{X_1\times X_2} \varphi_0(\mathbf{u}_1,\mathbf{u}_2)\, d\nu_\epsilon(\mathbf{u}_1,\mathbf{u}_2),
\end{aligned}$$

$$\sup_{\mathbf{u}_2\in X_2} \limsup_{T\to\infty} \frac{1}{T}\int_0^T \varphi_{02}(S(t,0)\mathbf{u}_2)\, dt \quad = \quad \int_{X_2} \varphi_{02}(\mathbf{u}_2)\, d\nu_0(\mathbf{u}_2).$$

Moreover, the extremal statistics are upper semi-continuous in parameter, i.e.,

$$\begin{aligned}&\limsup_{\epsilon\to 0}\sup_{(\mathbf{u}_1,\mathbf{u}_2)\in X_1\times X_2}\limsup_{T\to\infty}\frac{1}{T}\int_0^T\varphi_0(S(t,\epsilon)(\mathbf{u}_1,\mathbf{u}_2))\,dt\\ \le\ &\sup_{\mathbf{u}_2\in X_2}\limsup_{T\to\infty}\frac{1}{T}\int_0^T\varphi_0(F_1(S(t,0)\mathbf{u}_2,0),S(t,0)\mathbf{u}_2)\,dt\end{aligned}$$

Proof. The first half of the proof, saturation by ergodic measures, is exactly the same as the regular perturbation part (theorem 5).

For the second part, the upper semi-continuity, we have, assuming $\nu_\varepsilon \rightharpoonup \mathcal{L}\mu_0$,

$$\begin{aligned}&\limsup_{\epsilon\to 0}\sup_{(\mathbf{u}_1,\mathbf{u}_2)\in X_1\times X_2}\limsup_{T\to\infty}\frac{1}{T}\int_0^T\varphi_0(S(t,\epsilon)(\mathbf{u}_1,\mathbf{u}_2))\,dt\\ =\ &\limsup_{\epsilon\to 0}\int_{X_1\times X_2}\varphi_0(\mathbf{u}_1,\mathbf{u}_2)\,d\nu_\epsilon\quad\text{(saturation by ergodic measure)}\\ =\ &\int_{X_1\times X_2}\varphi_0(\mathbf{u}_1,\mathbf{u}_2)\,d\mathcal{L}\mu_0\quad\text{(weak convergence after lift)}\\ =\ &\int_{X_2}\varphi_0(F_1(\mathbf{u}_2,0),\mathbf{u}_2)\,d\mu_0\quad\text{(definition of lift)}\\ \le\ &\int_{X_2}\varphi_0(F_1(\mathbf{u}_2,0),\mathbf{u}_2)\,d\nu_0\quad\text{(definition of }\nu_0)\\ =\ &\lim_{T\to\infty}\frac{1}{T}\int_0^T\varphi_0(F_1(S(t,0)\mathbf{u}_2,0),S(t,0)\mathbf{u}_2)\,dt\quad(a.s.\ w.r.t.\ \nu_0,\ \text{ergodicity of }\nu_0)\\ =\ &\sup_{\mathbf{u}_2\in X_2}\limsup_{T\to\infty}\frac{1}{T}\int_0^T\varphi_0(F_1(S(t,0)\mathbf{u}_2,0),S(t,0)\mathbf{u}_2)\,dt.\quad\text{(definition of }\nu_0)\end{aligned}$$

This ends the proof of the theorem. □

4. **Remarks on applications and extensions.** Here we give a few remarks on the application of the abstract theorems to Rayleigh-Bénard convection at large Prandtl number, and convection in fluid saturated porous media (Darcy-Boussinesq system).

We recall the well-known **Boussinesq system for Rayleigh-Bénard convection (non-dimensional)** [3, 33]:

$$\begin{aligned}\frac{1}{Pr}(\frac{\partial\mathbf{u}}{\partial t}+(\mathbf{u}\cdot\nabla)\mathbf{u})+\nabla p &= \Delta\mathbf{u}+Ra\,\mathbf{k}\theta,\quad \nabla\cdot\mathbf{u}=0,\quad \mathbf{u}|_{z=0,1}=0,\\ \frac{\partial\theta}{\partial t}+\mathbf{u}\cdot\nabla\theta-u_3 &= \Delta\theta,\quad \theta|_{z=0,1}=0,\end{aligned}$$

where $\mathbf{u}$ is the fluid velocity field, p is the kinematic pressure, θ is the deviation of the temperature field from the pure conduction state $1-z$, $\mathbf{k}$ is the unit upward vector, Ra is the *Rayleigh number*, Pr is the *Prandtl number*, and the fluids occupy the (non-dimensionalized) region $\Omega=[0,L_x]\times[0,L_y]\times[0,1]$ with periodicity in the horizontal directions assumed for simplicity.

Although the Boussinesq system does not define a dynamical system due to the well-known difficulty associated with the three dimensional incompressible fluid Navier-Stokes system, we have eventual regularity at large Prandtl number [37, 38, 39]. Hence the generalized dynamical system can be treated as a usual dynamical

system in terms of long time behavior (global attractors, stationary statistical properties) at large Prandtl number. The phase space in this case is given by $H \times L^2$ where H is the divergence free subspace of $(L^2)^3$ with zero normal trace at $z = 0, 1$. The results presented in the previous sections (except the singular perturbation one) apply to the Boussinesq system at large Prandtl number. In particular, for the continuous test functional $\varphi_0(\mathbf{u}, \theta) = 1 + \frac{1}{|\Omega|} \int_\Omega u_3(\mathbf{x})\theta(\mathbf{x})\, d\mathbf{x}$ defined on $H \times L^2$, we have upper semi-continuity of the Nusselt number in the Rayleigh number and Prandtl number. We leave the details to the interested reader.

For the singular perturbation case, we have $X_1 = H$ and $X_2 = L^2$, and the limit problem is given by the **infinite Prandtl number model**

$$\begin{aligned} \nabla p^0 &= \Delta \mathbf{u}^0 + Ra\, \mathbf{k}\theta^0, \quad \nabla \cdot \mathbf{u}^0 = 0, \quad \mathbf{u}^0|_{z=0,1} = 0, \\ \frac{\partial \theta^0}{\partial t} + \mathbf{u}^0 \cdot \nabla \theta^0 - u_3^0 &= \Delta \theta^0, \quad \theta^0|_{z=0,1} = 0. \end{aligned}$$

The results regarding singular perturbation (theorems 6, 7) also apply with $\varepsilon = \frac{1}{Pr}$ and $F_1(\theta, y) = Ra\, A^{-1}(\mathbf{k}\theta) - A^{-1}(y)$ where A is the Stokes operator with the associated boundary conditions. The uniform dissipativity was derived in [38, 39]. The convergence of the slow variable was derived in [37]. The continuity of the slave relation is obvious and the formal smallness of the perturbation was verified in [39, 40]. In particular, we have upper semi-continuity on Prandtl number even in the singular limit of infinite Prandtl number. The interested reader is referred to [40] for more details.

Next, we consider an application of the abstract results to the case of convection in fluid saturated porous media. The governing equation is the following **Darcy-Oberbeck-Boussinesq system (non-dimensional)** [26]:

$$\begin{aligned} \gamma_a \frac{\partial \mathbf{v}}{\partial t} + \mathbf{v} + \nabla p &= Ra_D\, \mathbf{k}T, \quad \nabla \cdot \mathbf{v} = 0, \quad v_3|_{z=0,1} = 0, \\ \frac{\partial \theta}{\partial t} + \mathbf{v} \cdot \nabla \theta - v_3 &= \Delta \theta, \quad \theta|_{z=0,1} = 0 \end{aligned}$$

where $\mathbf{v}$ is the non-dimensional seepage velocity, p is the non-dimensional pressure, $T = 1 - z + \theta$ is the non-dimensional temperature. The parameters in the system are given by the *Prandtl-Darcy number* γ_a^{-1}, and the *Rayleigh-Darcy number* Ra_D. Again we assume the fluids occupy the (non-dimensionalized) region $\Omega = [0, L_x] \times [0, L_y] \times [0, 1]$ with periodicity in the horizontal directions assumed for simplicity.

The system is partially/weakly dissipative since there is no dissipation in the velocity equation. In particular, there is no compact absorbing ball in the phase space $H \times L^2$. However, the results presented in the previous sections (except the singular perturbation one) apply to the Darcy-Boussinesq system. In particular, for the continuous test functional $\varphi_0(\mathbf{v}, \theta) = 1 + \frac{1}{|\Omega|} \int_\Omega v_3(\mathbf{x})\theta(\mathbf{x})\, d\mathbf{x}$ defined on $H \times L^2$, we have upper semi-continuity of the Nusselt number in the Rayleigh-Darcy number and Prandtl-Darcy number. For the singular perturbation case, we have $X_1 = H$ and $X_2 = L^2$, and the limit problem is given by the **infinite Prandtl-Darcy number model**

$$\begin{aligned} \mathbf{u} + \nabla p &= Ra_D \mathbf{k}T, \quad \nabla \cdot \mathbf{u} = 0, \quad u_3|_{z=0,1} = 0, \\ \frac{\partial \theta}{\partial t} + \mathbf{u} \cdot \nabla \theta - u_3 &= \Delta \theta, \quad \theta|_{z=0,1} = 0. \end{aligned}$$

The results regarding singular perturbation (theorems 6, 7) also apply with $\varepsilon = \gamma_a$ and $F_1(\theta, y) = Ra_D\, P(\mathbf{k}\theta) - P(y)$ where P is the Leray-Hopf projection from $(L^2)^3$ to H. The interested reader is referred to [27] for more details.

Many questions remain open. For instance, although we have non-uniqueness in general in terms of invariant measures, it is very likely there is a unique physically observable one. A question then is to have physically interesting and verifiable conditions on the uniqueness of physically relevant invariant measures. Some interesting progress has been made on axiom-A systems [43]. However, it seems extremely difficult if not impossible to verify those conditions listed there on any practical fluid system. On the other hand, we may consider noisy system since our world is intrinsically noisy. Appropriate noise may lead to uniqueness of invariant measure (see for instance [9, 11, 18, 23, 17, 12]) since appropriate noise may help to connect different parts of the (deterministic) global attractor that are otherwise disconnected (having disjoint basin of attraction). Nevertheless, it is non-trivial to find physically relevant noises. Also, the possible selection of physically relevant invariant measure through zero noise limit is not well understood (see [43] for axiom-A system however). There are many advantages of having unique invariant measure including continuous dependence on parameters for uniformly dissipative systems. Even in the case with unique invariant measure, accurate and efficient methods for estimating leading order change of various statistical quantities are still not well understood (this is related to the linear response theory, see for instance [21]). More generally, we do not expect continuity of arbitrary statistical quantity. But can we expect continuity of physically relevant statistical quantities? If no continuity, can we classify the bifurcation phenomena for statistical behaviors? On a more practical side, we still do not have good a priori estimates on the (long) time interval needed for temporal average to effectively approximate spatial average (related to the rate of mixing). The study of numerical methods that are good for capturing stationary statistical properties is still in its infancy (see [4, 5]). Explicit physically relevant estimates on specific statistical quantities are also rare (see estimates on Nusselt number [6, 7, 10, 41]). Another practical (difficult) problem would be statistical inversion, i.e., finding relevant information on the system with available statistical observation (see for instance [16] for more details).

Acknowledgements. Part of the work was done during the author's visit to The Institute of Mathematical Sciences at The Chinese University of Hong Kong, and the School of Mathematics at Fudan University, Shanghai China.

REFERENCES

[1] M. B. Bekka and M. Mayer, "Ergodic Theory and Topological Dynamics of Group Actions on Homogeneous Spaces," London Mathematical Society Lecture Note Series, 269, Cambridge University Press, Cambridge, 2000.

[2] P. Billingsley, "Weak Convergence of Measures: Applications in Probability," Conference Board of the Mathematical Sciences Regional Conference Series in Applied Mathematics, No. 5, SIAM, Philadelphia, 1971.

[3] S. Chandrasekhar, "Hydrodynamic and Hydro-Magnetic Stability," The International Series of Monographs on Physics, Clarendon Press, Oxford, 1961.

[4] W. Cheng and X. Wang, *A semi-implicit scheme for stationary statistical properties of the infinite Prandtl number model*, SIAM Jour. Num. Anal., accepted July, 2008.

[5] W. Cheng and X. Wang, "A Uniformly Dissipative Scheme for Stationary Statistical Properties of the Infinite Prandtl Number Model," Applied Mathematics Letters, in press. doi:10.1016/j.aml.2007.07.036

[6] P. Constantin and C. R. Doering, *Heat transfer in convective turbulence*, Nonlinearity, **9** (1996), 1049–1060.

[7] P. Constantin and C. R. Doering, *Infinite Prandtl number convection*, J. Stat. Phys., **94** (1999), 159–172.

[8] P. Constantin and F. Ramos, *Inviscid limit for damped and driven incompressible Navier-Stokes equations in* $\mathcal{R}^2$, preprint 2007. arXiv:math/0611782

[9] G. Da Prato and J. Zabczyk, "Ergodicity for Infinite Dimensional Systems," London Mathematical Society Lecture Note Series, 229, Cambridge University Press, Cambridge, 1996.

[10] C. R. Doering, F. Otto and M. G. Reznikoff, *Bounds on vertical heat transport for infinite Prandtl number Rayleigh-Bénard convection*, J. Fluid Mech., **560** (2006), 229–241.

[11] E, Weinan, "Stochastic Hydrodynamics," Current developments in mathematics, 2000, 109–147, Int. Press, Somerville, MA, 2001.

[12] W. E and D. Li, *The Andersen thermostat in molecular dynamics*, preprint, 2007.

[13] P. Fabrie and B. Nicolaenko, *Exponential attractors for nondissipative systems modeling convection in porous media*, Asymptotic Analysis, **12** (1996), 295–327.

[14] C. Foias, O. Manley, R. Rosa and R. Temam, "Navier-Stokes Equations and Turbulence," Encyclopedia of Mathematics and its Applications, 83, Cambridge University Press, Cambridge, 2001.

[15] J. K. Hale, "Asymptotic Behavior of Dissipative Systems," Mathematical Surveys and Monographs, 25, American Mathematical Society, Providence, RI, 1988.

[16] J. Kaipio and E. Somersalo, "Statistical and Computational Inverse Problems," Applied Mathematical Sciences 160, Springer-Verlag, New York, 2005.

[17] S. B. Kuksin, "Randomly Forced Nonlinear PDEs and Statistical Hydrodynamics in 2 Space Dimensions," Zurich Lectures in Advanced Mathematics, European Mathematical Society (EMS), Zürich, 2006.

[18] A. Lasota and M. C. Mackey, "Chaos, Fractals, and Noise: Stochastic Aspects of Dynamics," 2nd ed., Applied Mathematical Sciences, 97, Springer-Verlag, New York, 1994.

[19] P. D. Lax, "Functional Analysis," Pure and Applied Mathematics (New York), Wiley-Interscience [John Wiley & Sons], New York, 2002.

[20] T.-T. Li and T. H. Qin, "Physics and Partial Differential Equations," Higher Education Press, 1997.

[21] A. J. Majda, R. Abramov and M. Grote, "Information Theory and Stochastics for Multiscale Nonlinear Systems," CRM monograph series 25, American Mathematical Society, RI, 2005

[22] A. J. Majda and X. Wang, "Non-linear Dynamics and Statistical Theory for Basic Geophysical Flows," Cambridge University Press, Cambridge, 2006.

[23] J. C. Mattingly, "On Recent Progress for the Stochastic Navier Stokes Equations," Journés Équations aux Dérivées Partielles" (Forges-les-Eaux, 2003), vol. XV (Summer, 2003), viii+298, Université de Nantes, Nantes.

[24] I. Moise, R. Rosa and X. Wang, *Attractors for non-compact semigroups via energy equations*, Nonlinearity, **11** (1998), 1369–1393.

[25] A. S. Monin and A. M. Yaglom, "Statistical Fluid Mechanics; Mechanics of Turbulence," English ed. updated, augmented and rev. by the authors, MIT Press, Cambridge, Mass., 1975.

[26] D. Nield and A. Bejan, "Convection in Porous Media, 2nd ed," Springer-Verlag, New York, 1999.

[27] R. D. Parshad and X. Wang, *Asymptotic behavior of Darcy-Boussinesq system at large Prandtl-Darcy number*, in preparation, 2008.

[28] G. Raugel, "Global Attractors in Partial Differential Equations," Handbook of dynamical systems, Vol. **2**, 885–982, North-Holland, Amsterdam, 2002.

[29] H. L. Royden, "Real Analysis, 3rd ed.," Macmillan Publishing Company, New York, 1988.

[30] Ya. G. Sinai, "Topics in Ergodic Theory," Princeton Mathematical Series, 44, Princeton University Press, Princeton, NJ, 1994.

[31] R. M. Temam, "Navier-Stokes Equations and Nonlinear Functional Analysis, 2nd edition," CBMS-NSF Regional Conference Series in Applied Mathematics, 66, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1995.

[32] R. M. Temam, "Infinite Dimensional Dynamical Systems in Mechanics and Physics, 2nd ed," Applied Mathematical Sciences, 68, Springer-Verlag, New York, 1997.

[33] D. J. Tritton, "Physical Fluid Dynamics," Second edition. Oxford Science Publications, The Clarendon Press, Oxford University Press, New York, 1988.

[34] M. I. Vishik and A. V. Fursikov, "Mathematical Problems of Statistical Hydromechanics," Kluwer Acad. Publishers, Dordrecht/Boston/London, 1988.

[35] P. Walters, "An Introduction to Ergodic Theory," Graduate Texts in Mathematics, 79, Springer-Verlag, New York-Berlin, 1982.

[36] X. Wang, *An energy equation for weakly damped driven nonlinear Schrödinger equations and its application to their attractors*, Physica D, **88** (1995), 167–175.

[37] X. Wang, *Infinite Prandtl number limit of Rayleigh-Bénard convection*, Comm. Pure. Appl. Math., **57** (2004), 1265–1282.

[38] X. Wang, *A note on long time behavior of solutions to the boussinesq system at large Prandtl number*, Nonlinear partial differential equations and related analysis, 315–323, Contemp. Math., 371, Amer. Math. Soc., Providence, RI, 2005.

[39] X. Wang, *Asymptotic behavior of global attractors to the Boussinesq system for Rayleigh-Bénard convection at large Prandtl number*, Comm. Pure and Appl. Math., **60** (2007), 1293–1318.

[40] X. Wang, *Stationary statistical properties of Rayleigh-Bénard convection at large Prandtl number*, Comm. Pure Appl. Math., **61** (2008), 789–815. DOI: 10.1002/cpa.20214.

[41] X. Wang, *Bound on the vertical heat transport at large Prandtl number*, Physica D, **237** (2008), 854–858. doi:10.1016/j.physd.2007.11.001

[42] X. Wang, *Lecture notes on elementary statistical theories with applications to fluid systems*, 2007 Shanghai Mathematics Summer School in Fudan University. To be published by *Higher Education* Press.

[43] L. S. Young, *What are SRB measures and which dynamical systems have them?* Dedicated to David Ruelle and Yasha Sinai on the occasion of their 65th birthdays, J. Stat. Phy., **108** (2002), 733–754.

Received November 2007; revised April 2008.

E-mail address: wxm@math.fsu.edu