NON-INTRUSIVE AND STRUCTURE PRESERVING MULTISCALE INTEGRATION OF STIFF ODES, SDEs AND HAMILTONIAN SYSTEMS WITH HIDDEN SLOW DYNAMICS VIA FLOW AVERAGING

MOLEI TAO, HOUMAN OWHADI AND JERROLD E. MARSDEN

Technical Report No. 2009-04
August 2009
Non-intrusive and structure preserving multiscale integration of stiff ODEs, SDEs and Hamiltonian systems with hidden slow dynamics via flow averaging.

Molei Tao\textsuperscript{1}, Houman Owhadi\textsuperscript{1,2}, and Jerrold E. Marsden\textsuperscript{1}.

August 9, 2009

Abstract

We introduce a new class of integrators for stiff ODEs as well as SDEs. An example of subclass of systems that we treat are ODEs and SDEs that are sums of two terms one of which has large coefficients. These integrators are (i) Multiscale: they are based on flow averaging and so do not resolve the fast variables but rather employ step-sizes determined by slow variables (ii) Basis: the method is based on averaging the flow of the given dynamical system (which may have hidden slow and fast processes) instead of averaging the instantaneous drift of assumed separated slow and fast processes. This bypasses the need for identifying explicitly (or numerically) the slow or fast variables. (iii) Non intrusive: A pre-existing numerical scheme resolving the microscopic time scale can be used as a black box and turned into one of the integrators in this paper by simply turning the large coefficients on over a microscopic timescale and off during a mesoscopic timescale. (iv) Convergent over two scales: strongly over slow processes and in the sense of measures over fast ones. We introduce the related notion of two scale flow convergence and analyze the convergence of these integrators under the induced topology. (v) Structure preserving: For stiff Hamiltonian systems (possibly on manifolds), they are symplectic, time-reversible, and symmetric (under the group action leaving the Hamiltonian invariant) in all variables. They are explicit and apply to arbitrary stiff potentials (that need not be quadratic). Their application to the Fermi-Pasta-Ulam problems shows accuracy and stability over 4 orders of magnitude of time scales. For stiff Langevin equations, they are symmetric (under a group action), time-reversible and Boltzmann-Gibbs reversible, quasi-symplectic on all variables and conformally symplectic with isotropic friction.

Contents

1 Introduction 3

1.1 Quick overview of the integrator \textsuperscript{.} 3

1.2 Related work \textsuperscript{.} 4

\textsuperscript{1}California Institute of Technology, Applied & Computational Mathematics, Control & Dynamical systems, MC 217-50 Pasadena, CA 91125

\textsuperscript{2}owhadi@caltech.edu
# 2 ODEs

2.1 Two scale flow convergence ........................................ 8
2.2 FLow AVeraging integratORS for stiff ODEs ........................ 9
  2.2.1 FLAVORS ........................................... 10
  2.2.2 Non intrusive FLAVORS ................................ 10
2.3 Convergence results and error bounds for FLAVORS .............. 11
2.4 Natural FLAVORS generalization to generic stiff ODEs .......... 12
2.5 Numerical experiments: systems of ODEs with hidden slow dynamics 14

# 3 Deterministic Hamiltonian Systems

3.1 Symplectic, symmetric (under a group action) and time-reversible FLA-
  VORS for mechanical systems on manifolds .......................... 15
  3.1.1 An example of symplectic FLAVOR .......................... 17
  3.1.2 An example of symplectic and time-reversible FLAVOR ........ 17
  3.1.3 An artificial FLAVOR .................................... 17
3.2 Stability ........................................................................ 18
3.3 Numerical error analysis of FLAVOR for nonlinear systems ....... 20
3.4 Numerical error analysis of FLAVOR for linear systems .......... 22
3.5 Nonlinear stiff and soft potentials ................................... 24
3.6 Fermi-Pasta-Ulam problem ............................................ 24
3.7 Nonlinear 2D Primitive Molecular Dynamics Example ........... 30
3.8 Nonlinear 2D Molecular Clipper ...................................... 31
3.9 Forced non-autonomous mechanical system: Kapitza’s Inverted Pendulum 33

# 4 SDEs

4.1 Two scale flow convergence for SDEs ............................... 35
4.2 Flow Averaging Integrators for stiff SDEs, Convergence theorem 37
4.3 Natural FLAVORS generalization to generic stiff SDEs .......... 39
4.4 Numerical experiments ................................................ 42

# 5 Stochastic Mechanical Systems: Langevin Equations

5.1 Quasi and conformally symplectic, symmetric and time-reversible FLA-
  VORS for stochastic mechanical systems on manifolds ............. 42
  5.1.1 Example of quasi-symplectic FLAVORS .................... 44
  5.1.2 Example of quasi-symplectic and time-reversible FLAVORS 44
  5.1.3 Example of Boltzmann-Gibbs reversible Metropolis-Adjusted FLA-
         VOR ..................................................... 44
5.2 Numerical implementation with slow noise and friction ........... 45
5.3 Numerical implementation with fast noise and friction .......... 46

# 6 Appendix

6.1 Proof of Theorem 2.2 ............................................. 47
6.2 Proof of Theorem 4.2 ............................................. 54
1 Introduction

1.1 Quick overview of the integrator.

Consider the following ODE on $\mathbb{R}^d$.

$$\dot{u}^\epsilon = G(u^\epsilon) + \frac{1}{\epsilon} F(u^\epsilon) \quad (1.1)$$

In Subsection 2.4, 3.1, 4.1, 4.3 and 5.1 we will consider the more general form (2.20), stiff deterministic Hamiltonian systems (3.1), SDEs ((4.1) and (4.15)) and Langevin equations ((5.1) and (5.2)) but for the sake of clarity we will start the description of our method with (1.1).

Because of the term involving $1/\epsilon$, the direct simulation of (1.1) requires a discretization in time with time steps of the order of $\epsilon$. For $\epsilon \ll 1$ such a direct simulation is impractical. Assume that there exists a diffeomorphism $\eta := (\eta^x, \eta^y)$, from $\mathbb{R}^d$ onto $\mathbb{R}^{d-p} \times \mathbb{R}^p$, separating slow and fast variables, i.e. such that the process $(x^\epsilon(t), y^\epsilon(t)) = (\eta^x(u^\epsilon(t)), \eta^y(u^\epsilon(t)))$ satisfies an ODE system of the form

$$\begin{cases}
\dot{x}^\epsilon = g(x^\epsilon, y^\epsilon), & x^\epsilon_0 = x_0 \\
\dot{y}^\epsilon = \frac{1}{\epsilon} f(x^\epsilon, y^\epsilon), & y^\epsilon_0 = y_0
\end{cases} \quad (1.2)$$

Then if the trajectories of the fast degrees of freedom are locally ergodic then it is known (we refer for instance to Theorem 14, Section 3 of Chapter II of [82]) that $x^\epsilon$ converges towards $x_t$ defined as the solution to the ODE

$$\dot{x} = \int g(x, y) \mu(x, dy), \quad x|_{t=0} = x_0 \quad (1.3)$$

and $\mu(x, dy)$ is the ergodic measure associated with the solution to the ODE

$$\dot{y} = f(x, y) \quad (1.4)$$

It follows that the slow behavior of solutions of (1.1) can be simulated over coarse time steps by first identifying the slow process $x^\epsilon$ and then using numerical approximations of solutions of (1.3) to approximate $x^\epsilon$. Two classes of integrators have been founded on this observation: The equation free method [52, 51] and the Heterogeneous Multiscale Method [30, 32, 29, 5]. One shared rationale of those integrators is, after identification of the slow variables, to use a micro-solver to approximate the effective drift in (1.3) by averaging the instantaneous drift $g$ with respect to numerical solutions of (1.4) over a time span larger than the mixing time of the solution to (1.4).

In this paper we propose a new method based on the averaging of the instantaneous flow of the ODE (1.1) with hidden slow and fast variables instead of the instantaneous drift of $x^\epsilon$ in ODE (1.2) with separated slow and fast variables. We have called the resulting class of numerical integrators FLow AVeraging integratORS (FLAVORS).

Since FLAVORS are directly applied to (1.1), hidden slow variables do not need to be identified, either explicitly or numerically. Furthermore FLAVORS can be implemented
using an arbitrary legacy integrator $\Phi^\frac{1}{\epsilon}$ for (1.1) in which the parameter $\frac{1}{\epsilon}$ can be controlled. More precisely, assume that there exists a constant $h_0 > 0$ such that $\Phi^\alpha$ satisfies for all $t \leq h_0 \min(\frac{1}{\alpha}, 1)$

$$|\Phi^\alpha_t(u) - u - tG(u) - \alpha tF(u)| \leq C t^2 (1 + \alpha)^2$$

(1.5)
then FLAVOR can be defined as the algorithm simulating the process

$$\bar{u}_t = (\Phi^\frac{1}{\epsilon}_{\delta - \tau} \circ \Phi^\frac{1}{\tau}_{\tau})^k(u_0) \quad \text{for} \quad k\delta \leq t < (k + 1)\delta$$

(1.6)
where $\tau$ is a fine time step resolving the “fast” time scale and $\delta$ is a mesoscopic time step independent of the “fast” time scale. We show that (1.6) is strongly accurate with respect to (hidden) slow variables and weakly (in the sense of measures) accurate with respect to (hidden) fast variables. Motivated by this observation, we introduce in Subsection 2.1 the related notion of “two scale flow convergence”. The accuracy of (1.6) (under the topology induced by two scale flow convergence) is established under the conditions $\tau \ll \epsilon \ll \delta \ll 1$ and $(\frac{\tau}{\epsilon})^2 \ll \delta \ll \frac{\tau}{\epsilon}$. In our numerical experiments we have used the “rule of thumb” $\delta \sim \gamma \frac{\tau}{\epsilon}$ where $\gamma$ is a small parameter (0.1 for instance). The rationale and mechanism behind FLAVORS is the approximation of the flow of (1.1) over a coarse time step $h$ (resolving the “slow” time scale) by the flow

$$\Phi_h := (\Phi^0_{\frac{h}{\tau} - \tau} \circ \Phi^\frac{1}{\tau}_\tau)^M$$

(1.7)
where $M$ is a positive integer corresponding to the number of “samples” used to average the flow ($\delta$ has to be identified with $\frac{h}{M}$). Since FLAVORS are obtained by flow-composition we will show in Section 3 and 5 that they inherit the structure preserving properties (for instance symplecticity and symmetries under a group action) of the legacy integrator for Hamiltonian systems and Langevin equations.

The speed up at fixed accuracy (of FLAVORS compared to their legacy integrator) scales like $O(\frac{1}{\epsilon})$ or like $O(e^{\frac{\tau}{\epsilon}})$. We refer to Remark 2.5 of Theorem 2.2 for a detailed discussion.

1.2 Related work.

Dynamical systems with multiple time scales pose a major problem in simulations because the small time steps required for stable integration of the fast motions lead to large numbers of time steps required for the observation of slow degrees of freedom [89, 45]. Traditionally stiff dynamical systems have been separated into two classes with distinct integrators: stiff systems with fast transients and stiff systems with rapid oscillations [6, 29]. The former has been solved using implicit schemes [38, 28, 45, 47], Chebyshev methods [57, 1] or the projective integrator approach [40]. These latter have been solved using filtering techniques [37, 53, 78] or Poincaré map techniques [39, 72]. It has been observed in [29], that, at the present time, there exists no unified strategy for dealing with both classes of problems. When slow variables can be identified effective equations
can be obtained by averaging the instantaneous drift driving those slow variables\textsuperscript{82}. Two classes of numerical methods have been built on this observation: The equation free method\textsuperscript{52, 51} and the Heterogeneous Multiscale Method\textsuperscript{30, 32, 29, 5}.

In this paper we introduce a new method based on averaging the flow of the stiff dynamics instead of their instantaneous drift. We refer to the resulting class of numerical integrators as \textbf{F}low \textbf{A}veraging \textbf{i}ntegrat\textbf{ORS} (FLAVORS). We will show that

- FLAVORS apply in a unified way to both stiff systems with fast transients and stiff systems with rapid oscillations, with or without noise (sections 2 and 4), with a mesoscopic integration timestep chosen independently from the stiffness.

- Hidden slow variables do not need to be identified, explicitly or numerically. FLAVORS can be directly applied to ODEs with mixed (hidden) slow and fast variables (provided that the fast variables are locally ergodic/mixing).

- FLAVORS are non-intrusive (Subsection 2.2.2). A pre-existing numerical scheme resolving the microscopic time scale can be used as a black box and turned into a FLAVOR by simply turning on and off stiff parameters over a microscopic timescale (on) and a mesoscopic timescale (off).

We will show, analogous to\textsuperscript{3}, that the flow of FLAVORS converge over two scales: strongly over slow variables and in the sense of measures over fast variables. We call two scale flow convergence (or $F$-convergence), this type of convergence for ODEs (and SDEs). Later on, this paper proves the accuracy of FLAVORS in the topology induced by $F$-convergence.

Since averaging with FLAVORS is obtained by flow composition, FLAVORS have a natural extension to multiscale structure preserving integrators for stiff Hamiltonian systems (section 3). Structure preserving numerical methods for Hamiltonian systems have been developed in the framework of geometric numerical integration\textsuperscript{15, 58} and variational integrators\textsuperscript{63, 60}. The subject of geometric numerical integration deals with numerical integrators that preserve geometric properties of the flow of a differential equation, and it explains how structure preservation leads to an improved long-time behavior\textsuperscript{41}. Variational integration theory derives integrators for mechanical systems from discrete variational principles and are characterized by a discrete Noether theorem. These methods have excellent energy behavior over long integration runs because they are symplectic, i.e., by backward error analysis, they simulate a nearby mechanical system instead of nearby differential equations. Furthermore, statistical properties of the dynamics such as Poincaré sections are well preserved even with large time steps\textsuperscript{13}.

These structure-preserving properties of geometric numerical integrators motivated their extension to multiscale Hamiltonian systems. We refer to\textsuperscript{26} for a recent review on numerical integrators for highly oscillatory Hamiltonian systems. \textit{Symplectic integrators are natural for the integration of Hamiltonian systems since they reproduce at the discrete level an important geometric property of the exact flow\textsuperscript{17}.} For symplectic integrators with stiff quadratic potentials we refer to the Impulse Method, the Mollified Impulse, and its variations\textsuperscript{42, 88, 36, 76, 86}. In the context of variational integrators,
by defining a discrete Lagrangian with an explicit trapezoidal approximation of the soft potential and a midpoint approximation for the fast potential, a symplectic (IMEX—IMplicit–EXplicit) scheme for stiff Hamiltonian systems has been proposed in [83]. The resulting scheme is explicit for quadratic potentials and implicit for non quadratic stiff potentials. We also refer to Le Bris and Legoll’s (Hamilton-Jacobi derived) homogenization method [17]. Asynchronous Variational Integrators [59] provide a way to derive conservative symplectic integrators for PDEs where the solution advances non-uniformly in time; however stiff potentials require a fine time step discretization over the whole time evolution.

Several approaches to the homogenization of Hamiltonian systems (in analogy with classical homogenization [9, 19]) have been proposed. We refer to M-convergence introduced in [79, 12], to the two scale expansion of solutions of the Hamilton-Jacobi form of Newton’s equations (with stiff quadratic potentials) [17] and to PDE methods in weak KAM theory [33]. Obtaining explicit symplectic integrators for Hamiltonian systems with non quadratic stiff potentials is known to be an important and non trivial problem. By using Verlet/Leap-frog macro-solvers, methods that are symplectic on slow variables (when those variables can be identified) have been proposed in the framework of HMM (the Heterogeneous Multiscale Method) in [80, 21]. More recently, a reversible multiscale integration method for mechanical systems has been proposed in [6] in the context of HMM. By tracking slow variables, [6] enforces reversibility in all variables as an optimization constraint at each coarse step when minimizing the distance between the effective drift obtained from the micro-solver (in the context of HMM) and the drift of the macro-solver. We are also refer to [77] for HMM symmetric methods for mechanical systems with a stiff potentials of the form $\frac{1}{\epsilon} \sum_{j=1}^{\nu} g_j(q)^2$.

In the context of stiff Hamiltonian systems we will show that

- FLAVORS are symplectic, time-reversible and symmetric under a group action (Theorem 3.2) in all variables.
- FLAVORS are explicit with arbitrary (non necessarily quadratic or a sum of squares) stiff potentials.
- Hidden slow variables do not need to be identified (explicitly or numerically).
- FLAVORS have a natural extension to mechanical systems on manifolds such as constrained mechanical systems.
- If the pre-existing numerical scheme resolving the microscopic time scale is structure preserving (for instance, symplectic) then the non intrusive FLAVOR obtained by turning on and off the stiff parameter is structure-preserving (symplectic).
- When applied to the Fermi-Pasta-Ulam problem [34], FLAVORS remain accurate over 4 orders of magnitude of time scales (writing $\omega$ the frequency of stiff springs): $O(\frac{1}{\omega})$ (in the sense of measures), $O(1)$, $O(\omega)$ and $O(\omega^2)$ (strongly on slow variables).
Asymptotic problems for stochastic differential equations arose and were solved simultaneously with the very beginnings of the theory of such equations \[82\]. Here we refer to the early work of Gikhman \[41\], Krylov \[54, 55\], Bogolyubov \[11\] and Papanicolaou-Kohler \[71\]. We refer in particular to Skorokhod’s detailed monograph \[82\]. As for ODEs, effective equations for stiff SDEs can be obtained by averaging the instantaneous coefficients (drift and the diffusivity matrix squared) with respect to the fast components, we refer to Chapter II, Section 3 of \[82\] for a detailed analysis including error bounds. Numerical methods such as HMM \[31\] and equations free methods \[7\] have been extended to SDEs based on this averaging principle. Implicit methods in general fail to capture the effective dynamics of the slow time scale because they cannot correctly capture non-Dirac invariant distributions \[61\] (we refer to non-Dirac invariant distribution as a measure of probability on the configuration space whose support is not limited to a single point). An other idea is to treat fast variables by conditioning, here we refer to optimal prediction \[23, 22, 24\] that also been used for model reduction. We also refer to \[8, 43, 87, 19, 20, 61, 2\].

Since FLAVORS are obtained via flow averaging they have a natural extension to SDEs. We refer to Section 4 for this extension. As for ODEs, FLAVORS are directly applied to SDEs with mixed (hidden) slow and fast variables without prior (analytical or numerical) identification of slow variables. Furthermore, they can be implemented using a pre-existing scheme by turning on and off the stiff parameters.

Since the foundational work of Bismut \[10\], the field of stochastic geometric mechanics is emerging in response to the demand for tools to analyze the structure of continuous and discrete mechanical systems with uncertainty \[81, 48, 90, 25, 68, 69, 67, 56, 62, 15, 14, 16\]. Like their deterministic counterparts, these integrators are structure preserving in terms of statistical invariants.

In Section 5, FLAVORS are extended to structure preserving integrators for stiff stochastic mechanical systems, i.e., stiff Langevin equations. In particular it will be shown that FLAVORS for Langevin equations are

- Non intrusive.
- Quasi symplectic as defined by Conditions RL1 and RL2 of \[70\]. With isotropic friction, they are conformally symplectic \[65\]. They can also be made symmetric under a group action and time-reversible in all variables.
- Like their deterministic counterparts, FLAVORS for Langevin equations are explicit with arbitrary (not necessarily quadratic) stiff potentials. Hidden slow variables do not need to be identified and they can be implemented on manifolds.

2 ODEs

In this section, for the sake of clarity, we will start the detailed description and analysis of our method with \[1.1\]. In Subsection 2.4 we will consider the more general form \[2.20\].
2.1 Two scale flow convergence.

To analyze the behaviors of solutions of (1.1) we introduce the following notion of two scale flow convergence in analogy with homogenization theory for elliptic PDEs. This definition is motivated by the fact that solutions of (1.2) are characterized by both the strong convergence of the slow process \( x^\epsilon_t \) towards \( x_t \) and the weak convergence (in the sense of measures) of the fast process \( y^\epsilon_t \) towards \( \mu(x_t,dy) \).

Two scale flow convergence. Let \((\xi^\epsilon_t)_{t \in \mathbb{R}^+}\) be a sequence of processes on \(\mathbb{R}^d\) (functions from \(\mathbb{R}^+\) to \(\mathbb{R}^d\)) indexed by \(\epsilon > 0\). Let \((X_t)_{t \in \mathbb{R}^+}\) be a process on \(\mathbb{R}^{d-p}\) \((p \geq 0)\). Let \(x \mapsto \nu(x,dz)\) be a function from \(\mathbb{R}^{d-p}\) into the space of probability measures on \(\mathbb{R}^d\).

**Definition 2.1.** We say that the process \(\xi^\epsilon F\)-converges towards \(\nu(X,dz)\) as \(\epsilon \downarrow 0\) and write \(\xi^\epsilon F\lim_{\epsilon \to 0} \nu(X,dz)\) if and only if for all functions \(\varphi\) bounded and uniformly Lipschitz-continuous on \(\mathbb{R}^d\), and for all \(t > 0\),

\[
\lim_{h \to 0} \lim_{\epsilon \to 0} \frac{1}{h} \int_t^{t+h} \varphi(\xi^\epsilon_s) \, ds = \int_{\mathbb{R}^d} \varphi(z) \nu(X_t,dz) \quad (2.1)
\]

**Condition 2.1.** The F-convergence of solutions of (1.1) as \(\epsilon \downarrow 0\) holds under the following conditions.

1. \(F\) and \(G\) are Lipschitz continuous.

2. There exists a diffeomorphism \(\eta := (\eta^x,\eta^y)\), from \(\mathbb{R}^d\) onto \(\mathbb{R}^{d-p} \times \mathbb{R}^p\), independent of \(\epsilon\), with uniformly bounded \(C^1, C^2\) derivatives, and such that for all \(\epsilon > 0\), the process \((x^\epsilon_t, y^\epsilon_t) = (\eta^x(u^\epsilon_t), \eta^y(u^\epsilon_t))\) satisfies the ODE

\[
\begin{cases}
\dot{x}^\epsilon = g(x^\epsilon, y^\epsilon), & x^\epsilon_0 = x_0 \\
\dot{y}^\epsilon = \frac{1}{\epsilon} f(x^\epsilon, y^\epsilon), & y^\epsilon_0 = y_0
\end{cases} \quad (2.2)
\]

3. There exists a family of probability measures \(\mu(x,dy)\) on \(\mathbb{R}^p\) indexed by \(x \in \mathbb{R}^{d-p}\) and a positive function \(T \mapsto E(T)\) such that \(\lim_{T \to \infty} E(T) = 0\) and such that for all \(x_0, y_0, T\) and \(\phi\) uniformly bounded and Lipschitz, the solution to

\[
\dot{Y}_t = f(x_0, Y_t) \quad Y_0 = y_0 \quad (2.3)
\]

satisfies

\[
\left| \frac{1}{T} \int_0^T \phi(Y_s) - \int \phi(y) \mu(x_0,dy) \right| \leq \chi(\|x_0, y_0\|) E(T)(\|\phi\|_{L^\infty} + \|\nabla \phi\|_{L^\infty}) \quad (2.4)
\]

where \(r \mapsto \chi(r)\) is bounded on compact sets.

4. For all \(u_0, T > 0\), the trajectories \((u^\epsilon_t)_{0 \leq t \leq T}\) are uniformly bounded in \(\epsilon\).
For $\pi$ an arbitrary measure on $\mathbb{R}^d$, we define $\eta^{-1} \ast \pi$ to be the push forward of the measure $\pi$ by $\eta^{-1}$.

**Theorem 2.1.** Consider the system of stiff ODEs (1.1). Assume that Conditions 2.1 are satisfied then

- $u^\epsilon_t$ $F$-converges towards $\eta^{-1} \ast (\delta_{X_t} \otimes \mu(X_t, dy))$ as $\epsilon \downarrow 0$ where $X_t$ is the solution to

$$\dot{X}_t = \int g(X_t, y) \mu(X_t, dy) \quad X_0 = x_0$$

**Remark 2.1.** The $F$ convergence of $u^\epsilon_t$ towards $\eta^{-1} \ast (\delta_{X_t} \otimes \mu(X_t, dy))$ can be restated as

$$\lim_{h \to 0} \lim_{\epsilon \to 0} \frac{1}{h} \int_t^{t+h} \varphi(u^\epsilon_s) \, ds = \int_{\mathbb{R}^d} \varphi(\eta^{-1}(X_t, y)) \mu(X_t, dy)$$

for all functions $\varphi$ bounded and uniformly Lipschitz-continuous on $\mathbb{R}^d$, and for all $t > 0$.

**Proof.** The proof of Theorem 2.1 is classical and similar to that of Theorem 2.2 below and in the Appendix. We also refer to Chapter II, Section 3 of [82]. \(\square\)

### 2.2 FLOW AVeraging integratORS for stiff ODEs.

Assume that we are given two mappings $\theta^\epsilon$, $\theta^G_h$ from $\mathbb{R}^d$ onto $\mathbb{R}^d$, the former approximating the flow of the whole system (1.1) for time $\tau$, and the latter approximating the flow of $v_t = G(v)$ for time $h$. More precisely we assume that $\theta^\epsilon$ and $\theta^G_h$ satisfy the following conditions.

**Condition 2.2.** We will prove the $F$-convergence of FLAVORS for solutions of (1.1) under the following conditions on $\theta^G$ and $\theta^\epsilon$.

1. There exists $h_0, C > 0$ such that for $h \leq h_0$,

$$|\theta^G_h(u) - u - hG(u)| \leq Ch^2$$

2. There exists $\tau_0, C > 0$, such that for $\frac{\tau}{\epsilon} \leq \tau_0$

$$|\theta^\epsilon(\tau) - u - \tau G(u) - \frac{\tau}{\epsilon} F(u)| \leq C\left(\frac{\tau}{\epsilon}\right)^2$$

3. For all $u_0$, $T > 0$, the discrete trajectories $((\theta^G_{\delta+\tau} \circ \theta^\epsilon)^k(u_0))_{0 \leq k \leq T/\delta}$ are uniformly bounded in $\epsilon$, $0 < \delta \leq h_0$, $\tau \leq \min(\tau_0, \delta)$.

**Remark 2.2.** If one is given a micro-scale integrator $\theta^\epsilon$ for (1.1) in which $1/\epsilon$ is a parameter that can be controlled, then $\theta^G_h$ can in principle be obtained from $\theta^\epsilon$ by setting $1/\epsilon$ at a value close to zero. This observation allows one to use FLAVORS with an existing code without having to modify it (see Subsection 2.2.2 below).
2.2.1 FLAVORS

Let $\epsilon \ll \delta < h_0$ and $\tau \in (0, \delta)$, $\tau \leq \epsilon \tau_0$. We define FLAVORS to be algorithms simulating the discrete process $(\bar{u}_{k\delta})_{k \in \mathbb{N}^*}$, where

$$\bar{u}_{k\delta} := \left( \theta^{G}_{h_{k\tau}} \circ \theta^{\epsilon}_{\tau} \right)^k (u_0). \quad (2.9)$$

Define the process $t \mapsto \bar{u}_t$ by

$$\bar{u}_t = \left( \theta^{G}_{h_{k\tau}} \circ \theta^{\epsilon}_{\tau} \right)^k (u_0) \text{ for } k\delta \leq t < (k+1)\delta. \quad (2.10)$$

Theorem 2.2 and 2.1 show that $u_{\epsilon t}$ (the solution of the full system (1.1) and $\bar{u}_t$ have the same $F$-limit but the computational cost of simulating $\bar{u}_t$ is much smaller than that of simulating $u_{\epsilon t}$. In particular the accuracy of the algorithm has been obtained (Theorem 2.2) for $\tau \ll \epsilon \ll \delta \ll 1$ and $(\frac{\tau}{\epsilon})^2 \ll \delta \ll \frac{\tau}{\epsilon}$. The condition $\tau \ll \epsilon$ is needed for the accuracy of the integrator $\theta^{\epsilon}_{\tau}$ (equation (2.8)). The condition $\delta \ll \frac{\tau}{\epsilon}$ is needed for the averaging of the hidden slow dynamics with respect to the (hidden) fast variables. The condition $(\frac{\tau}{\epsilon})^2 \ll \delta$ is needed only if $\eta$ is not the identity diffeomorphism, it allows for the control of the error on the slow dynamics induced the inaccuracy of $\theta^{\epsilon}_{\tau}$ (right hand side of (2.8)).

2.2.2 Non intrusive FLAVORS

Assume that we are given a mapping $\Phi^\alpha_t$ from $\mathbb{R}^d$ onto $\mathbb{R}^d$ approximating the flow of (1.1) for $\alpha = 1/\epsilon$. If the parameter $\alpha$ can be controlled then $\Phi^\alpha_t$ can be used as a black box for accelerating the computation of solutions of (1.1). The main difference with equation free methods is that here the acceleration can be obtained without prior identification of the slow variables. Indeed assume that there exists a constant $h_0 > 0$ such that $\Phi^\alpha_t$ satisfies for all $t \leq h_0 \min(\frac{1}{\tau}, 1)$

$$|\Phi^\alpha_t(u) - u - tG(u) - \alpha tF(u)| \leq Ct^2(1 + \alpha)^2 \quad (2.11)$$

then $\theta^\epsilon_t := \Phi^{\frac{1}{\epsilon}}_t$ and $\theta^{G\epsilon}_h := \Phi^0_t$ satisfy Conditions (2.2) and FLAVOR can be defined as the algorithm simulating the process

$$\bar{u}_t = \left( \Phi^0_{h_{k\tau}} \circ \Phi^{\frac{1}{\epsilon}}_{\tau} \right)^k (u_0) \text{ for } k\delta \leq t < (k+1)\delta. \quad (2.12)$$

Rationale and mechanism behind FLAVORS  We will now explain the rationale and mechanism behind FLAVORS. Let us start by considering the case where $\eta$ is the identity diffeomorphism. Let $\varphi^\epsilon$ be the flow of (2.2) and $\varphi^g$ be the flow of (2.2) with $y^\epsilon$ frozen, i.e.

$$\varphi^g(x, y) = (\hat{x}_t, y) \text{ where } \hat{x}_t \text{ solves } \frac{d\hat{x}}{dt} = g(\hat{x}, y), \quad \hat{x}_0 = x \quad (2.13)$$
The main effect of FLAVORS is to average the flow of (2.2) with respect to fast degrees of freedom via splitting and re-synchronization. By splitting we refer to the substitution of the flow $\phi_\epsilon$ by composition of $\varphi^g_{\delta - \tau}$ and $\varphi^\epsilon_\tau$ and by re-synchronization we refer to the distinct time-steps $\delta$ and $\tau$ whose effects are to advance the internal clock of fast variables by $\tau$ every $\delta$ steps. By averaging we refer to the fact that FLAVORS approximates the flow $\phi_h$ by the flow

$$\varphi_h := \left( \varphi^g_{\frac{\delta - \tau}{\delta}} \circ \varphi^\epsilon_\tau \right)^M$$

(2.14)

where $h$ is a coarse time step resolving the “slow” time scale associated with $x^\epsilon$, $M$ is a positive integer corresponding to the number of “samples” used to average the flow ($\delta$ has to identified with $hM$) and $\tau$ is a fine time step resolving the “fast” time scale, of the order of $\epsilon$, and associated with $y^\epsilon$. In general, analytical formulae are not available for $\varphi^g$ and $\varphi^\epsilon$ and numerical approximations are used instead.

In the situation where $\eta$ is not the identity diffeomorphism, simulating $\bar{u}_{n\delta}$ defined in (2.9) is equivalent to simulating $(\bar{x}_{n\delta}, \bar{y}_{n\delta})$ defined as

$$(\psi^g_{\delta - \tau} \circ \psi^\epsilon_\tau)(x_0, y_0)$$

(2.15)

and

$$(\varphi^g_h)(x, y) := \eta \circ \theta^G_h \circ \eta^{-1}(x, y)$$

(2.16)

The regularity of $\eta$ then leads to the accuracy of $\psi^\epsilon$ in resolving (2.2) over time steps $\tau \leq \tau_0 \epsilon$ and that of $\psi^g_h$ in resolving (2.2) over time steps $h \leq h_0$.

2.3 Convergence results and error bounds for FLAVORS.

The following theorem and Theorem 2.1 show that $\bar{u}_t$ (FLAVORS) and $u^\epsilon_t$ (the solution of (1.1)) share the same limit in terms of $F$-convergence. More precisely Theorem 2.2 establishes the accuracy of FLAVORS under the conditions $\tau \ll \epsilon \ll \delta \ll 1$ and $(\frac{\tau}{\epsilon})^2 \ll \delta \ll \frac{\tau}{\epsilon}$.

**Theorem 2.2.** Let $\bar{u}_t$ be the process defined in (2.10) or (2.12) and $X_t$ be the solution of (2.5). Assume that Conditions 2.1 and 2.2 are satisfied. Then

- $\bar{u}_t$ $F$-converges towards $\eta^{-1} \ast (\delta X_t \otimes \mu(X_t, dy))$ for $\epsilon \leq \delta/(C \ln \delta)$, $\delta \downarrow 0$, $\frac{\tau}{\epsilon} \downarrow 0$, $\frac{\tau}{\delta} \downarrow 0$ and $(\frac{\tau}{\epsilon})^2 \downarrow 0$.

- There exists $C > 0$ such that for $\delta < h_0$ and $\tau < \tau_0 \epsilon$,

$$|x^\epsilon_t - \eta^\epsilon(\bar{u}_t)| \leq Ce^{Ct} \Psi_1(u_0, \epsilon, \delta, \tau)$$

(2.17)

and

$$\left| \frac{1}{T} \int_t^{t+T} \varphi(\bar{u}_s) \, ds - \int_{\mathbb{R}^K} \varphi(\eta^{-1}(X_t, y)) \mu(X_t, dy) \right| \leq \Psi_2(u_0, \epsilon, \delta, \tau, T, t)(\|\varphi\|_{L^\infty} + \|\nabla \varphi\|_{L^\infty})$$

(2.18)
where $\Psi_1$ and $\Psi_2$ are functions converging towards zero as $\epsilon \leq \delta/(C \ln \delta)$, $\delta \downarrow 0$, $\frac{\tau}{\epsilon} \downarrow 0$, $\delta \delta \downarrow 0$ and $(\frac{\tau}{\epsilon})^{2 \frac{1}{3}} \downarrow 0$.

**Remark 2.3.** For $\epsilon \leq \delta/(C \ln \delta)$ and $\delta \downarrow 0$, the following holds

$$\Psi_1(u_0, \epsilon, \delta, \tau) \leq \sqrt{\delta} + \left(\frac{\tau}{\epsilon}\right)^2 \frac{1}{\delta} + E\left(\frac{1}{C \ln \frac{1}{\delta}}\right) + \left(\frac{\delta \epsilon}{\tau}\right)^2 \frac{1}{\tau} + E\left(\frac{1}{C \ln \left(\frac{\delta \epsilon}{\tau} + \frac{\tau}{\epsilon}\right)^{-1}}\right)$$

(2.19)

and $\Psi_2$ satisfies a similar inequality.

**Remark 2.4.** Through this paper $C$ will refer to an appropriately large enough constant independent from $\epsilon, \delta, \tau$. To simplify the presentation of our results we use the same letter $C$ for expressions such as $2CeC$ instead of writing it as a new constant $C_1$ independent from $\epsilon, \delta, \tau$.

**Remark 2.5.** Choosing $\tau \sim \gamma \epsilon$ and $\delta \sim \gamma \frac{\tau}{\epsilon}$, where $\gamma$ is a small constant independent from $\epsilon$, $\delta, \tau$. To simplify the presentation of our results we use the same letter $C$ for expressions such as $2CeC$ instead of writing it as a new constant $C_1$ independent from $\epsilon, \delta, \tau$.

Choosing $\tau \sim \gamma \epsilon$ and $\delta \sim \gamma \frac{\tau}{\epsilon}$, where $\gamma$ is a small constant independent from $\epsilon$, Theorem 2.2 shows that the approximation error of FLAVOR is bounded by a function of $\gamma$ converging towards zero as $\gamma \downarrow 0$. If follows that the speed up is of the order of $\delta \tau \sim \gamma \epsilon$, i.e. scales like $\frac{1}{\epsilon}$ at fixed accuracy. In order to be able to compare FLAVOR with integrators resolving all the fine timesteps we have limited the speed up in the numerical experiments to $200 \times$ (but the latter can be arbitrary large as $\epsilon \downarrow 0$). For sufficiently small $\epsilon$ we observe that FLAVORS with microstep $\tau$ and mesostep $\delta$ overperforms its associated legacy integrator with the same microstep $\tau$ over large simulation times (we refer to section 3.6 on the Fermi-Pasta-Ulam problem). This phenomenon is caused by an error accumulation at each tick (microstep) of the clock of fast variables. Since FLAVORS (indirectly, i.e. without identifying fast variables) slow down the speed of this clock from $\frac{1}{\epsilon}$ to a value $\frac{\tau}{\delta} \sim \frac{\tau}{\epsilon}$ independent from $\epsilon$ this error doesn’t blow up as $\epsilon \downarrow 0$ (as opposed to an integrator resolving all fine timesteps). Because of this reason, if this error accumulation on fast variables is exponential, then the speed up at fixed accuracy does not scale like $\frac{1}{\epsilon}$ but like $e^{\frac{T}{\epsilon}}$ where $T$ is the total simulation time (a consequence of this phenomenon can be seen in Figure 10 (associated with the FPU problem) where Velocity Verlet fails to capture the $O(\epsilon^{-1})$ dynamic with a timestep $h = 10^{-5}$ dynamic whereas FLAVOR remains accurate with $\tau = 10^{-4}$ and $\delta = 2 \cdot 10^{-3}$).

**Proof.** We refer to Subsection 6.1 of the appendix for the proof of Theorem 2.2.

### 2.4 Natural FLAVORS generalization to generic stiff ODEs.

FLAVORS have a natural generalization to systems of the form

$$\dot{u}^{\alpha, \epsilon} = F(u^{\alpha, \epsilon}, \alpha, \epsilon) \quad (2.20)$$

where $u \mapsto F(u, \alpha, \epsilon)$ is Lipshitz continuous.

**Condition 2.3.** $F$-convergence of $u^{\frac{1}{\epsilon}, \epsilon}$ as $\epsilon \downarrow 0$ holds under the following conditions.

1. $\epsilon \mapsto F(u, \alpha, \epsilon)$ is uniformly continuous in the neighborhood of 0.
Theorem 2.3. Let (2.24). Assume that Conditions 2.3 and 2.4 are satisfied then

\[ x^α = g(x^α, y^α) \quad x_0^α = x_0 \tag{2.21} \]

where \( g(x,y) \) is Lipschitz continuous in \( x \) and \( y \) on bounded sets.

3. There exists a family of probability measures \( \mu(x, dy) \) on \( \mathbb{R}^p \) such that for all \( x_0, y_0, T \) \((x_0, y_0) := \eta(u_0)\) and \( \varphi \) uniformly bounded and Lipschitz

\[
\left| \frac{1}{T} \int_0^T \varphi(y^α_s) ds - \int \varphi(y) \mu(x_0, dy) \right| \leq \chi(\|x_0, y_0\|) \left( E_1(T) + E_2(Te^\alpha) \right) \| \nabla \varphi \|_{L^\infty}
\]

where \( r \mapsto \chi(r) \) is bounded on compact sets and \( E_2(r) \to 0 \) as \( r \to \infty \) and \( E_1(r) \to 0 \) as \( r \to 0 \).

4. For all \( u_0, T \), the trajectories \((u^α_0)_{0 \leq t \leq T}\) are uniformly bounded in \( \alpha \geq 1 \).

Assume that we are given a mapping \( \Phi^α_{t, \nu} \) from \( \mathbb{R}^d \) onto \( \mathbb{R}^d \) approximating the flow of (2.20). If the parameter \( \alpha \) can be controlled then \( \Phi^α_{t, \nu} \) can be used as a black box for accelerating the computation of solutions of (2.20). The acceleration is obtained without prior identification of the slow variables.

Condition 2.4. We will prove the \( F \)-convergence of FLAVORS for (2.20) under the following conditions on \( \Phi^α_{t, \nu} \).

1. There exists a constant \( h_0 > 0 \) such that \( \Phi^α_{t, \nu} \) satisfies for all \( t \leq h_0 \min(\frac{1}{\alpha^2}, 1) \),
\[ 0 < \epsilon \leq 1 \leq \alpha \]
\[ |\Phi^α_{t, \nu}(u) - u - tF(u, \alpha, \epsilon)| \leq C(u)t^2(1 + \alpha^{2\nu}) \tag{2.23} \]

where \( C(u) \) is bounded on compact sets.

2. For all \( u_0, T \), the discrete trajectories \( \left( (\Phi^0_{t, \nu} \circ \Phi^1_{t+\delta})^k(u_0) \right)_{0 \leq k \leq T/\delta} \) are uniformly bounded in \( 0 < \epsilon \leq 1 \), \( 0 < \delta \leq h_0, \epsilon \leq \min(h_0 e^\nu, \delta) \).

FLAVOR can be defined as the algorithm given by the process

\[
\bar{u}_t = (\Phi^0_{t, \nu} \circ \Phi^1_{t+\delta})^k(u_0) \quad \text{for} \quad k\delta \leq t < (k + 1)\delta \tag{2.24}
\]

The theorem below shows the accuracy of FLAVORS for \( \delta \ll h_0 \), \( \tau \ll \epsilon^\nu \) and \( (\frac{\tau}{\epsilon})^2 \ll \delta \ll \frac{\tau}{\epsilon^\nu} \).

**Theorem 2.3.** Let \( u^1_{t, \epsilon} \) be the solution to (2.20) with \( \alpha = 1/\epsilon \) and \( \bar{u}_t \) be defined by (2.24). Assume that Conditions 2.3 and 2.4 are satisfied then
• \( u^{\frac{1}{2},\epsilon} \) converges towards \( \eta^{-1} \ast (\delta_{X_t} \otimes \mu(X_t,dy)) \) as \( \epsilon \downarrow 0 \) where \( X_t \) is the solution to
\[
\dot{X}_t = \int g(X_t,y) \mu(X_t,dy) \quad X_0 = x_0 \tag{2.25}
\]

• As \( \epsilon \downarrow 0 \), \( \tau \epsilon^{-\nu} \downarrow 0 \), \( \delta \epsilon^{\nu} \downarrow 0 \), \( \frac{x^2}{\epsilon^{2\nu}} \downarrow 0 \), \( \bar{u}_t \) converges towards \( \eta^{-1} \ast (\delta_{X_t} \otimes \mu(X_t,dy)) \) as \( \epsilon \downarrow 0 \) where \( X_t \) is the solution of \( \epsilon \) iterations of \( \Phi \).

**Proof.** The proof of Theorem 2.3 is similar to that of Theorem 2.2 and 4.2 given below. Only the idea of the proof will be given here. The condition \( \epsilon \ll 1 \) is needed for the approximation of \( u^{\alpha,\epsilon} \) by \( u^{\alpha,0} \) and for the F-convergence of \( u^{\frac{1}{2},\epsilon} \). Since \( y_t^n = \eta^n(u_t^{\alpha,0}) \) the condition \( \tau \ll \epsilon' \) is used along with equation (2.23) for the accuracy of \( \Phi \) in (locally) approximating \( y_t \). The condition \( \delta \ll \frac{1}{\epsilon'} \) allows for the averaging of \( g \) to take place prior to a significant change of \( x_t^{\alpha} \), more precisely it allows for \( m \gg 1 \) iterations of \( \Phi \) prior to a significant change of \( x_t^{\alpha} \). The condition \( \left( \frac{\epsilon}{\epsilon'} \right)^2 \ll \delta \) is required in order to control the error accumulated by \( m \) iterations of \( \Phi \).

\[\square\]

### 2.5 Numerical experiments: systems of ODEs with hidden slow dynamics.

Consider the following system ODEs
\[
\begin{align*}
\dot{r} &= \frac{1}{\epsilon} (r \cos \theta + r \sin \theta - \frac{1}{\epsilon} r^3 \cos^3 \theta) \cos \theta - \epsilon r \cos \theta \sin \theta \\
\dot{\theta} &= -\epsilon \cos^2 \theta - \frac{1}{\epsilon} (\cos \theta + \sin \theta - \frac{1}{3} r^2 \cos^3 \theta) \sin \theta
\end{align*}
\tag{2.26}
\]

where \( \epsilon \ll 1 \). Taking the transformation from polar coordinates to Cartesian coordinates by \( [x,y] = [r \cos \theta, r \sin \theta] \) as the local diffeomorphism, we obtained the hidden system:
\[
\begin{align*}
\dot{x} &= \frac{1}{\epsilon} (y + x - \frac{1}{3} x^3) \\
\dot{y} &= -\epsilon x
\end{align*}
\tag{2.27}
\]

Taking the second time derivative of \( x \), the system can also be written as the 2\textsuperscript{nd}-order ODE:
\[
\ddot{x} + x = \frac{1}{\epsilon} (1 - x^2) \dot{x}
\tag{2.28}
\]

The latter is a classical Van der Pol oscillator [91]. Non intrusive FLAVOR as defined by (2.24) can be directly applied to (2.26) (with hidden slow and fast processes) by turning on and off the stiff parameter \( \frac{1}{\epsilon} \). More precisely, defining \( \Phi^{\alpha,\epsilon}(r,\theta) \) by
\[
\Phi^{\alpha,\epsilon}_h(r,\theta) := \left( \begin{array}{c} r \\ \theta \end{array} \right) + \alpha h \left( \begin{array}{c} (r \cos \theta + r \sin \theta - \frac{1}{3} r^3 \cos^3 \theta) \cos \theta \\ -(\cos \theta + \sin \theta - \frac{1}{3} r^2 \cos^3 \theta) \sin \theta \end{array} \right) - \epsilon h \left( \begin{array}{c} r \cos \theta \sin \theta \\ \cos^2 \theta \end{array} \right)
\tag{2.29}
\]

FLAVOR is defined by (2.24) with \( \bar{u} := (\bar{r}, \bar{\theta}) \), i.e.
\[
(\bar{r}_t, \bar{\theta}_t) = \left( \Phi^{0,\epsilon}_{\bar{r},\tau} \circ \Phi^\frac{1}{2,\epsilon}_\tau \right)^k (r_0, \theta_0) \quad \text{for} \quad k\delta \leq t < (k + 1)\delta
\tag{2.30}
\]

14
We refer to Figure 1 for a comparison of integrations by Forward Euler used as a benchmark, and FLAVORS. FLAVORS gives trajectories close to Forward Euler and correctly captures the $O(\frac{1}{\epsilon})$ period \[91\] of the relaxation oscillation. Moreover, a 200 fold acceleration is achieved using FLAVOR.

![Figure 1: Over a timespan of $5/\epsilon$](image)

(a) Direct Forward Euler simulation of (2.27) with time steps resolving the fast time scale (b) (non intrusive (2.24)) FLAVOR simulation of (2.27) (c) Polar to cartesian image of the (non intrusive (2.24)) FLAVOR simulation of (2.26) with hidden slow and fast variables. Forward Euler uses timestep $h = 0.05 \epsilon = 0.00005$. The two FLAVORS simulations use $\delta = 0.01$ and $\tau = 0.00005$. Parameters are $\frac{1}{\epsilon} = 1000$, $x(0) = 1$, $y(0) = 1$.

3 Deterministic Hamiltonian Systems.

3.1 Symplectic, symmetric (under a group action) and time-reversible FLAVORS for mechanical systems on manifolds.

In this section we will apply FLAVORS to ODEs of the form

$$\dot{p} = -\partial_q H(p, q) \quad \dot{q} = \partial_p H(p, q)$$

where the Hamiltonian

$$H(q, p) := \frac{1}{2} p^T M^{-1} p + V(q) + \frac{1}{\epsilon} U(q)$$
represents the total energy of a mechanical system with Euclidean phase space $\mathbb{R}^d \times \mathbb{R}^d$ or a cotangent bundle $T^*M$ of a configuration manifold $M$.

We define $\Theta_\delta$ the FLAVOR discrete mapping approximating solutions of (3.1) over time steps $\delta \gg \epsilon$ by

$$\Theta_\delta(q_{n+1}\delta, p_{n+1}\delta) := \Theta_\delta(q_{n\delta}, p_{n\delta})$$

(3.3)

It is known that numerical methods that preserve the geometric properties of the flow of a differential equation are characterized by improved long time behavior and respect for conserved quantities[44]. The continuous flow of (3.1) with a Hamiltonian of the form (3.2) is known to be both symplectic and time-reversible. We will show below how a symplectic integrator resolving the fast time scale $\tau \ll \epsilon$ of (3.2) can be used as black box integrator and turned into a FLAVOR integrator with slow intrinsic time scale $\delta \gg \epsilon$ by simply turning on and off the parameter $\frac{1}{\epsilon}$ in the integrator. Similarly a symplectic and reversible integrator can be turned in a non-intrusive way into a symplectic, symmetric (under a group action), and time-reversible FLAVOR integrator.

First assume that we are given a mapping $\Phi^\alpha_t$ acting on the phase space such that for $t \leq h_0 \min(1, \alpha^{-\frac{1}{2}})$

$$\left| \Phi^\alpha_t(q, p) - (q, p) - t(M^{-1}p, -V(q) - \alpha U(q)) \right| \leq Ct^2(1 + \alpha)$$

(3.4)

then

$$\Theta_\delta := \Phi^0_{h_0 \tau^{-1}} \circ \Phi^\frac{1}{\tau}_{\delta}$$

(3.5)

defines a FLAVOR integrator for (3.1). Theorem 2.3 establishes the accuracy of this integrator under Conditions 2.3 and 2.4 provided that $\tau \ll \sqrt{\epsilon} \ll \delta$ and $\frac{\tau^2}{\epsilon} \ll \delta \ll \frac{\tau}{\sqrt{\epsilon}}$.

The following theorem, whose proof is straightforward, shows that $\Theta_\delta$ inherits the symmetries of the legacy integrator $\Phi^\epsilon_h$.

**Theorem 3.1.** If for all $h, \epsilon > 0$ $\Phi^\epsilon_h$ is symmetric under a group action then $\Theta_\delta$ is symmetric under the same group action.

Write

$$\Phi^\ast_h := (\Phi_{-h})^{-1}$$

(3.6)

Let us recall the following definition corresponding to definition 1.4 of the Chapter V of [45]

**Definition 3.1.** A numerical one-step method $\Phi_h$ is called time-reversible if it satisfies $\Phi^\ast_h = \Phi_h$.

The following theorem shows how derive a symplectic FLAVOR or a “symplectic and symmetric and time-reversible” FLAVOR from a symplectic integrator for (3.2). Since this derivation applies to manifolds it also leads to FLAVORS for constrained mechanical systems that are symplectic and time-reversible.
Theorem 3.2. If $\Phi^\alpha_t$ is symplectic on the tangent bundle $T^*M$ of a configuration manifold $M$ then $\Theta_\delta$ defined by (3.5) is symplectic on the tangent bundle $T^*M$. Moreover

$$\Theta_\delta := \Phi^{1/2\epsilon}_t \circ \Phi^{0}_{-\epsilon/2} \circ \Phi^{0}_{-\epsilon/2} \circ \Phi^{1/2}_{-\epsilon}$$

is time-reversible. Furthermore if, for all $h, \epsilon > 0$, $\Phi^1_h$ is symmetric under a group action then $\Theta_\delta$ is symmetric under the same group action.

Proof. The proof is straightforward. We also refer to [45].

Remark 3.1. Observe that (except for the first and last steps) iterating $\Theta_\delta$ defined by (3.7) is equivalent to iterating

$$\Theta_\delta := \Phi^{0}_{-\epsilon/2} \circ \Phi^{0}_{-\epsilon/2} \circ \Phi^{1/2}_{-\epsilon} \circ \Phi^{1/2}_{-\epsilon}$$

(3.8)

It follows that a symplectic, symmetric and reversible FLAVOR can be obtained in a non-intrusive way from a Störmer/Verlet integrator for (3.1) [44, 46, 92].

3.1.1 An example of symplectic FLAVOR

If the phase space is $\mathbb{R}^d \times \mathbb{R}^d$ then an example of symplectic FLAVOR is obtained from Theorem 3.2 by choosing $\Phi^\alpha_t$ to be the symplectic Euler (also known as Variational Euler or VE for short) integrator defined by

$$\Phi^\alpha_t(q,p) = \begin{pmatrix} q \\ p \end{pmatrix} + t \begin{pmatrix} M^{-1}(p - t(V(q) + \alpha U(q))) \\ -V(q) - \alpha U(q) \end{pmatrix}$$

(3.9)

and letting $\Theta_\delta$ be defined by (3.5).

3.1.2 An example of symplectic and time-reversible FLAVOR.

If the phase space is the Euclidean space $\mathbb{R}^d \times \mathbb{R}^d$ then an example of symplectic and time-reversible FLAVOR is obtained by letting $\Theta_\delta$ be defined by equation (3.7) of Theorem 3.2 by choosing $\Phi^\alpha_t$ to be the symplectic Euler integrator defined by (3.9) and

$$\Phi^{\alpha, \epsilon}_t(q,p) = \begin{pmatrix} q \\ p \end{pmatrix} + t \begin{pmatrix} M^{-1}p \\ -V(q + tM^{-1}p) - \alpha U(q + tM^{-1}p) \end{pmatrix}$$

(3.10)

3.1.3 An artificial FLAVOR.

There is not a unique way of averaging the flow of (3.2). We present below an alternative method based on the freezing and unfreezing of degrees of freedom associated with fast potentials. We have called this method “artificial” because it is intrusive. With this method the discrete flow approximating solutions of (3.1) is given by (3.3) with

$$\Theta_\delta := \theta^{tr}_{\delta - \epsilon} \circ \theta^c_{\epsilon} \circ \theta^V_{\delta}$$

(3.11)
where \( \theta^V_\delta \) is a symplectic map corresponding to the flow of \( H^{\text{slow}}(q, p) := V(q) \), approximating the effects of the soft potential on momentum over the mesoscopic time step \( \delta \) and defined by

\[
\theta^V_\delta(q, p) = (q, p - \delta \nabla V(q)).
\]  

(3.12)

\( \theta^\tau \) is a symplectic map approximating the flow of \( H^{\text{fast}}(q, p) := \frac{1}{2} p^T M^{-1} p + \frac{1}{\epsilon} U(q) \) over a microscopic time step \( \tau \):

\[
\theta^\tau(q, p) = (q + \tau M^{-1} p, p - \frac{\tau}{\epsilon} \nabla U(q + tM^{-1} p)).
\]  

(3.13)

\( \theta^{tr}_{\delta - \tau} \) is a map approximating the flow of the Hamiltonian \( H^{\text{free}}(q, p) := \frac{1}{2} p^T M^{-1} p \) under holonomic constraints imposing the freezing of stiff variables. Velocities along the direction of constraints have to be stored and set to be 0 before the constrained dynamics, i.e. freezed, and the stored velocities should be restored after the constrained dynamics, i.e. un-freezed; geometrically speaking one projects to the constrained subsymplectic manifold, runs the constrained dynamics, and lifts back to the original full space. Oftentimes the exact solution to the constrained dynamics can be found (examples given in Subsection 3.3, 3.4, 3.5, 3.6 and 3.7).

When the exact solution to the constrained dynamics can not be easily found, one may want to employ integrators for constrained dynamics such as SHAKE [74] or RATTLE [4] instead. This has to be done with caution, because symplecticity of the translational flow may be lost. The composition of projection onto the constrained manifold (freezing), evolution on the constrained manifold, and lifting from it to the unconstrained space (unfreezing) preserves symplecticity in the unconstrained space only if the evolution on the constrained manifold preserves the inherited symplectic form. A numerical integration preserves the discrete symplectic form on the constrained manifold, but not necessarily the projected continuous symplectic form.

**Remark 3.2.** This artificial FLAVOR is locally a perturbation of natural FLAVORS. By splitting theory \([60, 45]\),

\[
\theta_{\delta - \tau}^{tr} \circ \theta_\tau \circ \theta^V_\delta \approx \theta_{\delta - \tau}^{tr} \circ \theta^V_\delta \circ \theta_\tau \circ \theta^V_\delta \approx \theta_{\delta - \tau}^{tr} \circ \theta^V_\delta \circ \Phi^1_\tau
\]  

(3.14)

whereas \( \Phi_0^{\delta - \tau} \circ \Phi^1_\tau \approx \theta^{free}_{\delta - \tau} \circ \theta^V_\delta \circ \Phi^1_\tau \), where \( \theta^{free} \) is the flow of \( H^{\text{free}}(q, p) \) under no constraint. The only difference is that constraints are treated in \( \theta^{tr} \) but not in \( \theta^{free} \).

The advantage of this artificial FLAVOR lies in the fact that only \( \tau \ll \sqrt{\epsilon} \ll \delta \) and \( \delta \ll \sqrt{\epsilon} \) are required for its accuracy (and not \( \frac{\tau^2}{\epsilon} \ll \delta \)). We also observe that, in general, artificial FLAVOR overperforms non-intrusive FLAVOR in FPU long time \((O(\omega^2))\) simulations (we refer to Subsection 3.6).

### 3.2 Stability

We will consider the following test problem (see Figure 2(a))

\[
H(x, y, px, py) = \frac{1}{2} px^2 + \frac{1}{2} py^2 + \frac{1}{2} \omega^2 x^2 + \frac{1}{2} (y - x)^2
\]  

(3.15)
Theorem 3.3. The FLAVOR (3.5) derived the symplectic Euler integrator (3.9) with \(1/\tau \gg \omega \gg 1\) is stable for \(\delta \in (0, \sqrt{2})\).

Remark 3.3. Recall that Variational Euler is stable if and only if its timestep satisfies \(h \leq 2/\omega\).

Proof. (Theorem 3.3) The FLAVOR (3.5) derived from the symplectic Euler integrator (3.9) can be written as

\[
\begin{bmatrix}
    x_{n+1} \\
    y_{n+1} \\
    (p_x)_{n+1} \\
    (p_y)_{n+1}
\end{bmatrix} = T \begin{bmatrix}
    x_n \\
    y_n \\
    (p_x)_n \\
    (p_y)_n
\end{bmatrix}, \text{ where } T =
\begin{bmatrix}
    1 & 0 & \delta - \tau & 0 \\
    1 & 1 & 0 & 0 \\
    0 & 1 & \tau - \delta & \delta - \tau \\
    0 & 0 & 1 & 0 \\
    0 & 0 & \tau - \delta & \tau - \delta \\
    0 & 0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
    1 & 0 & \tau & 0 \\
    0 & 1 & 0 & 0 \\
    0 & 0 & 1 & 0 \\
    0 & 0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
    1 & 0 & 0 & 0 \\
    0 & 1 & 0 & \tau \\
    -\tau (\omega^2 + 1) & \tau & 0 & 0 \\
    \tau & -\tau & 0 & 1
\end{bmatrix}
\]

The characteristic polynomial of T is

\[
\lambda^4 + \left(-4 + 2\delta^2 - 4\delta^2 \tau^2 + 8\delta^3 \tau^3 - 4\tau^4 - \delta \tau \omega^2 - \delta^2 \tau^2 \omega^2 + 2\delta^3 \tau^3 \omega^2 - \tau^4 \omega^2\right)\lambda^3 + \left(6 - 4\delta^2 + 8\delta^2 \tau^2 - 16\delta^3 \tau^3 + 8\tau^4 - 2\delta \tau \omega^2 + \delta^3 \tau^3 \omega^2 + 2\delta^2 \tau^4 \omega^2 - 4\delta^3 \tau^3 \omega^2 - 2\delta^3 \tau^3 \omega^2 - 2\tau^4 \omega^2\right)\lambda^2
\]

\[
+ \left(4\delta^2 \tau^4 \omega^2 - 2\delta \tau^5 \omega^2\right)\lambda
\]

\[
+ \left(-4 + 2\delta^2 - 4\delta^2 \tau^2 + 8\delta^3 \tau^3 - 4\tau^4 + \delta \tau \omega^2 - \delta^2 \tau^2 \omega^2 + 2\delta^3 \tau^3 \omega^2 - \tau^4 \omega^2\right)\lambda + 1
\]

(3.17)

Since \(\omega \gg 1\), \(\tau \ll 1/\omega\), as long as \(\delta \lesssim 1\) roots to the above polynomial are close to roots to the asymptotic polynomial

\[
\lambda^4 + (2\delta^2 - 4)\lambda^3 + (6 - 4\delta^2)\lambda^2 + (2\delta^2 - 4)\lambda + 1
\]

(3.18)

which can be shown to be 1 with multiplicity 2 and \((1 - \delta^2 \pm \delta \sqrt{\delta^2 - 2})\).
It is easy to see that if and only if $|\delta| \leq \sqrt{2}$ all roots are complex numbers with moduli less or equal to 1, and hence FLAVORS is stable (for more details see for instance [85]).

**Theorem 3.4.** The artificial FLAVOR described in Subsection 3.1.3 with $1/\tau \gg \omega \gg 1$ is stable for $\delta \in (0, 2)$. 

**Proof.** (Theorem 3.4) The FLAVOR with the translational flow derived from the symplectic Euler integrator (3.9) can be written as

\[
\begin{bmatrix}
  x_{n+1} \\
  y_{n+1} \\
  (p_x)_{n+1} \\
  (p_y)_{n+1}
\end{bmatrix} = T \begin{bmatrix}
  x_n \\
  y_n \\
  (p_x)_n \\
  (p_y)_n
\end{bmatrix},
\]

where

\[
T = \begin{bmatrix}
  1 & 0 & 0 & 0 & 0 & 0 & 0 \\
  0 & 1 & 0 & \delta - \tau & -\tau \omega^2 & 0 & 0 \\
  0 & 1 & 0 & 0 & 0 & 1 & 0 \\
  0 & 0 & 1 & 0 & 0 & 0 & 1 \\
  0 & 0 & 0 & 1 & 0 & 0 & 0
\end{bmatrix}
\]

The characteristic polynomial of $T$ is

\[
\lambda^4 + (\omega^2 \tau^2 + \tau \delta + \delta^2 - 4)\lambda^3 + (6 - 2\delta^2 - 2\delta \tau - 2\tau^2 \omega^2 + \delta^2 \tau^2 \omega^2)\lambda^2 + (\omega^2 \tau^2 + \tau \delta + \delta^2 - 4)\lambda + 1
\]

Since $\omega \gg 1$, $\tau \ll 1/\omega$, as long as $\delta \lesssim 1$ roots to the above polynomial are close to roots to the asymptotic polynomial

\[
\lambda^4 + (\delta^2 - 4)\lambda^3 + (6 - 2\delta^2)\lambda^2 + (\delta^2 - 4)\lambda + 1
\]

which can be shown to be 1 with multiplicity 2 and $\frac{1}{2}(2 - \delta^2 \pm \delta \sqrt{\delta^2 - 4})$.

It is easy to see that if and only if $|\delta| \leq 2$ all roots are complex numbers with moduli less or equal to 1, and hence FLAVORS is stable (for more details see for instance [85]).

### 3.3 Numerical error analysis of FLAVOR for nonlinear systems

In this subsection we will consider the example illustrated in Figure 2(b) associated with the Hamiltonian

\[
H(x, y, z, p_x, p_y, p_z) = \frac{1}{2}p_x^2 + \frac{1}{2}p_y^2 + \frac{1}{2}p_z^2 + x^4 + \epsilon^{-1} \frac{\omega_1}{2}(y - x)^2 + \epsilon^{-1} \frac{\omega_2}{2}(z - y)^2
\]

Thus, the potential is $U = \frac{1}{2}(y - x)^2 + \frac{\omega_2}{2}(z - y)^2$ and $V = x^4$. Here $\frac{x+y+z}{3}$ acts as a slow degree of freedom and $y - x$ and $z - y$ act as fast degrees of freedom. Figure 3 illustrates $t \mapsto \frac{x(t) + y(t) + z(t)}{3}$ (slow variable) computed with symplectic Euler and with the induced symplectic FLAVOR ((3.5) and (3.9)). Define $q := (x, y, z)$. To illustrate the $F$-convergence property of FLAVOR, we fix $h$ and vary the mesostep $\delta = h/M$ by changing
Figure 3: Comparison between trajectories integrated by Variational Euler and FLAVOR (defined by (3.5) and (3.9)). FLAVOR uses mesostep $\delta = 0.01$ and microstep $\tau = 0.0005$, and Symplectic Euler uses timestep $\tau = 0.0005$. FLAVOR accelerated the computation by roughly 20-fold ($\delta = 20\tau$). In this experiment $\epsilon = 10^{-6}$, $\omega_1 = 1.1$, $\omega_2 = 0.97$, $x(0) = 0.8$, $y(0) = 0.811$, $z(0) = 0.721$, $p_x(0) = 0$, $p_y(0) = 0$ and $p_z(0) = 0$. Simulation time $T = 50$.

$M$ and show the Euclidean norm error of the difference between $\frac{1}{M} \sum_{i=0}^{M-1} q(T - ih/M)$ computed with FLAVOR and computed with symplectic Euler in Figure 4(a). Notice without an averaging over time length $h$ the error will be no longer monotonically but oscillatorily decreasing as $\delta$ changes (see for example Figure 5(d)), because fast variables are captured only in the sense of measure. As shown in Figure 4(a) the error scales linearly with $\frac{1}{M}$ for $M$ not too small. Figure 4(b) shows that the error in general grows linearly with the total simulation time, and this linear growth of the error has been observed for a simulation time larger than $\omega (\epsilon^{-1/2})$. Figure 4(c) shows that the error doesn’t depend on $\omega (\epsilon^{-1/2})$ for a fixed $\delta$, as long as $\epsilon$ is not too large (i.e., $\omega$ not too small) is not too large. These results together suggest that FLAVORS is a first order method with error independent from $\omega$ and hence a homogenization method. The fact that the error scales linearly with total simulation time is a stronger result than our error analysis for FLAVORS (in which the error is bounded by a term growing exponentially with the total simulation time). We believe that the linear growth of the error is a consequence of the fact that FLAVOR is symplectic. A rigorous analysis of the effects of the structure preservation of FLAVORS on long term behavior remains to be done.
3.4 Numerical error analysis of FLAVOR for linear systems

Consider a simplified, linear version of (3.22) given by the Hamiltonian

$$H(x, y, p_x, p_y) = \frac{1}{2} p_x^2 + \frac{1}{2} p_y^2 + \frac{1}{2} x^2 + \frac{1}{2} \omega^2 (y - x)^2$$  (3.23)

where $\frac{x+y}{2}$ is the slow variable. The flow of (3.23) has been explicitly obtained and compared with solutions obtained from non-intrusive FLAVOR based on symplectic Euler ((3.5) and (3.9)) and with artificial FLAVOR (3.11).

The total simulation time is $T = 0$, and absolute errors have been computed with respect to the Euclidean norm of the difference in positions between analytical and numerical solutions. Stabilities have been investigated using the same technique used in Section 3.2. We have chosen a linear system because of the availability of explicit solutions.

Figure 5(a) and 5(b) illustrate errors as functions of mesostep $\delta$ and renormalized small step $\tau/\epsilon$. We observe that given $\delta$ errors are minimized at specific values of $\tau/\epsilon$ for both integrators, but the accuracy of non-intrusive FLAVOR is less sensitive to $\tau/\epsilon$.

Figure 5(c) and 5(d) plot the optimal value of $\tau/\epsilon$ as a function of $\delta$ and the associated error. We observe that for non-intrusive FLAVOR the dependence of the optimal value of $\tau/\epsilon$ on $\delta$ is weak, whereas for artificial FLAVOR the optimal value of $\tau/\epsilon$ roughly scales linearly with $\delta$. Figure 5(e) has to be put in correspondence with the condition $\delta << \tau/\epsilon$ required for accuracy. The weak dependence of the error on $\tau/\epsilon$ for a fix $\delta$ show that one doesn’t have to choose the microstep with too much care or optimize the integrator with respect to its value. As a matter of fact, all the numerical experiments illustrated in this paper (besides Figures 5(c) and 5(d)) have been performed without any tuning of the value $\tau/\epsilon$. We have simply and only used the rule of thumb $\delta \sim \gamma \frac{\tau}{\epsilon}$ where $\gamma$ is a small parameter (0.1 for instance).

Figure 6(a) and 6(b) illustrate the domain of stability of non-intrusive FLAVOR and artificial FLAVOR, i.e. values of $\delta$ and $\tau/\epsilon$ ensuring stable numerical integrations. We observe that artificial FLAVOR has a much larger stability domain than non-intrusive FLAVOR. Furthermore, for non-intrusive FLAVOR and large values of $\delta$, $\tau/\epsilon$ has to be...
Figure 5: Error analysis of (3.23). Parameters are $\omega = 10^3$, $x(0) = 0.8$ and $y(0) = x(0) + 1.1/\omega$. 

(a) Error of non-intrusive FLAVOR as a function of $\delta$ and $\tau/\epsilon$

(b) Error of artificial FLAVOR as a function of $\delta$ and $\tau/\epsilon$

(c) Optimal $\tau/\epsilon$ and error of non-intrusive FLAVOR as functions of $\delta$

(d) Optimal $\tau/\epsilon$ and error of artificial FLAVOR as functions of $\delta$

(e) Error dependence on $\tau/\epsilon$ for a given $\delta$: non-intrusive FLAVOR

(f) Error dependence on $\tau/\epsilon$ for a given $\delta$: artificial FLAVOR
smaller than the step length that guarantees a stable symplectic Euler integration of the fast dynamics.

Therefore, it appears that the benefits of artificial FLAVORS lie in their superior accuracy and stability.

3.5 Nonlinear stiff and soft potentials.

In this subsection we will apply the Symplectic Euler FLAVOR defined by (3.5) and (3.9) to the mechanical system illustrated in Figure 2(a) whose Hamiltonian is

$$H(x, y, p_x, p_y) := \frac{1}{2} p_x^2 + \frac{1}{2} p_y^2 + \epsilon^{-1} x^6 + (y - x)^4$$  \hfill (3.24)

The left spring is stiff, the right spring is soft and both non-harmonic. $x$ is the distance between wall and the left mass, $y$ is the distance between wall and the right mass. $p_x$ and $p_y$ are their momentums. Here $U = x^6$ and $V = (y - x)^4$.

Hence $x$ is a fast variable and $y - x$ is the combination of a slow variable and a fast variable. Figure 7 illustrates $t \mapsto x(t)$ (fast variable), $t \mapsto y(t) - x(t)$, and $t \mapsto H(t)$ computed with: Symplectic Euler, the induced symplectic FLAVOR ((3.5) and (3.9)), and IMEX [83].

3.6 Fermi-Pasta-Ulam problem

In this subsection we will consider the Fermi-Pasta-Ulam (FPU) Problem [34] illustrated by Figure 8 and associated with the Hamiltonian

$$H(q, p) := \frac{1}{2} \sum_{i=1}^{m} (p_{2i-1}^2 + p_{2i}^2) + \frac{\omega^2}{4} \sum_{i=1}^{m} (q_{2i} - q_{2i-1})^2 + \sum_{i=0}^{m} (q_{2i+1} - q_{2i})^4$$  \hfill (3.25)

The FPU problem is a well known benchmark problem [61] [45] for multiscale integrators because it exhibits different behaviors over widely separated timescales. The stiff
Figure 7: In this experiment $\epsilon = 10^{-6}$, $x(0) = 1.1$, $y(0) = 2.2$, $p_x(0) = 0$ and $p_y(0) = 0$. Simulation time $T = 2$. FLAVOR (defined by (3.5) and (3.9)) uses mesostep $\delta = 10^{-3}$ and microstep $\tau = 10^{-5}$. Variational Euler uses small timestep $\tau = 10^{-5}$, and IMEX uses mesostep $\delta = 10^{-3}$. Since the fast potential is nonlinear, IMEX is an implicit method and nonlinear equations have to be solved at every step, and IMEX turns out to be slower than Variational Euler. FLAVOR is strongly accurate with respect to slow variables and accurate in the sense of measures with respect to fast variables. Comparing to Symplectic Euler, FLAVOR accelerated the computation by roughly 100-fold.

springs (nearly) behave like harmonic oscillator with period $\sim O(\omega^{-1})$. Then the centres of masses linked by stiff springs (i.e. the middle points of stiff springs) change over a timescale $O(1)$. The third timescale $O(\omega)$ is associated with the rate of energy exchange between stiff springs. The fourth timescale $O(\omega^2)$ corresponds to the synchronization of energy exchange between stiff springs. As the number of stiff springs increases, the interactive energy exchange pattern becomes more and more complicated, and this pattern extends to $O(\omega^\infty)$ timescale when $m \to \infty$. On the other hand, the total energy of stiff springs behave almost like a constant. This wide separation of timescales can be seen in Figure 9, 10, and 12, where 4 subplots address different scales: Subplot1 shows scaled expansions of three stiff springs $(q_2 - q_{2i-1})/\sqrt{2}$, which are fast variables; Subplot2 shows scaled middle point position of the first stiff spring $(q_2 + q_1)/\sqrt{2}$, which is one of the slow variables; Subplot3 shows the energy transferring pattern among stiff springs, which is even slower; Subplot4 shows the near-constant total energy of three stiff springs. All 4 subplots are time-series. A comprehensive survey on FPU problem, including discussions on timescales and numerical recipes, can be found in [45].

Figures 9(b) and 9(a) compare symplectic Euler (with time steps fine enough to resolve FPU over the involved long time scale) and with the artificial FLAVOR (3.11). On
a timescale $O(\omega)$ FLAVOR captured slow variable’s periodic behavior with the correct period and phase, as well as the slower process of energy transferring. In this simulation, the time-span is much longer than $O(1)$ which is commonly considered in numerical error analysis. At the same time, FLAVOR accelerated the computation by roughly 40-fold (since $\delta = 40\tau'$).

Artificial FLAVOR’s thicker curves in energy transferring pattern do not affect the global pattern and are caused by the numerical error associated with microstep $\tau$. This can be inferred by using the artificial FLAVOR introduced in Subsection 3.1.3 with $\theta'_{\tau}$ corresponding to the exact flow of $H_{fast}$ (rather than its Variational Euler approximation: this specific artificial Euler resembles Impulse Method, but Impulse Method will yield unbounded trajectories if one runs even longer time simulations, whereas FLAVORS don’t seem to have an error growing exponentially with the total simulation time). See Figure 11 for obtained thin energy curves.

Now, we integrate over a longer total simulation time of the order of $\omega^2$ to investigate different integrators’ performances in capturing the synchronized energy exchanging pattern (Figure 10).

There is a significant difference among stiff spring energy transferring patterns produced by Velocity Verlet, FLAVOR, IMEX and the Impulse Method. Here there is no exact or provably accurate solution to compare to. More precisely Velocity Verlet is no longer accurate over this very large time scale (even with $h = 10^{-5}$) and shows distortions in the energy transfer pattern of the stiff springs. IMEX and the Impulse method show similar distortions in the energy transfer pattern of the stiff springs. Based on the following observations, FLAVOR appears to be the only integrator giving accurate and un-deformed trajectories on this very long time scale. Mathematically it can be shown that the dynamical system admits periodic orbits (for example, by Poincaré-Bendixson theorem); intuitively half a period roughly goes from the beginning to around time 1000, in which three stiff springs alternatively obtain their maximal energies, and these maximal energies should be of fixed values. In addition, if we change the slow potential to be quadratic the system is still very similar to non-harmonic FPU however could be solved analytically. The resulting energy exchanging pattern (Figure 12) resembles FLAVORS’s result of the non-harmonic system but not the other integrators’. Notice that if ran on the modified quadratic FPU problem however, FLAVORS, Velocity-Verlet, IMEX and Impulse Method all obtain perfect results (plots omitted).

On the other hand, Velocity-Verlet, IMEX and Impulse Method gradually lost the ability to obtain the maximum energy that should be contained in the some stiff spring and hence the periodic pattern of the energy exchanging. Mollified Impulse Method can’t
(a) By Variational Euler with small timestep $\tau' = 5 \times 10^{-5} = 0.05/\omega$. 38 periods in sub-figure 2 with zoomed-in time axis (∼380 in total over the whole simulation span)

(b) By artificial FLAVOR (3.11) with mesostep $\delta = 0.002$ and microstep $\tau = 10^{-4} = 0.1/\omega$. 38 periods in sub-figure 2 with zoomed-in time axis (∼380 in total over the whole simulation span)

Figure 9: Simulations of FPU problem over $T = 2\omega$. Subplot2 has a zoomed-in time-axis so that whether phase lag or any other distortion of trajectory exists could be closely investigated. In this experiment $m = 3$, $\omega = 10^3$, $x(0) = [0.4642, -0.4202, 0.0344, 0.1371, 0.0626, 0.0810]$ and $y(0) = [0, 0, 0, 0, 0]$ are randomly chosen.
(a) By Velocity Verlet with small timestep $h = 10^{-5}$.

(b) By artificial FLAVOR (3.11) with mesostep $\delta = 0.002$ and microstep $\tau = 0.0005 = 0.1/\omega$.

(c) By IMEX with mesostep $\delta = 0.002$.

(d) By Impulse Method with mesostep $\delta = 0.002$.

Figure 10: Simulations of FPU problem over $T = \frac{1}{2}\omega^2$. Initial conditions are $x(0) = [1, 0, 0, 1/\omega, 0, 0]$ and $y(0) = [0, 0, 0, 0, 0, 0]$ so that energy starts concentrated on the leftmost soft and stiff springs, propagates to the right, bounces back, and oscillates among springs. We chose a smaller $\omega = 200$ because with a larger $\omega$ it would take weeks to run Velocity Verlet on a laptop.
get the first period right and hence not plotted here. The fact that implicit methods cannot correctly capture non-Dirac invariant distributions \[61\] may explain the distortions observed with IMEX.

It is worth discussion why Velocity-Verlet with tiny timestep is still not satisfactory. Being a second order method it has an error bound of \(O(e^T h^2)\); on the other hand backward error analysis guarantees that the energy of the integrated trajectory oscillates around the true conserved energy, hence eliminating the possibility that numerical artifacts blow up the integrated solution. Nevertheless, these two analytical results don’t guarantee a long term accuracy on the stiff springs’ energies in a long time simulation. The energy exchange among stiff springs is in fact an delicate phenomenon, and a slight distortion in stiff spring lengths could easily wreck its period, if not the whole of it.

These numerical observation seem to indicate that symplectic FLAVORS have special long time properties. Notice that most error bounds for numerical integrations are only valid on \(T = O(1)\), and error will grow exponentially as simulation time increases. Our general theory for FLAVORS gets rid of \(\epsilon\) dependence in the error bound, but its error bound still exponentially depends on \(T\). In the previous example we have observed a linear growth of the error with FLAVORS. We couldn’t quantify the error here because there is no benchmark to compare to when the total simulation time is \(O(\omega^2)\), but the long term behavior seems to indicate accuracy and stability. A rigorous investigation on FLAVORS’s mysterious long time behavior remains to be done.

Figure 13 summarizes FLAVOR’s performance on various timescales in a comparison to Velocity Verlet.

Notice there are many sophisticated methods designed for integrating FPU problem (for a review see [45] for example), as well as general multiscale methods that can be applied to FPU problem. HMM as one state-of-art method in the latter category,
FLAVOR (3.11) captures the fastest timescale in the sense of measure, while Velocity Verlet can’t accurately capture the slowest ($O(\omega^2)$) timescale despite of the small timestep it uses. Here FLAVOR is 200 times faster than Velocity Verlet. All parameters are the same as in Figure 10(a) and 10(b), e.g. $\omega = 200$, $\delta = 0.002$, $\tau = 0.0005$, and $h = 10^{-5}$.

Figure 13: Quantities of interest in integrations of FPU over different timescales. FLAVOR captures the fastest timescale in the sense of measure, while Velocity Verlet can’t accurately capture the slowest ($O(\omega^2)$) timescale despite of the small timestep it uses. Here FLAVOR is 200 times faster than Velocity Verlet. All parameters are the same as in Figure 10(a) and 10(b), e.g. $\omega = 200$, $\delta = 0.002$, $\tau = 0.0005$, and $h = 10^{-5}$.

REMARK: (on resonance) In the FPU problem there is Takens resonance [84], because the eigenfrequencies of the strong potential are identical. Therefore there is no homogenized equation describing the $\omega \to \infty$ limiting behavior [12]. Nevertheless it is still possible to capture the solution trajectories given any large value of $\omega$ using FLAVORS with mesostep $\delta \gg 1/\omega$ independent of $\omega$.

3.7 Nonlinear 2D Primitive Molecular Dynamics Example

Now consider a 2D 2DOF example in which a point mass is linked through a spring to a massless fixed hinge at the origin. While the spring as well as the point mass are allowed to rotate around the hinge (the spring remains straight), the more the spring-mass tilts away from its equilibrium angle the more restorative force it will experience. This example is a simplified version of prevailing molecular dynamics models, in which bond lengths and angles between neighboring bonds are both spring like; other potential energy terms that are ignored.

Denote by $x$ and $y$ the Euclidean coordinates of the mass, and $p_x$, $p_y$ corresponding momentums. Also introduce polar coordinates $(r, \theta)$, defined through $x = r \cos \theta$ and $y = r \sin \theta$. Then the Hamiltonian writes
\[ H = \frac{1}{2}p_x^2 + \frac{1}{2}p_y^2 + \frac{1}{2}\omega^2(r - r_0)^2 + (\cos \theta)^2 \]
\[ = \frac{1}{2}p_x^2 + \frac{1}{2}p_y^2 + \frac{1}{2}\omega^2(\sqrt{x^2 + y^2} - r_0)^2 + \frac{x^2}{x^2 + y^2} \]  
(3.26)

where \( r_0 \) is equilibrium bond length parameter and \( \omega \) is large number denoting bond oscillation frequency.

Remark 3.4. This seemingly-trivial example is not easy to integrate.

1. If the system is viewed in Euclidean coordinates \((x, y, p_x, p_y)\) it is completely nonlinear with a nonpolynomial potential, and hence Impulse Method or its derivatives \[42, 88, 36, 76, 86\], or IMEX \[83\], or the homogenization method introduced in \[17\] can’t be applied.

2. If the Hamiltonian is rewritten in generalized coordinates \((r, \theta, p_r, p_\theta)\), \( H = \frac{1}{2}p_r^2 + \frac{1}{2}\omega^2(r - r_0)^2 + \frac{1}{2}\cos(\theta)^2 \), a fast quadratic potential can be identified.

However, the mass matrix \[\begin{bmatrix} 1 & 0 \\ 0 & r^2 \end{bmatrix}\] is not a constant but rapidly oscillating, and hence methods that work for quasi-quadratic fast potential (i.e. "harmonic oscillator" with a slowly changing frequency) \[17, 86\] can’t be applied.

Figure 14 compares symplectic Euler with the induced symplectic FLAVOR (3.5) and (3.9)) applied to (3.26) in Euclidean coordinates. FLAVOR reproduced the slow variable \((\theta)\) trajectory while accelerating the simulation time by roughly 50-fold (since \(\delta = 50\tau\)). It can also be seen from both energy fluctuations and the trajectory of the fast variable that the fast process’ amplitude is well captured although itself has been stretched along the time axis.

3.8 Nonlinear 2D Molecular Clipper.

We now consider a united-atom representation of 3-atom polymer with 2 bonds (e.g. propane molecule or water molecule). This is a simplified version of several prevailing molecular dynamics force fields (for example, CHARMM \[18\], AMBER \[27\], or a simpler example of butane \[73, 75\]). Due to conservation of angular momentum we can assume that the polymer remains in a 2D plane. Introduce both Cartesian coordinates \((x_1, y_1, x_2, y_2, x_3, y_3)\), as well as generalized coordinates \(r_1 = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}\) and \(r_2 = \sqrt{(x_3 - x_2)^2 + (y_3 - y_2)^2}\) for bond lengths and \(\theta\) for the angle between the two bonds (Figure 15). The kinetic energy writes as

\[ K.E. = \frac{1}{2}m_1(\dot{x}_1^2 + \dot{y}_1^2) + \frac{1}{2}m_2(\dot{x}_2^2 + \dot{y}_2^2) + \frac{1}{2}m_3(\dot{x}_3^2 + \dot{y}_3^2) \]  
(3.27)

where \(m_1, m_2, m_3\) denotes the masses of the atoms (for propane \(m_1 = 15, m_2 = 14, m_3 = 15\)).
Figure 14: Simulation of (3.26). Symplectic Euler uses small timestep $\tau = 0.0002$ and the induced symplectic FLAVOR ((3.5) and (3.9)) uses mesostep $\delta = 0.01$ and microstep $\tau = 0.0002$. In this simulation $\omega = 500$, $x(0) = 1.1$, $y(0) = 0.8$, $p_x(0) = 0$, $p_y(0) = 0$, simulation time $T = 100$.

The potential energy consists a bond term and a bond angle term, both of which are harmonic oscillator-like:

$$P.E. = V_{\text{bond}} + V_{\text{angle}}$$  \hspace{1cm} (3.28)

$$V_{\text{bond}} = \frac{1}{2} K_r [(r_1 - r_0)^2 + (r_2 - r_0)^2]$$  \hspace{1cm} (3.29)

$$V_{\text{angle}} = \frac{1}{2} K_\theta (\cos(\theta) - \cos(\theta_0))^2$$  \hspace{1cm} (3.30)

where for propane $r_0 = 1.53 \text{Å}$, $K_r = 83.7 \text{kcal/(molÅ}^2\text{)}$, $\theta_0 = 109.5^\circ$, and $K_\theta = 43.1 \text{kcal/(molÅ}^2\text{)}$. [73] Notice the system is in fact fully nonlinear: if written in generalized coordinates the kinetic energy will correspond to a mass matrix nonlinear and position dependent, whereas in Cartesian coordinates both terms in the potential energy are non-polynomial functions in positions.

The above system is characterized by a separation of timescales to some extent: bond stretching and bond-angle bending are characterized by $10^{14}$, $10^{13}$ Hz vibrational frequencies respectively [93]. In this numerical experiment we have exaggerated this difference and set $K_r$ to be 8370 and $K_\theta$ to be 4.31.

In this system, the bond potential is the fast potential and the bond-angle potential is the slow one. It is well known that using a large timestep at the timescale corresponding
to the bond-angle potential by freezing bond lengths produces biased results [35], and many physics-based methods have been proposed to remedy this difficulty (see a review in [93] for example). On the other hand, few multiscale methods work for this fully nonlinear system.

Figure 16 compares symplectic Euler with the induced symplectic FLAVOR ((3.5) and (3.9)) applied in Euclidean coordinates. 10 fold acceleration is achieved. A simulation movie is also available at http://www.cds.caltech.edu/~mtao/Propane.avi and http://www.acm.caltech.edu/~owhadi/.

3.9 Forced non-autonomous mechanical system: Kapitza’s Inverted Pendulum.

As the famous Kapitza’s inverted pendulum shows [50], the up state of a single pendulum can be stabilized if the pivot base of the pendulum experiences an external forcing in the form of vertical oscillation. Specifically, if the pivot base has its position in $y$ direction as a harmonic motion $y = \sin(\omega t)$, by Lagrangian mechanics the system is governed by

$$l\ddot{\theta} = [g + \omega^2 \sin(2\pi \omega t)] \sin \theta$$

(3.31)

where $\theta$ denotes the clockwise angle of the pendulum from the positive $y$ direction, $l$ is the length of the pendulum and $g$ is the gravitational constant. In this case, the rapid vibration causes the pendulum to oscillate slowly around the positive $y$ direction with a $O(1)$ frequency.

A single scale integration of this system could be done by Variational Euler with discrete d’Alembert principle for external forces [63]

$$\begin{align*}
 f_i & = \omega^2 \sin(2\pi \omega h) \\
 p_{i+1} & = p_i + h[g + f]\sin \theta_i \\
 \theta_{i+1} & = \theta_i + hp_{i+1}/l
\end{align*}$$

(3.32)
Figure 16: Simulations of propane molecule with $K_r = 8370$ and $K_\theta = 4.31$ (Subsection 3.8). Symplectic Euler uses $h = 0.01$ and the induced symplectic FLAVOR (3.5) and (3.9) parameters are $\delta = 0.1$ and $\tau = 0.01$. Initial conditions are $[x_1, y_1, x_2, y_2, x_3, y_3] = [0, 0.1253, 0.2877, -1.1465, 1.1909]$.

where the timestep length $h$ has to be smaller than $O(1/\omega)$.

FLAVOR is given by

$$
\begin{align*}
q_{n\delta+\tau} &= q_{n\delta} + \tau p_{n\delta}/l \\
p_{n\delta+\tau} &= p_{n\delta} + \tau g \sin(q_{n\delta+\tau}) + \omega^2 \sin(2\pi \omega n\tau) \\
q_{(n+1)\delta} &= q_{n\delta+\tau} + (\delta - \tau)p_{n\delta+\tau}/l \\
p_{(n+1)\delta} &= p_{n\delta+\tau} + (\delta - \tau)g \sin(q_{(n+1)\delta})
\end{align*}
$$

(3.33)

Observe that the time dependent force is synchronized on the $\tau$ time scale instead of the $\delta$ time scale, specifically $\omega^2 \sin(2\pi \omega n\tau)$ instead of $\omega^2 \sin(2\pi \omega n\delta)$ in (3.33).

Numerical results are illustrated in Figure 17 (also available as a movie at http://www.cds.caltech.edu/~mtao/InvertedPendulum.avi and http://www.acm.caltech.edu/~owhadi/). Notice in this example $\theta$ being the only degree of freedom contains a combination of slow and fast dynamics. FLAVOR could only capture the fast dynamics in the sense of measures, and this is why dents appear as modulation on the slow oscillation of $\theta$. On the other hand, although this forced system doesn’t admit a conserved energy, the value of the Hamiltonian should oscillate periodically due to the periodic external driving force. While a non mechanics-based method such as Forward Euler
Figure 17: Simulations of inverted pendulum. The integration by Variational Euler + d’Alembert principle uses timestep \( h = 0.2/\omega/\sqrt{l} \approx 0.000067 \), while FLAVOR (defined by (3.33)) uses \( \delta = 0.002 \) and \( \tau = 0.2/\omega/\sqrt{l} \). \( g = 9.8 \), \( l = 9 \), \( \theta(0) = 0.2 \), \( \dot{\theta}(0) = 0 \). \( \omega = 1000 \)

often produces an unbounded growth or a decrease in the energy, FLAVORS don’t have this drawback.

4 SDEs

4.1 Two scale flow convergence for SDEs.

Consider the following SDE on \( \mathbb{R}^d \).

\[
du^\epsilon_t = \left( G(u^\epsilon_t) + \frac{1}{\epsilon} F(u^\epsilon_t) \right) dt + \left( H(u^\epsilon_t) + \frac{1}{\sqrt{\epsilon}} K(u^\epsilon_t) \right) dW_t, \quad u^\epsilon_0 = u_0 \quad (4.1)
\]

where \( (W_t)_{t \geq 0} \) is a \( d \)-dimensional Brownian Motion; \( F, G \) are uniformly Lipschitz vector fields on \( \mathbb{R}^d \); \( H, K \) are uniformly Lipschitz \( d \times d \) matrix fields on \( \mathbb{R}^d \).

In Subsection 4.3 we will consider the more general form (4.15) but for the sake of clarity we will start the description of with (4.1).

We will first extend the definition of two scale flow convergence introduced in Subsection 2.1 to stochastic processes.
Two scale flow convergence for SDEs. Let \((\xi^\epsilon(t, \omega))_{t \in \mathbb{R}^+ \times \Omega}\) be a sequence of stochastic processes on \(\mathbb{R}^d\) (progressively measurable mappings from \(\mathbb{R}^+ \times \Omega\) to \(\mathbb{R}^d\)) indexed by \(\epsilon > 0\). Let \((X_t)_{t \in \mathbb{R}^+}\) be a (progressively measurable) stochastic process on \(\mathbb{R}^{d-p}\) \((p \geq 0)\). Let \(x \mapsto \nu(x, dz)\) be a function from \(\mathbb{R}^{d-p}\) into the space of probability measures on \(\mathbb{R}^d\).

**Definition 4.1.** We say that the process \(\xi^\epsilon\) F-converges towards \(\nu(X, dz)\) as \(\epsilon \downarrow 0\) and write \(\xi^\epsilon \xrightarrow{F} \nu(X, dz)\) if and only if for all function \(\varphi\) bounded and uniformly Lipschitz-continuous on \(\mathbb{R}^d\), and for all \(t > 0\)

\[
\lim_{h \to 0} \lim_{\epsilon \to 0} \frac{1}{h} \int_{t}^{t+h} \mathbb{E}[\varphi(\xi^\epsilon_s)] \, ds = \mathbb{E}\left[\int_{\mathbb{R}^d} \varphi(z) \nu(X_t, dz)\right]
\]

(4.2)

**Condition 4.1.** The F-convergence of solutions of (4.1) holds under the following conditions

1. \(F, G, H\) and \(K\) are uniformly bounded and Lipschitz continuous.
2. There exists a diffeomorphism \(\eta := (\eta^x, \eta^y)\), from \(\mathbb{R}^d\) onto \(\mathbb{R}^{d-p} \times \mathbb{R}^p\), independent of \(\epsilon\), with uniformly bounded \(C^1, C^2, C^3\) derivatives, such that the process \((x^\epsilon_t, y^\epsilon_t) = (\eta^x(u^\epsilon_t), \eta^y(u^\epsilon_t))\) satisfies the SDE

\[
\begin{aligned}
&dx^\epsilon = g(x^\epsilon, y^\epsilon) \, dt + \sigma(x^\epsilon, y^\epsilon) \, dW_t, \quad x^\epsilon_0 = x_0 \\
&dy^\epsilon = \frac{1}{\sqrt{\epsilon}} f(x^\epsilon, y^\epsilon) \, dt + \frac{1}{\sqrt{\epsilon}} Q(x^\epsilon, y^\epsilon) \, dW_t, \quad y^\epsilon_0 = y_0
\end{aligned}
\]

(4.3)

where \(g\) is \(d - p\) dimensional vector field; \(f\) a \(p\)-dimensional vector field; \(\sigma\) is a \((d - p) \times d\)-dimensional matrix field; \(Q\) a \(p \times d\)-dimensional matrix field and \(W_t\) a \(d\)-dimensional Brownian Motion.

3. Let \(Y_t\) be the solution to

\[
\begin{aligned}
dY_t &= f(x_0, Y_t) \, dt + Q(x_0, Y_t) \, dW_t \quad Y_0 = y_0
\end{aligned}
\]

(4.4)

there exists a family of probability measures \(\mu(x, dy)\) on \(\mathbb{R}^p\) indexed by \(x \in \mathbb{R}^{d-p}\) and a positive function \(T \mapsto E(T)\) such that \(\lim_{T \to \infty} E(T) = 0\) and such that for all \(x_0, y_0, T\) and \(\phi\) with uniformly bounded \(C^r\) derivatives for \(r \leq 3\),

\[
\left| \frac{1}{T} \int_0^T \mathbb{E}[\phi(Y_s)] - \int \phi(y) \mu(x_0, dy) \right| \leq \chi(||(x_0, y_0)||) E(T) \max_{r \leq 3} ||\phi||_{C^r}
\]

(4.5)

where \(r \mapsto \chi(r)\) is bounded on compact sets.

4. For all \(u_0, T > 0\), \(\sup_{0 \leq t \leq T} \mathbb{E}\left[\chi(||u^\epsilon_t||)\right]\) is uniformly bounded in \(\epsilon\).

**Remark 4.1.** As in the proof of Theorem 2.2, the uniform regularity of \(F, G, H, K\) can be relaxed to local regularity by adding a control on the rate of escape of the process towards infinity. To simplify the presentation we will use the global uniform regularity.
Remark 4.2. A comprehensive monograph on the limit behavior of SDEs of type (4.3) can be found in Chapter II of [82]. When $\sigma = 0$, we refer to [31] for an error analysis of HMM for SDE (4.3) with separated slow and fast variables under exponential mixing; we also refer to Subsection 3.3 and 3.4 of Chapter II of [82] for an asymptotic analysis including local errors.

Our goal here is to show that solutions of (4.1) $F$-converge and can be approximated (in terms of two scale flow-convergence) with FLAVORS directly applied to SDE (4.1) with hidden slow and fast processes. When FLAVORS are applied to SDE (4.3) with explicitly separated slow and fast processes, they lead to integrators that are locally in the neighborhood of those obtained with equation free or HMM methods with only one microstep per macrostep and with a reinitialization of the fast variables at macrotime $n$ by their final value at macrotime step $n-1$ [31]. However because HMM uses averaged drift while FLAVORS use composition of flows, FLAVORS and HMM are characterized by different global behaviors when applied to structured systems (such as Hamiltonian systems, see Section 3 and Section 5).

Theorem 4.1. Let $u^\epsilon$ be the solution to (4.1). Assume that Conditions 4.1 are satisfied then

- $u^\epsilon_t$ $F$-converges towards $\eta^{-1} * (\delta_{X_t} \otimes \mu(X_t, dy))$ as $\epsilon \downarrow 0$ where $X_t$ is the solution to

$$dX_t = \int g(X_t, y) \mu(X_t, dy) dt + \bar{\sigma}(X_t) dB_t \quad X_0 = x_0 \quad (4.6)$$

where $\bar{\sigma}$ is a $(d-p) \times (d-p)$ matrix field defined by

$$\bar{\sigma} T = \int \sigma T(x, y) \mu(x, dy) \quad (4.7)$$

and $B_t$ a $(d-p)$-dimensional Brownian Motion.

Remark 4.3. $x^\epsilon_t$ is converging weakly towards $X_t$ but if $\sigma \neq 0$ and is not constant it is not converging strongly. A simple counter-example can be constructed by taking $g = 0$ and letting $Q, f$ and $\sigma$ depend only on $y$.

Proof. The proof of Theorem 4.1 is classical and can be deduced from that of Theorem 4.2. In particular we also refer to Chapters II of [82].

4.2 Flow Averaging Integrators for stiff SDEs. Convergence theorem.

Let $\omega$ be a random sample from a probability space $(\Omega, F, \mathbb{P})$ and $\theta^G(\cdot, \omega)$ a random mapping from $\mathbb{R}^d$ onto $\mathbb{R}^d$ approximating in distribution the flow of (4.1) over time steps $\tau \ll \epsilon$. Let $\theta_h^G(\cdot, \omega)$ a random mapping from $\mathbb{R}^d$ onto $\mathbb{R}^d$ approximating in distribution the flow of

$$dv^\epsilon_t = G(v^\epsilon_t) dt + H(v^\epsilon_t) dW_t \quad (4.8)$$

over time steps $h \ll 1$. More precisely, we will assume that $\theta^G$ and $\theta^F$ satisfy the following conditions.
**Condition 4.2.** We will prove the $F$-convergence of FLAVORS for (4.1) under the following additional assumptions on $G$ and $\epsilon$.

1. There exists $h_0, C > 0$ and a $d$-dimensional centered Gaussian vector $\xi(\omega)$ with identity covariance matrix such that for $h \leq h_0$,
   \[
   \left( \mathbb{E} \left[ \left| \theta^G_h(u, \omega) - u - hG(u) - \sqrt{hH(u)}\xi(\omega) \right|^2 \right] \right)^{\frac{1}{2}} \leq Ch^{\frac{3}{2}} \tag{4.9}
   \]

2. There exists $\tau_0, C > 0$ and a $d$-dimensional centered Gaussian vector $\xi(\omega)$ with identity covariance matrix such that for $\tau \epsilon \leq \tau_0$,
   \[
   \left( \mathbb{E} \left[ \left| \theta^\epsilon_\tau(u, \omega) - u - \tau G(u) - \frac{\tau}{\epsilon} F(u) - \sqrt{\tau H(u)\xi(\omega)} - \frac{\tau}{\epsilon} K(u)\xi(\omega) \right|^2 \right] \right)^{\frac{1}{2}} \leq C(\frac{\tau}{\epsilon})^{\frac{3}{2}} \tag{4.10}
   \]

3. For all $u_0$, $T > 0$, $\sup_{0 \leq n \leq T/\delta} \mathbb{E} \left[ \chi(\|\bar{u}_n\|) \right]$ is uniformly bounded in $\epsilon$, $0 < \delta \leq h_0$, $\tau \leq \min(\tau_0, \epsilon)$, where $\bar{u}$ is defined by (4.11).

**FLAVORS** Let $\delta \leq h_0$ and $\tau \in (0, \delta)$ such that $\tau \leq \tau_0 \epsilon$. We define FLAVORS as the class of algorithms simulating the stochastic process $t \mapsto \bar{u}_t$ defined by

\[
\begin{cases}
\bar{u}_0 = u_0 \\
\bar{u}_{(k+1)\delta} = \theta^G_{\delta-t}(., \omega_k) \circ \theta^\epsilon_\tau(\bar{u}_{k\delta}, \omega_k) \\
\bar{u}_t = \bar{u}_{k\delta} & \text{for } k\delta \leq t < (k+1)\delta
\end{cases} \tag{4.11}
\]

where $\omega_k, \omega'_k$ are i.i.d samples from the probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

**Non intrusive FLAVORS** Assume that we are given a random mapping $\Phi^\alpha_t(., .)$ from $\mathbb{R}^d$ onto $\mathbb{R}^d$ approximating the flow of (4.1) for $\alpha = 1/\epsilon$. If the parameter $\alpha$ can be controlled then $\Phi^\alpha_t$ can be used as a black box for accelerating the computation of solutions of (4.1) without prior identification of slow variables. Indeed assume that there exists a constant $h_0 > 0$ and a normal random vector $\xi(\omega)$ such that for $t \leq h_0 \min(\frac{1}{\alpha}, 1)$

\[
\left( \mathbb{E} \left[ \left| \Phi^\alpha_t(u, \omega) - u - tG(u) - \alpha t F(u) - \sqrt{tH(u)\xi(\omega)} - \sqrt{\alpha t K(u)\xi(\omega)} \right|^2 \right] \right)^{\frac{1}{2}} \leq C t^{\frac{3}{2}} (1 + \alpha)^{\frac{3}{2}} \tag{4.12}
\]

then $\theta^\epsilon_t := \Phi^\alpha_t$ and $\theta^G_{\delta-t} := \Phi^0_t$ satisfy Conditions (4.2) and FLAVOR can be defined as the algorithm simulating the stochastic process

\[
\begin{cases}
\bar{u}_0 = u_0 \\
\bar{u}_{(k+1)\delta} = \Phi^0_{\delta-t}(., \omega'_k) \circ \Phi^\frac{1}{\epsilon}_t(\bar{u}_{k\delta}, \omega_k) \\
\bar{u}_t = \bar{u}_{k\delta} & \text{for } k\delta \leq t < (k+1)\delta
\end{cases} \tag{4.13}
\]
The following Theorem 4.2 shows that the flow averaging integrator is accurate with respect to $F$-convergence for $\tau \ll \epsilon \ll \delta$ and

$$\left( \frac{\tau}{\epsilon} \right)^{\frac{2}{3}} \ll \delta \ll \frac{\tau}{\epsilon}$$

\textbf{Theorem 4.2.} Consider the stochastic process $\bar{u}_t$ defined by (4.11) or (4.13). Assume that Conditions 4.1 and 4.2 are satisfied then as $\epsilon \downarrow 0$, $\tau \leq \delta$, $\frac{\tau}{\epsilon} \downarrow 0$, $\frac{\delta}{\epsilon} \downarrow 0$ and $\left( \frac{\tau}{\epsilon^2} \right)^{\frac{1}{3}} \downarrow 0$

- $\bar{u}_t$ $F$-converges towards $\eta^{-1} \ast (\delta X_t \otimes \mu(X_t, dy))$ where $X_t$ is the solution to (4.6).

\textbf{Proof.} We refer to Subsection 6.2 of the appendix for the proof of Theorem 4.2. \hfill $\square$

### 4.3 Natural FLAVORS generalization to generic stiff SDEs.

FLAVORS for stochastic systems have a natural generalization to SDEs on $\mathbb{R}^d$ of the form

$$d u^{\alpha, \epsilon} = F(u^{\alpha, \epsilon}, \alpha, \epsilon) \, dt + K(u^{\alpha, \epsilon}, \alpha, \epsilon) \, dW_t$$

where $(W_t)_{t \geq 0}$ is a $d$-dimensional Brownian Motion, $F$ and $K$ are Lipschitz continuous in $u$.

\textbf{Condition 4.3.} The $F$-convergence of $u^{1, \epsilon}$ as $\epsilon \downarrow 0$ holds under the following conditions.

1. $\gamma \mapsto F(u, \alpha, \gamma)$ and $\gamma \mapsto K(u, \alpha, \gamma)$ are uniformly continuous in the neighborhood of $0$.

2. There exists a diffeomorphism $\eta := (\eta^x, \eta^y)$ from $\mathbb{R}^d$ onto $\mathbb{R}^{d-p} \times \mathbb{R}^p$, independent from $\epsilon, \alpha$, with uniformly bounded $C^1, C^2, C^3$ derivatives, and such that the stochastic process $(x_t^{\alpha}, y_t^\alpha) = (\eta^x(u_t^{\alpha,0}), \eta^y(u_t^{\alpha,0}))$ satisfies for all $\alpha \geq 1$ the SDE

$$dx^{\alpha} = g(x^{\alpha}, y^{\alpha}) \, dt + \sigma(x^{\alpha}, y^{\alpha}) \, dW_t \quad x_0^{\alpha} = x_0$$

where $g$ is $d-p$ dimensional vector field, $\sigma$ is a $(d-p) \times d$-dimensional matrix field, $g$ and $\sigma$ are uniformly bounded and Lipschitz continuous in $x$ and $y$.

3. There exists a family of probability measures $\mu(x, dy)$ on $\mathbb{R}^p$ such that for all $x_0, y_0, T$ $(x_0, y_0) := \eta(u_0))$ and $\varphi$ with uniformly bounded $C^r$ derivatives for $r \leq 3$,

$$\left( \frac{1}{T} \right) \int_0^T \mathbb{E} [\varphi(y_t^\alpha)] \, ds - \int_0^T \varphi(y) \mu(x_0, dy) \leq \chi \|x_0, y_0\| \max_{r \leq 3} \|\varphi\|_{C^r}$$

where $r \mapsto \chi(r)$ is bounded on compact sets and $E_2(r) \to 0$ as $r \to \infty$ and $E_1(r) \to 0$ as $r \to 0$.

4. For all $u_0$, $T > 0$, $\sup_{0 \leq t \leq T} \mathbb{E} \left[ \chi(\|u_t^{\alpha,0}\|) \right]$ is uniformly bounded in $\alpha \geq 1$.

\textbf{Remark 4.4.} As in the proof of Theorem 2.2 the uniform regularity of $g$ and $\sigma$ can be relaxed to local regularity by adding a control on the rate of escape of the process towards infinity. To simplify the presentation we have used the global uniform regularity.
Let \( \omega \) be a random sample from a probability space \((\Omega, \mathcal{F}, \mathbb{P})\) and \( \Phi_{\alpha, \epsilon}^{t}(., \omega) \) a random mapping from \( \mathbb{R}^{d} \) onto \( \mathbb{R}^{d} \) approximating in distribution the flow of (4.15) over time steps \( \tau \ll \epsilon \). If the parameter \( \alpha \) can be controlled then \( \Phi_{\alpha, \epsilon}^{t} \) can be used as a black box for accelerating the computation of solutions of (4.15). The acceleration is obtained without prior identification of the slow variables.

**Condition 4.4.** We will prove the \( F \)-convergence of FLAVORS under the following conditions on \( \Phi_{\alpha, \epsilon}^{t} \).

1. There exists \( h_{0}, C > 0 \) and a \( d \)-dimensional centered Gaussian vector \( \xi(\omega) \) with identity covariance matrix such that for \( h \leq h_{0} \), \( 0 < \epsilon \leq 1 \leq \alpha \) and \( t \leq h_{0} \min(\frac{1}{\alpha^{2}}, 1) \)

\[
\left( \mathbb{E} \left[ \left| \Phi_{t}^{\alpha, \epsilon}(u) - u - t F(u, \alpha, \epsilon) - \sqrt{t} \xi(\omega) K(u, \alpha, \epsilon) \right|^2 \right] \right)^{\frac{1}{2}} \leq C \left( 1 + \frac{3}{\alpha \epsilon^{2}} \right) \tag{4.18}
\]

2. For all \( u_{0}, T > 0 \), \( \sup_{0 \leq n \leq T/\delta} \mathbb{E} \left[ \chi(\|\bar{u}_{n}\|) \right] \) is uniformly bounded in \( \epsilon, 0 < \delta \leq h_{0}, \tau \leq \min(h_{0} \epsilon^{\nu}, \delta) \), where \( \bar{u} \) is defined by (4.19).

**FLAVORS** Let \( \delta \leq h_{0} \) and \( \tau \in (0, \delta) \) such that \( \tau \leq \tau_{0} \epsilon^{\nu} \). We define FLAVORS as the class of algorithms simulating the stochastic process \( t \mapsto \bar{u}_{t} \) defined by

\[
\left\{ \begin{array}{l}
\bar{u}_{0} = u_{0} \\
\bar{u}_{(k+1)\delta} = \Phi_{\delta-\tau}^{0, \epsilon}(., \omega_{k}') \circ \Phi_{\delta}^{\epsilon} \left( \bar{u}_{k\delta}, \omega_{k} \right) \\
\bar{u}_{t} = \bar{u}_{k\delta} \quad \text{for} \quad k\delta \leq t < (k+1)\delta
\end{array} \right. \tag{4.19}
\]

where \( \omega_{k}, \omega_{k}' \) are i.i.d samples from the probability space \((\Omega, \mathcal{F}, \mathbb{P})\).

The following theorem shows that the flow averaging integrator is accurate with respect to \( F \)-convergence for \( \tau \ll \epsilon^{\nu} \ll \delta \) and

\[
\left( \frac{\tau}{\epsilon^{\nu}} \right)^{\frac{3}{2}} \ll \delta \ll \frac{\tau}{\epsilon^{\nu}} \tag{4.20}
\]

**Theorem 4.3.** Let \( \bar{u}^{\frac{1}{\epsilon}, \epsilon}_{t} \) be the solution to (4.15) with \( \alpha = 1/\epsilon \) and \( \bar{u}_{t} \) be defined by (4.19). Assume that Conditions 4.3 and 4.4 are satisfied then

- \( \bar{u}^{\frac{1}{\epsilon}, \epsilon}_{t} \) \( F \)-converges towards \( \eta^{-1} * (\delta X_{t} \otimes \mu(X_{t}, dy)) \) as \( \epsilon \downarrow 0 \) where \( X_{t} \) is the solution to

\[
dX_{t} = \int g(X_{t}, y) \mu(X_{t}, dy) + \bar{\sigma}(X_{t}) dB_{t} \quad X_{0} = x_{0} \tag{4.21}
\]

where \( \bar{\sigma} \) is a \((d - p) \times (d - p)\) matrix field defined by

\[
\bar{\sigma} \bar{\sigma}^{T} = \int \sigma \sigma^{T}(x, y) \mu(x, dy) \tag{4.22}
\]

and \( B_{t} \) a \((d - p)\)-dimensional Brownian Motion.
• As $\epsilon \downarrow 0$, $\tau \epsilon^{-\nu} \downarrow 0$, $\delta \epsilon^{\nu} \downarrow 0$, $(\frac{\tau}{\epsilon})^{2} \frac{1}{\delta} \downarrow 0$, $\bar{u}_t$ F-converges towards $\eta^{-1} * (\delta X_t \otimes \mu(X_t, dy))$ as $\epsilon \downarrow 0$ where $X_t$ is the solution to (4.24).

Proof. The proof of Theorem 4.3 is similar to the proof of Theorem 4.2. The condition $\epsilon \ll 1$ is needed for the approximation of $u^{\alpha, \epsilon}$ by $u^{\alpha, 0}$ and for the F-convergence of $u^{\epsilon, 0}$. Since $y_t^\alpha = \eta^\alpha(u_t^{\alpha, 0})$ the condition $\tau \ll \epsilon^\nu$ is used along with equation (4.18) for the accuracy of $\Phi^{1, \epsilon}$ in (locally) approximating $y_t^\alpha$. The condition $\delta \ll \frac{\tau}{\epsilon^\nu}$ allows for the averaging of $g$ and $\sigma$ to take place prior to a significant change of $x_{\alpha t}$, more precisely it allows for $m \gg 1$ iterations of $\Phi^{1, \epsilon}$ prior to a significant change of $x_{\alpha t}$. The condition $(\frac{\tau}{\epsilon^\nu})^\frac{3}{2} \ll \delta$ is required in order to control the error accumulated by $m$ iterations of $\Phi^{1, \epsilon}$.

Figure 18: (a) Integration of (4.23) by non intrusive FLAVOR (4.13) using mesostep step $\delta = 0.01$ (b) Integration of (4.23) by Euler-Maruyama using fine time step $h = 10^{-4}$ (c) Integration of (4.23) by Euler-Maruyama using the same small step $h = 10^{-4}$. Expectations of the slow variable (whether or not hidden) are obtained by empirically averaging over an ensemble of 100 independent sample trajectories. $\epsilon = 10^{-4}$, $x(0) = 1 + \epsilon$, $y(0) = 1$, $T = 2$ (the expectation of the real solution will blow up around $T = 3$). We have chosen $c = 10$ so that the transformation is a diffeomorphism.
4.4 Numerical experiments.

Consider the following artificial non-autonomous SDE system

\[
\begin{align*}
\frac{du}{dt} &= \frac{4}{3(u+v)^2} \left(-\frac{1}{2} \left(\frac{v-u}{2}\right)^2 + 5 \sin(2\pi t)\right) dt - \frac{1}{\epsilon} \left(\frac{u+v}{2}\right)^3 + c - \frac{v-u}{2} \right) dt - \sqrt{\frac{2}{\epsilon}} dW_t \\
\frac{dv}{dt} &= \frac{4}{3(u+v)^2} \left(-\frac{1}{2} \left(\frac{v-u}{2}\right)^2 + 5 \sin(2\pi t)\right) dt + \frac{1}{\epsilon} \left(\frac{u+v}{2}\right)^3 + c - \frac{v-u}{2} \right) dt + \sqrt{\frac{2}{\epsilon}} dW_t
\end{align*}
\]

where \(c\) is a positive constant and two \(dW_t\)'s refer to the same Brownian Motion. The system (4.23) can be converted via the local diffeomorphism

\[
\begin{align*}
u &= (x - c)^{1/3} - y \\
v &= (x - c)^{1/3} + y
\end{align*}
\]

into the following “hidden” system separating slow and fast variables.

\[
\begin{align*}
dx &= -\frac{1}{2}y^2 dt + 5 \sin(2\pi t) dW_t \\
dy &= \frac{1}{\epsilon}(x - y) dt + \sqrt{\frac{2}{\epsilon}} dW_t
\end{align*}
\]

Non intrusive FLAVOR (4.13) can be directly applied to (4.23) using a timestep \(\delta \gg \epsilon\) without prior identification of the slow and fast variables, i.e. without prior identification of the slow variable \(x\) or of the system (4.25). The expected values of solutions of (4.23) integrated by FLAVORS with mesostep \(\delta\) and Euler-Maruyama with a small timestep \(\tau\) are presented in Figure 18. FLAVOR has accelerated the computation by 100 fold.

5 Stochastic Mechanical Systems: Langevin Equations.

5.1 Quasi and conformally symplectic, symmetric and time-reversible FLAVORS for stochastic mechanical systems on manifolds.

In the section we will apply FLAVORS to Langevin equations, SDEs of the form

\[
\begin{align*}
dq &= M^{-1}p \\
dp &= -\nabla V(q) dt - \frac{1}{\epsilon} \nabla U(q) dt - C p dt + \sqrt{2\beta^{-1}C^2} dW_t
\end{align*}
\]

and of the form

\[
\begin{align*}
dq &= M^{-1}p \\
dp &= -\nabla V(q) dt - \frac{1}{\epsilon} \nabla U(q) dt - \frac{C}{\tau} p dt + \sqrt{2\beta^{-1}} C^2 \sqrt{\frac{\epsilon}{\tau}} dW_t
\end{align*}
\]

where \(C\) is a positive symmetric \(d \times d\) matrix.

Remark 5.1. Provided that hidden fast variables remain locally ergodic one can also consider Hamiltonian with a mixture of both slow and fast noise and friction, but for the sake of clarity we have restricted our presentation to (5.1) and (5.2).
Equations (5.1) and (5.2) model a mechanical system with Hamiltonian
\[ H(q, p) := \frac{1}{2} p^T M^{-1} p + V(q) + \frac{1}{\epsilon} U(q) \] (5.3)
with phase space the Euclidean space \( \mathbb{R}^d \times \mathbb{R}^d \) or a cotangent bundle \( T^* \mathcal{M} \) of a configuration manifold \( \mathcal{M} \).

**Remark 5.2.** If \( C \) is non constant and \( \mathcal{M} \) is not the usual \( \mathbb{R}^d \times \mathbb{R}^d \) Euclidean space one should use the Stratonovich integral instead of the Ito integral.

As in Section 3 we assume that we are given a mapping \( \Phi_1^\alpha \) acting on the phase space such that for \( t \leq h_0 \min(1, \alpha^{-\frac{1}{2}}) \)
\[ \left| \Phi_1^\alpha(q, p) - (q, p) - t(M^{-1}p, -V(q) - \alpha U(q)) \right| \leq C \epsilon^2 (1 + \alpha) \] (5.4)

Next consider the following Ornstein-Uhlenbeck equations:
\[ dp = -\alpha C p dt + \sqrt{2 \beta} \sqrt{C} dW_t \] (5.5)
The stochastic flow of (5.5) is defined by the following stochastic evolution map:
\[ \Psi_{t_1, t_2}^\alpha(q, p) = \left( q, e^{-C \alpha (t_2 - t_1)} p + \sqrt{2 \beta} \int_{t_1}^{t_2} e^{-C \alpha (t_2 - s)} dW_s \right) \] (5.6)

Let \( \delta \leq h_0 \) and \( \tau \in (0, \delta) \) such that \( \tau \leq \tau_0 / \sqrt{\alpha} \). FLAVOR for (5.1) can then be defined by
\[ \begin{cases} (\bar{q}_0, \bar{p}_0) = (q_0, p_0) \\ (\bar{q}(k+1)\delta, \bar{p}(k+1)\delta) = \Phi_{\delta - \tau}^0 \circ \Psi_{k\delta, \tau, (k+1)\delta}^1 \circ \Phi_{\tau}^1 \circ \Psi_{\delta, k\delta + \tau}^1 (q, p) \end{cases} \] (5.7)
and FLAVOR for (5.2) can be defined by
\[ \begin{cases} (\bar{q}_0, \bar{p}_0) = (q_0, p_0) \\ (\bar{q}(k+1)\delta, \bar{p}(k+1)\delta) = \Phi_{\delta - \tau}^0 \circ \Phi_{\tau}^1 \circ \Psi_{\delta, k\delta + \tau}^1 (q, p) \end{cases} \] (5.8)

**Remark 5.3.** Observe that if \( \Psi_{\tau}^1 \) is symmetric under a group action for all \( \epsilon > 0 \), then the resulting FLAVOR is symmetric under the same group action.

Theorem 4.3 establishes the accuracy of these integrators under Conditions 4.3 and 4.4 provided that \( \tau \ll \sqrt{\epsilon} \ll \delta \) and \( \left( \frac{\tau}{\sqrt{\epsilon}} \right)^2 \ll \delta \ll \frac{\tau}{\sqrt{\epsilon}} \).

**Theorem 5.1.** We have the following properties for FLAVOR
- If \( \Phi_1^\alpha \) is symplectic then the FLAVORS defined by (5.7) and (5.8) are quasi-symplectic as defined in Conditions RL1 and RL2 of [70] (it degenerates to a symplectic method if friction is set equal to zero and the Jacobian of the flow map is independent of \((q, p)\)).
• If in addition \( C \) is isotropic then FLAVOR defined by (5.7) is conformally symplectic, i.e. it preserves the precise symplectic area change associated to the flow of inertial Langevin processes [65].

Proof. Those properties are a consequence of the fact that FLAVORS are splitting schemes. The quasi-symplecticity and symplectic conformallity of GLA has been obtained in a similar way in [14].

5.1.1 Example of quasi-symplectic FLAVORS.

An example of quasi-symplectic FLAVOR can be obtained by choosing \( \Phi^\alpha_t \) to be the symplectic Euler integrator defined by (3.9). This integrator is also conformally symplectic if \( C \) is isotropic and friction is slow.

5.1.2 Example of quasi-symplectic and time-reversible FLAVORS.

Defining \( \Phi^\alpha_t \) by (3.9) and \( \Phi^\alpha,*,t \) by (3.10), an example of quasi-symplectic and time-reversible FLAVOR can be obtained by using the symmetric Strang splitting FLAVOR can then be defined by

\[
(\bar{q}(k+1),\bar{p}(k+1)) = \Psi^1_{k\delta+\delta/2}(q,p) \circ \Phi^{1,*}_{\delta/2} \circ \Phi^0_{\delta/2} \circ \Phi^1_{\delta} \circ \Phi^1_{k\delta+\delta/2}(q,p) \quad (5.9)
\]
for (5.1) and

\[
(\bar{q}(k+1),\bar{p}(k+1)) = \Psi^1_{(k+1)\delta-\delta/2}(q,p) \circ \Phi^{1,*}_{\delta/2} \circ \Phi^0_{\delta/2} \circ \Phi^1_{\delta} \circ \Phi^1_{(k+1)\delta+\delta/2}(q,p) \quad (5.10)
\]
for (5.2). This integrator is also conformally symplectic if \( C \) is isotropic and friction is slow.

5.1.3 Example of Boltzmann-Gibbs reversible Metropolis-Adjusted FLAVOR

Since the probability density of \( \Psi_{t_1,t_2} \) can be explicitly computed it follows that the probability densities of (5.9) and (5.10) be explicitly computed and these algorithms can be metropolized and made reversible with respect to the Gibbs distribution as it has been shown in [16] for the Geometric Langevin Algorithm [14]. This metropolization leads to stochastically stable (and ergodic if the noise applied on momentum is not degenerate) algorithms. We refer to [16] for details. Observe that if the proposed move is rejected, the momentum has to be flipped and the acceptance probability involves a momentum flip. A remarkable result of [16] is that GLA remains strongly accurate after a metropolization involving local momentum flips. Whether this preservation of accuracy over trajectories transfers in a weak sense (in distributions) to FLAVORS remains to be investigated.
Numerical implementation with slow noise and friction.

In this subsection we consider the one dimensional, two degrees of freedom system illustrated in Figure 2(a) and modeled by the SDEs (now both springs are quartic rather than harmonic):

\[
\begin{align*}
\frac{dx}{dt} &= dp_x \\
\frac{dy}{dt} &= dp_y \\
\frac{dp_x}{dt} &= -\epsilon^{-1}x^3 dt - 4(x - y)^3 dt - cp_x dt + \sigma dW_1^t \\
\frac{dp_y}{dt} &= -4(y - x)^3 dt - cp_y dt + \sigma dW_2^t \\
\end{align*}
\] (5.11)

We compare several autocorrelation functions and time-dependent moments of this stochastic process integrated by FLAVORS (5.7) and (3.9) and Geometric Langevin Algorithm [14]. Quasi-symplectic FLAVOR (Subsection 5.1.1) and GLA gave results in agreement (Figure 19 [20(a), 20(b)]. Since Geometric Langevin Algorithm is weakly-convergent and Boltzmann-Gibbs preserving, this is a numerical evidence that quasi-symplectic FLAVOR is so too.

Expectations are empirically calculated by averaging over an ensemble of 100 sample trajectories. \( T = 30, \epsilon = 10^{-8}, \tau = 0.001, \delta = 0.01. \ x(0) = 2.1/\omega \) (with \( \omega := 1/\sqrt{\epsilon} \)), \( y(0) = x(0) + 1.8, c = 0.1, \sigma = 0.5. \) GLA uses timestep \( h = 0.001. \) Noise and friction
are slow here in the sense that they are not of the order $O(\omega)$ or larger.

As shown in the plots, in the regime dominated by deterministic dynamics (roughly from $T=0$ to $T=8$) various moments calculated empirically by FLAVORS and GLA are in agreement, indicating that the same rate of convergence towards Boltzmann-Gibbs distribution is obtained. And in that regime autocorrelation functions of the slow variable agree, serving as a numerical evidence that FLAVORS is weakly converging towards the SDE solution, whereas autocorrelation functions of the fast variable agree only in the sense of measures (after time averaging over a mesoscopic ($\langle 1 \rangle$) time span). The fluctuations between FLAVORS and GLA for large time are an effect of the finite number of samples (100) used to compute sample averages.

Recall that if the noise is applied to slow variables, FLAVORS don’t converge strongly but only in the sense of distributions.

5.3 Numerical implementation with fast noise and friction.

Consider a system with the same “topology” as above (Figure 2(a)). The difference is that the soft spring oscillates at a frequency nonlinearly dependent on the stiff spring’s length, and the left mass experiences strong friction and noise while the right mass doesn’t. The Hamiltonian is

$$\begin{align*}
H(x, y, p_x, p_y) &= \frac{1}{2} p_x^2 + \frac{1}{2} p_y^2/2 + \frac{1}{4} \omega^4 x^4 + e^x (y - x)^2 \\
&= H(x, y, p_x, p_y) \quad (5.12)
\end{align*}$$

and the governing SDEs are:

$$\begin{align*}
 dx &= dp_x \\
 dy &= dp_y \\
 dp_x &= -\omega^4 x^3 dt - (2 + x - y)(x - y)e^x dt - \omega^2 cp_x dt + \omega \sigma dW_t \\
 dp_y &= -2(y - x)e^x dt \\
\end{align*} \quad (5.13)$$
In this system, the deterministic dynamics and the effects of noise and friction both involve a $O(1/\omega^2)$ timescale. We have implemented the fast noise and friction version of FLAVORS ([5.8] and [5.9]).

In Figure 21, we have plotted the first and second moments of the slow variable $y(t) - x(t)$ as well as the first moment of the fast variable $x(t)$ as functions of time. Moments of the slow variable integrated by quasi-symplectic FLAVOR (Subsection 5.1.1) and GLA [14] concur, numerically suggesting weak convergence and preservation of Boltzmann-Gibbs. 100 fold computational acceleration is achieved.

Figure 21: $\mathbb{E}[y(t) - x(t)]$, $\mathbb{E}[x(t)]$, and $\mathbb{E}[(y(t) - x(t))^2]$ obtained by GLA and quasi-symplectic FLAVOR (Subsection 5.1.1). Expectations are empirically calculated by averaging over an ensemble of 50 sample trajectories. $T = 10$, $\omega = 100$, $\tau = 10^{-4}$, $\delta = 0.01$. $x(0) = 1.1/\omega$, $y(0) = x(0) + 1.8$, $c = 0.1$, $\sigma = 1$. GLA uses timestep $h = 10^{-4}$.

6 Appendix

6.1 Proof of Theorem 2.2

Define the process $t \mapsto (x_t, y_t)$ by

$$(\tilde{x}_t, \tilde{y}_t) := \eta(\tilde{u}_t)$$

(6.1)

It follows from the regularity of $\eta$ that it is sufficient to prove the $F$-convergence of $(\tilde{x}_t, \tilde{y}_t)$ towards $\delta X_t \otimes \mu(X_t, dy)$. Moreover it is also sufficient to prove the following
inequalities (6.2), (6.3) in order to obtain inequalities (2.17) and (2.18)

\[ |x_t^\epsilon - \bar{x}_t| \leq C\epsilon^C \psi_1(u_0, \epsilon, \delta, \tau) \]  

and

\[
\frac{1}{T} \int_0^{t+T} \varphi(\bar{x}_s, \bar{y}_s) \, ds - \int_{\mathbb{R}^p} \varphi(X_t, y) \mu(X_t, dy) \leq \psi_2(u_0, \epsilon, \delta, \tau, T, t)(\|\varphi\|_{L^\infty} + \|\nabla \varphi\|_{L^\infty})
\]  

(6.3)

Now define \( \psi_\tau^\epsilon \) by

\[ \psi_\tau^\epsilon(x, y) := \eta \circ \theta_\tau^\epsilon \circ \eta^{-1}(x, y) \]  

(6.4)

Define \( \psi_h^g \) by

\[ \psi_h^g(x, y) := \eta \circ \theta_h^G \circ \eta^{-1}(x, y) \]  

(6.5)

**Proposition 6.1.** The vector fields \( f \) and \( g \) associated with the system of equations (2.2) are Lipschitz continuous. We also have

\[ (\bar{x}_t, \bar{y}_t) = (\psi_h^g \circ \psi_\tau^\epsilon) \, (x_0, y_0) \quad \text{for} \quad k\delta \leq t \leq (k + 1)\delta \]  

(6.6)

Moreover there exists \( C > 0 \) such that for \( h \leq h_0 \) and \( \tau \epsilon \leq \tau_0 \) we have

\[ |\psi_\tau^\epsilon(x, y) - (x, y) - \tau(g(x, y), 0) - \frac{\tau}{\epsilon}(0, f(x, y))| \leq C(\frac{\tau}{\epsilon})^2 \]  

(6.7)

and

\[ |\psi_h^g(x, y) - (x, y) - h(g(x, y), 0)| \leq Ch^2 \]  

(6.8)

Furthermore, given \( x_0, y_0 \), the trajectories of \( (x_t^\epsilon, y_t^\epsilon) \), \( (\bar{x}_t, \bar{y}_t) \) are uniformly bounded in \( \epsilon, \delta \leq h_0 \), \( \tau \leq \min(\tau_0, \epsilon) \).

**Proof.** Since \( (x, y) = \eta(u) \), we have

\[ \dot{x} = (G + \frac{1}{\epsilon}F)\nabla \eta_x \circ \eta^{-1}(x, y) \]  

(6.9)

\[ \dot{y} = (G + \frac{1}{\epsilon}F)\nabla \eta_y \circ \eta^{-1}(x, y) \]  

(6.10)

Hence we deduce from equation (2.2) of Condition 2.1 that

\[ g(x, y) = G\nabla \eta_x \circ \eta^{-1}(x, y) \]  

(6.11)

\[ f(x, y) = F\nabla \eta_y \circ \eta^{-1}(x, y) \]  

(6.12)

we deduce the regularity of \( f \) and \( g \) from the regularity of \( G \), \( F \) and \( \eta \). Equation (6.6) is a direct consequence of the definition of \( \psi_\tau^\epsilon \) and \( \psi_h^g \) and equation (2.10) (we write \( (x_0, y_0) := \eta(u_0) \)). Observe that equation (2.2) of Condition 2.1 also requires that

\[ F\nabla \eta_x = 0 \quad G\nabla \eta_y = 0 \]  

(6.13)

48
Now observe that
\[ \psi^\epsilon_{\tau}(x, y) = (x, y) - \left( f(x, y), x, y \right) \tau - \left( g(x, y), 0 \right) \tau \epsilon \]
\[ = \left( \eta \circ \theta^\epsilon_{\tau} - \eta(x, y) - \tau(G \nabla \eta^\epsilon, 0) - \tau \epsilon \left( 0, F \nabla \eta^\epsilon \right) \right) \circ \eta^{-1}(x, y) \] (6.14)

Using (6.13), (2.8), Taylor expansion and the regularity of \( \eta \) we obtain (6.7). Similarly
\[ \psi^g_{\tau}(x, y) = (x, y) - h(g(x, y), 0) := \left( \eta \circ \theta^G_{\tau} - h \eta(x, y) - h(G \nabla \eta^x, 0) \right) \circ \eta^{-1}(x, y) \] (6.15)

Using (6.13), (2.7), Taylor expansion and the regularity of \( \eta \) we obtain (6.8). The uniform bound (depending on \( x_0, y_0 \)) on the trajectories of \( (x_{\epsilon t}, y_{\epsilon t}) \) and \( (\bar{x}_t, \bar{y}_t) \) is a consequence of the uniform bound (given \( u_0 \)) on the trajectories of \( u^\epsilon_t \) and \( \bar{u}_t \).

It follows from Proposition 6.1 that it is sufficient to prove Theorem 2.2 in the situation where \( \eta \) is the identity diffeomorphism. More precisely the \( F \)-convergence of \( \bar{u}_t \) is a consequence of the \( F \)-convergence of \( (\bar{x}_t, \bar{y}_t) \) and the regularity of \( \eta \). Furthermore from the uniform bound (depending on \( x_0, y_0 \)) on the trajectories of \( (x^\epsilon_t, y^\epsilon_t) \) and \( (\bar{x}_t, \bar{y}_t) \) we deduce that \( g \) and \( f \) are uniformly bounded and Lipschitz continuous (in \( \epsilon, \delta \leq h_0, \tau \leq \min(\tau_0 \epsilon, \delta) \)) over those trajectories.

Define
\[ \bar{g} := \int g(x, y) \mu(x, dy) \]
where \( \mu \) is the family of measures introduced in Condition 2.1. Let us prove the following lemma.

**Lemma 6.1.**

\[ |x^\epsilon_{n \delta} - \bar{x}_{n \delta}| \leq C e^{C n \delta} \left( \delta + \left( \frac{\tau}{\epsilon} \right)^{2 \frac{1}{\delta}} \sup_{1 \leq l \leq n} |J(l)| \right) \] (6.16)

with \( J(k) = J_1(k) + J_2(k) \),
\[ J_1(k) := \sum_{n=0}^{k-1} \left( \int_{n \delta}^{(n+1) \delta} g(x^\epsilon_{n \delta}, y^\epsilon_{s}) \, ds - \delta \bar{g}(x^\epsilon_{n \delta}) \right) \] (6.17)
and
\[ J_2(k) := \sum_{n=0}^{k-1} \delta \left( \bar{g}(x^\epsilon_{n \delta}) - g(\bar{x}_{n \delta}, \bar{y}_{n \delta}) \right) \] (6.18)

**Proof.** Observe that
\[ x_{(n+1) \delta}^\epsilon = x_{n \delta}^\epsilon + \int_{n \delta}^{(n+1) \delta} g(x^\epsilon_{n \delta}, y^\epsilon_{s}) \, ds + \int_{n \delta}^{(n+1) \delta} \left( g(x^\epsilon_{s}, y^\epsilon_{s}) - g(x^\epsilon_{n \delta}, y^\epsilon_{s}) \right) \, ds \] (6.19)
Hence
\[ x_{(n+1) \delta}^\epsilon - \bar{x}_{(n+1) \delta} = x_{n \delta}^\epsilon - \bar{x}_{n \delta} + I_1 + I_2(n) + I_3 + I_4(n) + I_5 \] (6.20)
with
\[ I_1 := \int_{n\delta}^{(n+1)\delta} (g(x_n^\epsilon, y_n^\epsilon) - g(x_{n\delta}^\epsilon, y_n^\epsilon)) \]  \hspace{1cm} (6.21)
\[ I_2(n) := \int_{n\delta}^{(n+1)\delta} g(x_{n\delta}^\epsilon, y_n^\epsilon) \, ds - \delta\bar{g}(x_{n\delta}^\epsilon) \]  \hspace{1cm} (6.22)
\[ I_3 := \delta(\bar{g}(x_{n\delta}^\epsilon) - \bar{g}(\bar{x}_{n\delta})) \]  \hspace{1cm} (6.23)
\[ I_4(n) := \delta(\bar{g}(\bar{x}_{n\delta}) - g(\bar{x}_{n\delta}, \bar{y}_{n\delta})) \]  \hspace{1cm} (6.24)
\[ I_5 := \delta\bar{g}(\bar{x}_{n\delta}, \bar{y}_{n\delta}) - (\bar{x}_{(n+1)\delta} - \bar{x}_{n\delta}) \]  \hspace{1cm} (6.25)

Now observe that
\[ |I_1| \leq \|\nabla_x g\|_{L^\infty} \|g\|_{L^\infty} \delta^2 \]  \hspace{1cm} (6.26)

Next
\[ |I_3| \leq \delta\|\nabla_x g\|_{L^\infty} |x_{n\delta}^\epsilon - \bar{x}_{n\delta}| \]  \hspace{1cm} (6.27)

Using (6.8), (6.7) we obtain that
\[ |I_5| \leq C(\delta^2 + \left(\frac{T}{\epsilon}\right)^2) \]  \hspace{1cm} (6.28)

Combining the previous equations, we have obtained that
\[ x_{(n+1)\delta}^\epsilon - \bar{x}_{(n+1)\delta} \leq x_{n\delta}^\epsilon - \bar{x}_{n\delta} + C\left(\delta^2 + \left(\frac{T}{\epsilon}\right)^2\right) + C\delta|x_{n\delta}^\epsilon - \bar{x}_{n\delta}| + (I_2 + I_4)(n) \]  \hspace{1cm} (6.29)

and
\[ x_{(n+1)\delta}^\epsilon - \bar{x}_{(n+1)\delta} \geq x_{n\delta}^\epsilon - \bar{x}_{n\delta} - C\left(\delta^2 + \left(\frac{T}{\epsilon}\right)^2\right) - C\delta|x_{n\delta}^\epsilon - \bar{x}_{n\delta}| + (I_2 + I_4)(n) \]  \hspace{1cm} (6.30)

Write
\[ J(n) := \sum_{k=0}^{n-1} (I_2 + I_4)(k) \]  \hspace{1cm} (6.31)

Adding up the first \( n \) inequalities (6.29) and (6.30)
\[ x_{n\delta}^\epsilon - \bar{x}_{n\delta} \leq C\left(\delta^2 + \left(\frac{T}{\epsilon}\right)^2\right)n + C\delta \sum_{k=0}^{n-1} |x_{k\delta}^\epsilon - \bar{x}_{k\delta}| + J(n) \]  \hspace{1cm} (6.32)
\[ x_{n\delta}^\epsilon - \bar{x}_{n\delta} \geq -C\left(\delta^2 + \left(\frac{T}{\epsilon}\right)^2\right)n - C\delta \sum_{k=0}^{n-1} |x_{k\delta}^\epsilon - \bar{x}_{k\delta}| + J(n) \]  \hspace{1cm} (6.33)

Hence
\[ |x_{n\delta}^\epsilon - \bar{x}_{n\delta}| \leq C\left(\delta^2 + \left(\frac{T}{\epsilon}\right)^2\right)n + C\delta \sum_{k=0}^{n-1} |x_{k\delta}^\epsilon - \bar{x}_{k\delta}| + |J(n)| \]  \hspace{1cm} (6.34)
And we obtain by induction

\[
|x_{n\delta}^\epsilon - \bar{x}_{n\delta}| \leq C\left(\delta^2 + \frac{(\tau)}{\epsilon^2}\right)\left(n + C\delta \sum_{k=1}^{n}(n-k)(1+C\delta)^{k-1}\right)
\]

(6.35)

\[
|J(n)| + C\delta \sum_{l=2}^{n-l} (1+C\delta)^{l-2} |J(n-l+1)|
\]

Equation (6.35) concludes the proof of Lemma 6.1.

We now need to control \( J_1(k) \) and \( J_2(k) \). First let us prove the following lemma.

**Lemma 6.2.** For \( N \in \mathbb{N}^* \) we have

\[
|J_1(k)| \leq (\delta k)C\left(\delta e^{\frac{\delta}{N}} + E\left(\frac{\delta}{N}\right)\right)
\]

(6.36)

**Proof.** Define \( \hat{y}_t^\epsilon \) such that \( \hat{y}_t^\epsilon = y_t^\epsilon \) for \( t = (n+j/N)\delta, \; j \in \mathbb{N}^* \), and

\[
\frac{d\hat{y}_t^\epsilon}{dt} = \frac{1}{\epsilon} f(x_{n\delta}^\epsilon, \hat{y}_t^\epsilon) \quad \text{for} \quad (n+j/N)\delta \leq t < (n+(j+1)/N)\delta
\]

(6.37)

Using the regularity of \( f \) and \( g \) we obtain that

\[
|\hat{y}_t^\epsilon - y_t^\epsilon| \leq C\delta e^{\frac{\delta}{N}}
\]

(6.38)

First observe that

\[
\frac{1}{\delta} \int_{n\delta}^{(n+1)\delta} g(x_{n\delta}^\epsilon, y_s^\epsilon) \, ds - \bar{g}(x_{n\delta}^\epsilon) = K_1 + K_2
\]

(6.39)

with

\[
K_1 := \frac{1}{\delta} \sum_{j=0}^{N-1} \int_{(n+j/N)\delta}^{(n+(j+1)/N)\delta} \left( g(x_{n\delta}^\epsilon, y_s^\epsilon) - g(x_{n\delta}^\epsilon, \hat{y}_s^\epsilon) \right) \, ds
\]

(6.40)

and

\[
K_2 := \frac{1}{N} \sum_{j=0}^{N-1} \left( \frac{N}{\delta} \int_{(n+j/N)\delta}^{(n+(j+1)/N)\delta} g(x_{n\delta}^\epsilon, \hat{y}_s^\epsilon) \, ds - \bar{g}(x_{n\delta}^\epsilon) \right)
\]

(6.41)

We have

\[
|K_1| \leq \|\nabla_x g\|_{L^\infty} \frac{1}{N} \sum_{j=0}^{N-1} \sup_{(n+j/N)\delta \leq s \leq (n+(j+1)/N)\delta} |y_s^\epsilon - \hat{y}_s^\epsilon|
\]

(6.42)

Hence, we obtain from (6.38) that

\[
|K_1| \leq C\delta e^{\frac{\delta}{N}}
\]

(6.43)
Moreover we obtain from properties 3 and 4 of Condition 2.1 that

\[
|K_2| \leq CE\left(\frac{\delta}{N\epsilon}\right)
\]  
(6.44)

This concludes the proof of Lemma 6.2.

Lemma 6.3. We have for \( m \in \mathbb{N}^* \)

\[
|J_2(k)| \leq C\delta k\left(m\delta + E\left(\frac{m\tau}{\epsilon}\right) + \left(\frac{\tau}{\epsilon} + m\delta + m(\frac{r}{\epsilon})^2\right)e^{C\frac{m\tau}{\epsilon}}\right)
\]  
(6.45)

Proof. Let \( m \in \mathbb{N}^* \). Define \((\tilde{x}_s, \tilde{y}_s)\) such that for \( j \in \mathbb{N}^*, n \in \mathbb{N}^*, \)

\[
\begin{align*}
\frac{d\tilde{x}_s}{dt} &= g(\tilde{x}_s, \tilde{y}_s) & \text{for} & \quad jm\delta \leq s < (j+1)m\delta \\
\frac{d\tilde{y}_s}{dt} &= \frac{1}{\epsilon}f(\tilde{x}_s, \tilde{y}_s) & \text{for} & \quad n\delta \leq s < n\delta + \tau \\
\tilde{y}_s &= \tilde{y}_{n\delta + \tau} & \text{for} & \quad n\delta + \tau \leq s < (n+1)\delta \\
\tilde{y}_{(n+1)\delta} &= \tilde{y}_{n\delta + \tau} & \text{for} & \quad (n+1)\delta \neq jm \\
(\tilde{x}_{jm}, \tilde{y}_{jm}) &= (\bar{x}_{jm\delta}, \bar{y}_{jm\delta})
\end{align*}
\]  
(6.46)

Define \( \tilde{y}_a^a \) by

\[
\begin{align*}
\frac{d\tilde{y}_a^a}{dt} &= \frac{1}{\epsilon}f(\bar{x}_{jm\delta}, \tilde{y}_a^a) & \text{for} & \quad jm\tau \leq t < (j+1)m\tau \\
\tilde{y}_{jm\tau}^a &= \tilde{y}_{jm\delta}
\end{align*}
\]  
(6.47)

Define \( \tilde{x}_n^a \) by

\[
\tilde{x}_n^a = \bar{x}_{jm\delta} \quad \text{for} \quad jm \leq n < (j+1)m
\]  
(6.48)

Observe that

\[
J_2(k) = K_3 + K_4 + K_5 + K_6 + K_7
\]  
(6.49)

with

\[
K_3 := \sum_{n=0}^{k-1} \int_{n\delta}^{(n+1)\delta} g(\bar{x}_s, \bar{y}_s) \, ds - \delta g(\bar{x}_{n\delta}, \bar{y}_{n\delta})
\]  
(6.50)

\[
K_4 := \sum_{n=0}^{k-1} \frac{1}{\tau} \int_{n\tau}^{(n+1)\tau} g(\tilde{x}_a^a, \tilde{y}_a^a) \, ds - \frac{1}{\delta} \int_{n\delta}^{(n+1)\delta} g(\bar{x}_s, \bar{y}_s) \, ds
\]  
(6.51)

\[
K_5 := \frac{\delta}{\tau} \sum_{n=0}^{k-1} \left( \tau \bar{g}(\tilde{x}_a^a) - \int_{n\tau}^{(n+1)\tau} g(\tilde{x}_a^a, \tilde{y}_a^a) \, ds \right)
\]  
(6.52)

\[
K_6 := \delta \sum_{n=0}^{k-1} \left( \bar{g}(\bar{x}_{n\delta}) - \bar{g}(\tilde{x}_n^a) \right)
\]  
(6.53)

Using the regularity of \( g \) we obtain that

\[
|K_6| \leq \delta kC \delta m
\]  
(6.54)
Assembling the right hand side of (6.51) into groups of \( m \) terms corresponding to the intervals of (6.47) we obtain from Property 3 of Condition 2.1 and Property 3 of Condition 2.2 that
\[
|K_5| \leq Ck\delta E\left(\frac{m\tau}{\epsilon}\right) \tag{6.55}
\]

Using (6.48) and the regularity of \( f \) and \( g \) we obtain the following inequality
\[
|\tilde{y}_{\tau\epsilon} - \tilde{y}_{\epsilon}| \leq Cm\delta e^{C\frac{m\tau}{\epsilon}} \tag{6.56}
\]

It follows that
\[
|K_4| \leq C\delta km\delta e^{C\frac{m\tau}{\epsilon}} \tag{6.57}
\]

Similarly using (6.8) and (6.7) we obtain the following inequalities
\[
|\tilde{y}_{n\delta} - \tilde{y}_{n\delta}| \leq C\left(\frac{\tau}{\epsilon} + m\delta + m\left(\frac{\tau}{\epsilon}\right)^2\right) e^{C\frac{m\tau}{\epsilon}} \tag{6.58}
\]
\[
|\tilde{x}_{n\delta} - \tilde{x}_{n\delta}| \leq Cm\left(\delta + \left(\frac{\tau}{\epsilon}\right)^2\right) \tag{6.59}
\]

It follows that
\[
|K_3| \leq C\delta k\left(\frac{\tau}{\epsilon} + m\delta + m\left(\frac{\tau}{\epsilon}\right)^2\right) e^{C\frac{m\tau}{\epsilon}} \tag{6.60}
\]

This concludes the proof of Lemma 6.3. \( \square \)

Combining Lemma 6.1, 6.2 and 6.3 we have obtained that
\[
|x^\epsilon_{n\delta} - \tilde{x}_{n\delta}| \leq Ce^{C\delta n}\left(\delta + \left(\frac{\tau}{\epsilon}\right)^2\frac{1}{\delta} + \delta e^{C\frac{\delta}{N\epsilon}} + E\left(\frac{\delta}{N\epsilon}\right) + E\left(\frac{m\tau}{\epsilon}\right)\right)
\]
\[
+ \left(\frac{\tau}{\epsilon} + m\delta + m\left(\frac{\tau}{\epsilon}\right)^2\right) e^{C\frac{m\tau}{\epsilon}} \tag{6.61}
\]

Choosing \( N \) such that \( e^{C\frac{\delta}{N\epsilon}} \sim \delta^{-\frac{1}{2}} \) (observe that we need \( \epsilon \leq \delta/(C\ln \delta) \)) and \( m \) such that \( \frac{m\tau}{\epsilon} e^{C\frac{m\tau}{\epsilon}} \sim \left(\frac{\delta}{\tau} + \frac{\tau}{\epsilon}\right)^{-\frac{1}{2}} \) we obtain for \( \frac{\delta}{\tau} + \frac{\tau}{\epsilon} \leq 1 \) that
\[
|x^\epsilon_{n\delta} - \tilde{x}_{n\delta}| \leq Ce^{C\delta n}\left(\sqrt{\delta} + \left(\frac{\tau}{\epsilon}\right)^2\frac{1}{\delta} + E\left(\frac{1}{C\ln \frac{1}{\delta}}\right) \right.
\]
\[
+ \left(\frac{\delta}{\tau}\right)^{\frac{1}{2}} \left(\frac{\tau}{\epsilon}\right)^{\frac{1}{2}} + E\left(\frac{1}{C\ln \left(\frac{\delta}{\tau} + \frac{\tau}{\epsilon}\right)^{-1}}\right) \tag{6.62}
\]

This concludes the proof of inequality (6.2). The proof of (6.3) is similar and is also a consequence of (6.2).

53
6.2 Proof of Theorem 4.2

Define the process \( t \mapsto (\bar{x}_t, \bar{y}_t) \) by

\[
(\bar{x}_t, \bar{y}_t) := \eta(\bar{u}_t) \tag{6.63}
\]

It follows from the regularity of \( \eta \) that it is sufficient to prove the \( F \)-convergence of

\( (\bar{x}_t, \bar{y}_t) \) towards \( \delta_{X_t} \otimes \mu(X_t, dy) \). Now define \( \psi_{\epsilon} \) by

\[
\psi_{\epsilon}^\eta(x, y, \omega) := \eta \circ \theta_{\epsilon}^\eta(\cdot, \omega) \circ \eta^{-1}(x, y) \tag{6.64}
\]

Define \( \psi_{\eta}^\omega \) by

\[
\psi_{\eta}^\omega(x, y, \omega) := \eta \circ \theta_{\eta}^\omega(\cdot, \omega) \circ \eta^{-1}(x, y) \tag{6.65}
\]

**Proposition 6.2.** The vector fields \( f, g \) and matrix fields \( \sigma, Q \) associated with the system of equations \([4.3]\) are uniformly bounded and Lipschitz continuous. We also have

\[
\begin{align*}
\left( \mathbb{E} \left[ |\psi_{\eta}^\omega(x, y, \omega) - (x, y) - h(g(x, y), 0) - \sqrt{h}(\sigma(x, y)\xi'(\omega), 0)|^2 \right] \right)^{\frac{1}{2}} & \leq Ch\frac{1}{\epsilon} \tag{6.66}
\end{align*}
\]

\[
\begin{align*}
\left( \mathbb{E} \left[ |\psi_{\epsilon}^\eta(x, y, \omega) - (x, y) - \tau(g(x, y), 0) - \sqrt{\tau}(\sigma(x, y)\xi''(\omega), 0) \right. \right. \\
\left. \left. \quad - \sqrt{\frac{\tau}{\epsilon}}(0, Q(x, y)\xi'''(\omega))|^2 \right| \right)^{\frac{1}{2}} & \leq C \left( \frac{\tau}{\epsilon} \right)^{\frac{3}{2}} \tag{6.67}
\end{align*}
\]

**Proof.** Since \( (x, y) = \eta(u) \), we obtain from \([4.1]\) and Ito’s formula

\[
dx = \left( (G + \frac{1}{\epsilon} F)\nabla \eta^x \circ \eta^{-1}(x, y) \right) dt + \left( \nabla \eta^x (H + \frac{1}{\sqrt{\epsilon}} K) \right) \circ \eta^{-1}(x, y) dW_t
\]

\[
\quad + \frac{1}{2} \sum_{ij} \partial_i \partial_j \eta^x ((H + \frac{1}{\sqrt{\epsilon}} K)(H + \frac{1}{\sqrt{\epsilon}} K)^T)_{ij} dt \tag{6.68}
\]

\[
dy = \left( (G + \frac{1}{\epsilon} F)\nabla \eta^y \circ \eta^{-1}(x, y) \right) dt + \left( \nabla \eta^y (H + \frac{1}{\sqrt{\epsilon}} K) \right) \circ \eta^{-1}(x, y) dW_t
\]

\[
\quad + \left( \frac{1}{2} \sum_{ij} \partial_i \partial_j \eta^y ((H + \frac{1}{\sqrt{\epsilon}} K)(H + \frac{1}{\sqrt{\epsilon}} K)^T)_{ij} \right) \circ \eta^{-1} dt \tag{6.69}
\]
Hence we deduce from equation (4.3) of Condition 4.1 that
\[
g(x, y) = \left( G \nabla \eta^x + \frac{1}{2} \sum_{ij} \partial_i \partial_j \eta^x (HH)^{ij} \right) \circ \eta^{-1}(x, y) \quad (6.71)
\]
\[
\sigma(x, y) = \left( \nabla \eta^x H \right) \circ \eta^{-1}(x, y) \quad (6.72)
\]
\[
f(x, y) = \left( F \nabla \eta^y + \frac{1}{2} \sum_{ij} \partial_i \partial_j \eta^y (KK)^{ij} \right) \circ \eta^{-1}(x, y) \quad (6.73)
\]
\[
Q(x, y) = \left( \nabla \eta^y K \right) \circ \eta^{-1}(x, y) \quad (6.74)
\]

**Remark 6.1.** Observe that equation (4.3) of Condition 4.1 requires that
\[
F \nabla \eta^x = 0 \quad (6.75)
\]
\[
\sum_{ij} \partial_i \partial_j \eta^x (KK)^{ij} = 0 \quad (6.76)
\]
\[
\sum_{ij} \partial_i \partial_j \eta^y (HH)^{ij} = 0 \quad (6.77)
\]
\[
\sum_{ij} \partial_i \partial_j \eta^x (KK)^{ij} = 0 \quad (6.78)
\]
\[
\sum_{ij} \partial_i \partial_j \eta^y (KK)^{ij} = 0 \quad (6.79)
\]

In particular, equations (6.78) and (6.79) are satisfied if \( KH^T \) which translates into the fact that for all \( u \) the ranges of \( H(u) \) and \( K(u) \) are orthogonal, i.e. the noise with amplitude \( 1/\sqrt{\epsilon} \) is applied to degrees of freedom orthogonal to those with \( O(1) \) noise.

We deduce the regularity of \( f, g, \sigma \) and \( Q \) from the regularity of \( G, F, H, K \) and \( \eta \). Equation (6.6) is a direct consequence of the definition of \( \psi^x_\epsilon \) and \( \psi^y_\epsilon \) and equation (6.66). Now observe that
\[
\psi^x_\epsilon(x, y, \omega) - (x, y) - \tau (g(x, y), 0) - \frac{\tau}{\epsilon} (0, f(x, y)) - \sqrt{\tau} (\sigma(x, y) \xi'(\omega), 0)
\]
\[
- \frac{\tau}{\epsilon} (0, Q(x, y) \xi'(\omega)) = \left( \eta \circ \theta^x_\epsilon - \eta - \tau (G \nabla \eta^x + \frac{1}{2} \sum_{ij} \partial_i \partial_j \eta^x (HH)^{ij}, 0) \right)
\]
\[
- \frac{\tau}{\epsilon} (0, F \nabla \eta^y + \frac{1}{2} \sum_{ij} \partial_i \partial_j \eta^y (KK)^{ij}) - \sqrt{\tau} (\nabla \eta^x H \xi'(\omega), 0)
\]
\[
- \sqrt{\tau} (0, \nabla \eta^y \xi'(\omega)) \circ \eta^{-1}(x, y) \quad (6.80)
\]

Using equations (6.75), (6.76), (6.77), (6.78) and (6.79), the Taylor-Ito expansion of \( \eta \circ \theta^x_\epsilon \), the regularity of \( \eta \), and Setting \( \xi' \) equal to \( \xi \) defined in equation (4.10) we obtain equation (6.68). The proof of equation (6.67) is similar. \( \square \)
It follows from Proposition 6.2 that it is sufficient to prove Theorem 4.2 in the situation where $\eta$ is the identity diffeomorphism. More precisely, the $F$-convergence of $\bar{u}_t$ is a consequence of the $F$-convergence of $(\bar{x}_t, \bar{y}_t)$ and the regularity of $\eta$.

Let $x \mapsto \varphi(x)$ be a function with continuous and bounded derivatives up to order 3. Let us prove the following lemma.

**Lemma 6.4.** We have

\[
\mathbb{E}[\varphi(\bar{x}_{(n+1)\delta})] - \mathbb{E}[\varphi(\bar{x}_{n\delta})] = \\
\delta \mathbb{E}\left[g(\bar{x}_{n\delta}, \bar{y}_{n\delta})\nabla \varphi(\bar{x}_{n\delta}) + \sigma T(\bar{x}_{n\delta}, \bar{y}_{n\delta}) : \text{Hess} \varphi(\bar{x}_{n\delta})\right] + I_0
\]

with

\[
|I_0| \leq C\left(\delta^2 + \left(\frac{\tau}{\epsilon}\right)^3\right)
\]

**Proof.** Write $(\bar{x}_{n\delta+\tau}, \bar{y}_{n\delta+\tau}) := \psi_\epsilon^{(n)}(\bar{x}_{n\delta}, \bar{y}_{n\delta}, \omega_n)$. Using equation (6.68) we obtain that there exists an $\mathcal{N}(0, 1)$ random vector $\xi_n$ independent from $(\bar{x}_{n\delta}, \bar{y}_{n\delta})$ and such that

\[
\bar{x}_{n\delta+\tau} - \bar{x}_{n\delta} = g(\bar{x}_{n\delta}, \bar{y}_{n\delta})\tau + \sqrt{\tau}\sigma(\bar{x}_{n\delta}, \bar{y}_{n\delta})\xi_n + I_1
\]

with

\[
\left(\mathbb{E}[(I_1)^2]\right)^{\frac{1}{2}} \leq C\left(\frac{\tau}{\epsilon}\right)^{\frac{3}{2}}
\]

Hence

\[
\left|\mathbb{E}[\varphi(\bar{x}_{n\delta+\tau})] - \mathbb{E}[\varphi(\bar{x}_{n\delta})] - \tau \mathbb{E}\left[g(\bar{x}_{n\delta}, \bar{y}_{n\delta})\nabla \varphi(\bar{x}_{n\delta})
\right.ight.

\[
\left. + \sigma T(\bar{x}_{n\delta}, \bar{y}_{n\delta}) : \text{Hess} \varphi(\bar{x}_{n\delta})\right]\right| \leq C\left(\frac{\tau}{\epsilon}\right)^{\frac{3}{2}}
\]

Similarly, using equation (6.67) we obtain that there exists an $\mathcal{N}(0, 1)$ random vector $\xi_n'$, independent from $(\bar{x}_{n\delta+\tau}, \bar{y}_{n\delta+\tau})$, and such that

\[
\bar{x}_{(n+1)\delta} - \bar{x}_{n\delta+\tau} = g(\bar{x}_{n\delta+\tau}, \bar{y}_{n\delta+\tau})(\delta - \tau) + \sigma(\bar{x}_{n\delta+\tau}, \bar{y}_{n\delta+\tau})\sqrt{\delta - \tau}\xi_n' + I_2
\]

with

\[
\left(\mathbb{E}[(I_2)^2]\right)^{\frac{1}{2}} \leq C(\delta - \tau)^{\frac{3}{2}}
\]

Whence

\[
\left|\mathbb{E}[\varphi(\bar{x}_{(n+1)\delta})] - \mathbb{E}[\varphi(\bar{x}_{n\delta+\tau})] - (\delta - \tau)\mathbb{E}\left[g(\bar{x}_{n\delta+\tau}, \bar{y}_{n\delta+\tau})\nabla \varphi(\bar{x}_{n\delta+\tau})
\right.ight.

\[
\left. + \sigma T(\bar{x}_{n\delta+\tau}, \bar{y}_{n\delta+\tau}) : \text{Hess} \varphi(\bar{x}_{n\delta+\tau})\right]\right| \leq C(\delta - \tau)^{\frac{3}{2}}
\]
Using the regularity of \(\sigma\) we obtain that

\[
\left( \mathbb{E} \left[ |\sigma(\bar{x}_{n+\tau}, \bar{y}_{(n+1)\delta}) - \sigma(\bar{x}_{n\delta}, \bar{y}_{n\delta})|^2 \right] \right)^{\frac{1}{2}} \leq C(\delta^\frac{1}{2} + \sqrt{\tau}) \tag{6.89}
\]

The proof of (6.81) follows from \(6.85, 6.88, 6.89, 6.68\) and the regularity of \(g\) and \(\varphi\).

**Lemma 6.5.** We have

\[
\left| \frac{\mathbb{E}[\varphi(\bar{x}_{n\delta})] - \varphi(x_0)}{n\delta} - L\varphi(x_0) \right| \leq J_5 \tag{6.90}
\]

with (for \(\delta \leq C\tau/\epsilon\))

\[
|J_5| \leq C\left( (\frac{\delta\epsilon}{\tau})^\frac{1}{2} + (\frac{\tau}{\epsilon})^\frac{1}{2} + \frac{1}{\sqrt{\epsilon}} \right) + CE\left( \frac{1}{\nu} \ln \frac{\tau}{\delta\epsilon} \right) \tag{6.91}
\]

**Proof.** Define \(\tilde{B}_t\) by \(\tilde{B}_0 = 0\) and

\[
\tilde{B}_t - \tilde{B}_n\tau = B_{n\delta+t} - B_{n\delta} \quad \text{for} \quad n\tau \leq t \leq (n+1)\tau \tag{6.92}
\]

Define \(\tilde{y}_s\) by \(\tilde{y}_0 = y_0\) and

\[
d\tilde{y}_t = \frac{1}{\epsilon} f(x_0, \tilde{y}_t) dt + \frac{1}{\sqrt{\epsilon}} Q(x_0, \tilde{y}_t) d\tilde{B}_t \tag{6.93}
\]

Write

\[
\bar{g}(x_0) := \int g(x_0, y) \mu(x_0, dy) \tag{6.94}
\]

Using Lemma 6.4 we obtain that

\[
\frac{\mathbb{E}[\varphi(\bar{x}_{n\delta})] - \varphi(x_0)}{n\delta} = L\varphi(x_0) + J_1 + J_2 + J_3 + J_4 \tag{6.95}
\]

with

\[
L\varphi(x_0) := \bar{g}(x_0)\nabla\varphi(x_0) + \bar{\sigma}\bar{\sigma}^T(x_0) : \text{Hess}\ \varphi(x_0) \tag{6.96}
\]

\[
J_1 = \frac{1}{n} \sum_{k=0}^{n-1} \mathbb{E} \left[ g(\bar{x}_{k\delta}, \bar{y}_{k\delta})\nabla\varphi(\bar{x}_{k\delta}) + \sigma\sigma^T(\bar{x}_{k\delta}, \bar{y}_{k\delta}) : \text{Hess}\ \varphi(\bar{x}_{k\delta}) \right] \tag{6.97}
\]

\[
- \frac{1}{n} \sum_{k=0}^{n-1} \mathbb{E} \left[ g(\bar{x}_0, \bar{y}_{k\delta})\nabla\varphi(\bar{x}_0) + \sigma\sigma^T(\bar{x}_0, \bar{y}_{k\delta}) : \text{Hess}\ \varphi(\bar{x}_0) \right] \tag{6.98}
\]

\[
J_2 = \frac{1}{n} \sum_{k=0}^{n-1} \mathbb{E} \left[ g(\bar{x}_0, \bar{y}_{k\delta})\nabla\varphi(\bar{x}_0) + \sigma\sigma^T(\bar{x}_0, \bar{y}_{k\delta}) : \text{Hess}\ \varphi(\bar{x}_0) \right] \tag{6.98}
\]

\[
- \frac{1}{\tau} \int_{k\tau}^{(k+1)\tau} \mathbb{E} \left[ g(x_0, \tilde{y}_s)\nabla\varphi(x_0) + \sigma\sigma^T(x_0, \tilde{y}_s) : \text{Hess}\ \varphi(x_0) \right] ds \]
\[ J_3 = \frac{1}{n\tau} \int_0^{n\tau} \mathbb{E}\left[ g(x_0, \tilde{y}_s) \nabla \varphi(x_0) + \sigma \sigma^T(x_0, \tilde{y}_s) : \text{Hess} \varphi(x_0) \right] ds - L\varphi(x_0) \]  

(6.99)

\[ |J_4| \leq C \left( \delta^2 + \left( \frac{\tau}{\epsilon} \right)^{\frac{3}{2}} \frac{1}{\delta} \right) \]  

(6.100)

Using the regularity of \( \sigma, g, \varphi \), (6.6) and (6.7) we obtain that

\[ |J_1| \leq C \left( (n\delta)^{\frac{1}{2}} + n\delta + n \left( \frac{\tau}{\epsilon} \right)^{\frac{3}{2}} \right) \]  

(6.101)

Using Property 3 of Condition 4.1 and Property 3 of Condition 4.2 we obtain that

\[ |J_3| \leq C \left( \sqrt{\frac{\tau}{\epsilon}} + \left( n\delta \right)^{\frac{1}{2}} + n\delta + n \left( \frac{\tau}{\epsilon} \right)^{\frac{3}{2}} \right) e^{Cn\frac{\tau}{\epsilon}} \]  

(6.102)

Using (6.67) and (6.68), we obtain the following inequality

\[ \left( \mathbb{E}\left[ \left| \tilde{y}_{n\delta} - \tilde{y}_{n\tau} \right|^2 \right] \right)^{\frac{1}{2}} \leq C \left( \sqrt{\frac{\tau}{\epsilon}} + (n\delta)^{\frac{1}{2}} + n\delta + n \left( \frac{\tau}{\epsilon} \right)^{\frac{3}{2}} \right) \frac{n\tau}{\epsilon} e^{Cn\frac{\tau}{\epsilon}} \]  

(6.103)

which leads to

\[ |J_2| \leq C \left( \sqrt{\frac{\tau}{\epsilon}} + (n\delta)^{\frac{1}{2}} + n\delta + n \left( \frac{\tau}{\epsilon} \right)^{\frac{3}{2}} \right) e^{Cn\frac{\tau}{\epsilon}} \]  

(6.104)

Hence we have obtained that

\[ \left| \mathbb{E}\left[ \varphi(\bar{x}_{n\delta}) \right] - \varphi(x_0) \right| \leq J_5 \]  

(6.105)

with

\[ |J_5| \leq C \left( \sqrt{\frac{\tau}{\epsilon}} + (n\delta)^{\frac{1}{2}} + n\delta + n \left( \frac{\tau}{\epsilon} \right)^{\frac{3}{2}} \right) e^{Cn\frac{\tau}{\epsilon}} + \mathbb{E}\left[ \frac{n\tau}{\epsilon} \right] + C \left( \frac{\tau}{\epsilon} \right)^{\frac{3}{2}} \frac{1}{\delta} \]  

(6.106)

Choosing \( n \) such that \( \sqrt{\frac{n\tau}{\epsilon}} e^{Cn\frac{\tau}{\epsilon}} \sim \left( \frac{\tau}{\delta^2} \right)^{\frac{1}{2}} \) we obtain (6.91) for \( \delta \leq C\tau/\epsilon \).

We now combine Lemma 6.5 with Theorem 1 of Chapter 2 of [82] which states that the uniform convergence (in \( x_0, y_0 \)) of \( \mathbb{E}\left[ \varphi(\bar{x}_{n\delta}) \right] \) towards \( L\varphi(x_0) \) as \( \epsilon \downarrow 0 \), \( \tau \leq \delta \), \( \frac{x}{\tau} \downarrow 0 \), \( \frac{\delta}{\tau} \downarrow 0 \) and \( \left( \frac{\tau}{\delta} \right)^{\frac{3}{2}} \frac{1}{\delta} \downarrow 0 \) implies the convergence in distribution of \( \bar{x}_{n\delta} \) towards the Markov process generated by \( L \).

The \( F \)-convergence of \( (\bar{x}_t, \bar{y}_t) \) can be deduced from the convergence in distribution of \( \bar{x}_t \) and equation (4.5) of Condition 4.1. The proof follows the same lines as above which will not be repeated here.
References


