

Limit theorems for infinite-dimensional piecewise deterministic Markov processes. Applications to stochastic excitable membrane models*

Martin G. Riedler[†] Michèle Thieullen[‡] Gilles Wainrib[§]

Abstract

We present limit theorems for a sequence of Piecewise Deterministic Markov Processes (PDMPs) taking values in a separable Hilbert space. This class of processes provides a rigorous framework for stochastic spatial models in which discrete random events are globally coupled with continuous space-dependent variables solving partial differential equations, e.g., stochastic hybrid models of excitable membranes. We derive a law of large numbers which establishes a connection to deterministic macroscopic models and a martingale central limit theorem which connects the stochastic fluctuations to diffusion processes. As a prerequisite we carry out a thorough discussion of Hilbert space valued martingales associated to the PDMPs. Furthermore, these limit theorems provide the basis for a general Langevin approximation to PDMPs, i.e., stochastic partial differential equations that are expected to be similar in their dynamics to PDMPs. We apply these results to compartmental-type models of spatially extended excitable membranes. Ultimately this yields a system of stochastic partial differential equations which models the internal noise of a biological excitable membrane based on a theoretical derivation from exact stochastic hybrid models.

Keywords: Piecewise Deterministic Markov Processes; infinite-dimensional stochastic processes; law of large numbers; central limit theorem; excitable membrane models; random excitable media.

AMS MSC 2010: 60J25; 60B12; 60F05; 92C20.

Submitted to EJP on April 12, 2012, final version accepted on July 12, 2012.

1 Introduction

In this study we present limit theorems for sequences of Piecewise Deterministic Markov Processes (PDMPs) with values in a separable Hilbert space. PDMPs are a particular class of càdlàg, strong Markov processes which combine continuous deterministic time evolution and discontinuous, instantaneous, random ‘jump’ events originally introduced in [15, 16, 45] in filtering theory and stochastic control problems. We note that in view of applications this paper is ultimately motivated by the interest in the

*Supported by ANR MANDy “Mathematical Analysis of Neuronal Dynamics” ANR-09-BLAN-0008.

[†]Johannes Kepler Universität Linz, Austria. E-mail: martin.riedler@jku.at

[‡]Université Pierre et Marie Curie, France. E-mail: michele.thieullen@upmc.fr

[§]Université Paris 13, France. E-mail: wainrib@math.univ-paris13.fr

derivation of a justifiable Langevin approximation to spatio-temporal stochastic hybrid models of excitable membranes, e. g., neuronal membranes. This is accomplished by the limit theorems we present in the following.

We start briefly introducing the general idea of our framework and the main results which are made precise in the subsequent sections. We consider a family of fully coupled, Hilbert space-valued PDMPs indexed by $n \in \mathbb{N}$. Here fully coupled means that the PDMPs which split into a continuously moving and a piecewise constant component are such that the jump rates of the processes depend on the state of the full system and the continuous dynamics depend on the state of the jump component. For the limit theorems we rely on two key assumptions. Firstly, jumps possess heights decreasing to zero for $n \rightarrow \infty$ but occur at an increasing frequency roughly inversely proportional to the jump size. We are therefore in the *fluid limit* setting, cf. [31, 32]. Secondly, we assume that for each n the continuous dynamics in between jumps depend on the piecewise constant component only via a *finite* set of (Hilbert space-valued) functions thereof, which we call *coordinate functions*. It is the sequence of coordinate functions coupled to the continuous component for which we derive limits. The first limit theorem we present is a weak law of large numbers for PDMPs in infinite-dimensional Hilbert spaces where the deterministic limit is given by a solution of an abstract evolution equation. Next we proceed to the presentation of a central limit theorem for the martingales associated with a PDMP. This central limit theorem gives the basis for an approximation of PDMPs by diffusion processes which are solutions of stochastic partial differential equations. Finally, we show how to represent the stochastic process arising as the limit in the central limit theorem as a solution of a stochastic partial differential equation (SPDE) which then yields a Langevin approximation for PDMPs by a system of SPDEs. The new results presented extend previous results for PDMPs and pure jump processes in Euclidean space [32, 18, 37]. The difficulties in extending the fluid limit theorems in [31, 32, 37] to processes taking values in infinite-dimensional Hilbert spaces lie, on the one hand, in the appropriate treatment of Hilbert space-valued martingales. These arise by splitting a PDMP, being a semi-martingale, into a sum of a part with finite variation and a local martingale. As these considerations are essential we have devoted a full section, Section 3, to the discussion of the martingales. On the other hand, the more intricate existence theory of solutions to abstract evolution equations compared to solutions of ordinary differential equations in Euclidean space demands for additional technical rigour.

We apply our theoretical findings to spatially extended hybrid models of excitable membranes. A first hybrid formulation of one such model in the context of neuroscience was presented in [5] and reformulated and extended as examples for PDMPs taking values in infinite-dimensional Hilbert spaces in [13]. For example, the Hodgkin-Huxley model is a deterministic, macroscopic model for the coupled evolution of the neuronal membrane potential and the averaged gating dynamics of ion channels [23]. More realistically, the membrane potential, which is the macroscopically observed variable of interest, arises from the stochastic dynamics of finitely many ion channels. Thus the application of our limit theorems shows that the Hodgkin-Huxley is obtained as the limit of a sequence of stochastic microscopic models taking the form of Hilbert space valued PDMPs in the sense of a law of large numbers. Conceptually, here the fluid limit corresponds to increasing the number of ion channels while simultaneously decreasing the individual influence of an individual channel on the total current. The martingale central limit theorem can then be used to define the Langevin approximation providing a relatively simple stochastic version of the Hodgkin-Huxley model incorporating internal fluctuations.

Concluding this introduction, we comment on related work to fluid limits in the

infinite-dimensional setting. Averaging for PDMPs in infinite-dimension, in particular for the neuron model introduced in [5], wherein also a law of large numbers was considered, has been recently considered in [22]. For a model of linear chemical reactions by jump Markov processes a law of large numbers [4] and a central limit theorem [27] have been proven based on the original work of [31, 32] for finite-dimensional jump-processes. In these cases the deterministic limit is a reaction-diffusion partial differential equation and the central limit theorem yields diffusion processes given by stochastic partial differential equations. Limit theorems for variations of this model have been investigated in two series of studies, cf. [28, 29, 30] and [7, 8, 9, 10, 11]. A central difference between spatial models of excitable media to models of chemical reactions is that the latter exhibit diffusive motion of the reactants (\sim channels) which is absent in the former. Additionally excitable media equations exhibit non-local interaction of channels as their dynamics are coupled globally via the membrane potential. The limit theorems we establish have to account for these differences. Further, there is also a difference on the technical side. The technique employed in [27] and in all subsequent publications cited above is based on the semigroup approach to stochastic / deterministic evolution equations. In contrast, we pursue in the present paper the approach of a weak formulation. At large, the weak formulation of evolution problems allows to consider more general equations as when dealing with mild, strong or classical solutions, cf. a discussion of this aspect in [46, Chap. 23.1]. Finally, we also mention a central limit theorem for Hilbert-valued martingales in [36] and a diffusion approximation of SPDEs on nuclear spaces driven by Poisson random measures in [25]. The methods of proof we employ for the theoretical results in this study are motivated by the two last references, but differ as the classes of stochastic processes considered therein and in the present manuscript are different.

The remainder of the paper is organised as follows. We first briefly define PDMPs in Section 2 and precisely state the structure for a sequence of such PDMPs to allow for a limit. Then we discuss in detail the associated martingale process in Section 3. Limit theorems and the diffusion approximation are presented in Sections 4 and 5. We have deferred the proofs of the main results to Section 6. Next in Section 7 we discuss applications of these limit theorems to compartmental-type models of excitable membranes where the proofs of the conditions are deferred to Appendix B. The paper is concluded in Section 8 with a brief discussion and an outlook on further developments and applications. Finally, the Appendix A of the paper contains the proof of the technical Theorem 3.1 that guarantees the square-integrability of the associated Hilbert space valued martingales and establishes an appropriate Itô-isometry.

2 Piecewise Deterministic Markov Processes

In the first subsection we briefly define PDMPs and, in particular, discuss the specific subclass of PDMPs for which we present limit theorems in this study. For a general discussion of PDMPs we refer to the monographs [17, 24] and, specifically, for Hilbert space valued PDMPs associated to solutions of partial differential equations we refer to [13, 42]. In the second subsection we present the sequence of PDMPs for which the limits are analyzed in this study. Finally, a notational remark: in this paper pairings (\cdot, \cdot) and $\langle \cdot, \cdot \rangle$ denote the inner product or the duality pairing, respectively, with respect to a certain Hilbert space which is usually indicated with a subindex. Further, $*$ is used to denote dual spaces.

2.1 PDMPs on Hilbert spaces

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ denote a filtered probability space satisfying the usual conditions, $X \subset H \subset X^*$ be an evolution triple of separable real Hilbert spaces and K be a countable set of isolated states. The product $H \times K$ serves as the state space for a PDMP. Then, a PDMP is a càdlàg strong Markov process $X_t(\omega) = (U_t(\omega), \Theta_t(\omega)) \in H \times K$ for all $t \geq 0$ which consists of two components. The first, U_t , takes values in H , possesses continuous sample paths and is denoted the *continuous component* of the PDMP. The second, Θ_t , taking values in K and possessing right-continuous, piecewise constant sample paths, we call its *jump component*. We say a PDMP is *regular* if the number of jumps of Θ_t is a.s. finite in every finite time interval $[0, T]$. In this study PDMPs are always regular.

We next state the mechanisms which govern the time evolution of the paths of such a PDMP. Firstly, there exist for each $\theta \in K$ an abstract evolution equation

$$\dot{u} = A(\theta)u + B(\theta, u) \tag{2.1}$$

where $A(\theta) : X \rightarrow X^*$ is a linear and $B(\theta, \cdot) : X \rightarrow X^*$ a (possibly nonlinear) operator. We assume that the family of abstract evolution equations (2.1) is well-posed, i.e., given any $\theta \in K$ and any initial condition $u \in H$ there exists a unique global weak solution $\phi(\cdot, (u, \theta)) \in L^2((0, T), X) \cap H^1((0, T), X^*)$ depending continuously on the initial condition. Note, that the regularity implies $\phi(\cdot, (u, \theta)) \in C([0, T], H)$, cf. [40, Chap. 11]. Then the trajectory of the continuous component U_t follows in between jumps of the jump component Θ_t the weak solution to (2.1) corresponding to the parameter θ given by the current state of the jump component. That is, for $\tau_k, k \in \mathbb{N}$, denoting the jump times of the PDMP we have that

$$U_t = \phi(t - \tau_k, (U_{\tau_k}, \Theta_{\tau_k})) \quad \forall t \in [\tau_k, \tau_{k+1}).$$

Secondly, describing the stochastic transition dynamics of the jump component Θ_t there exist measurable transition rates $\Lambda : H \times K \rightarrow \mathbb{R}_+$ that define the distributions of the random jump time of Θ_t in the sense that for all $\theta \in K$

$$\mathbb{P}[\Theta_{t+s} = \theta, 0 \leq s \leq \Delta t \mid \Theta_t] = \exp\left(-\int_0^{\Delta t} \Lambda(U_{t+s}, \Theta_t) ds\right). \tag{2.2}$$

In view of (2.2) we assume that Λ is integrable along the solutions of (2.1) on any finite time interval, i.e.,

$$\int_0^T \Lambda(\phi(t, (u, \theta)), \theta) dt < \infty \quad \forall T < \infty$$

for all $\theta \in K$ and all initial conditions $u \in H$, but diverging as $T \rightarrow \infty$. We note that in applications we usually find that the transition rate Λ is bounded which implies the regularity of the PDMP. Finally, there exists a Markov kernel μ on $H \times K$ into K that gives the distribution of the post jump value, i.e.,

$$\mathbb{P}[\Theta_t = \xi \mid \Theta_t \neq \Theta_{t-}] = \mu((U_t, \Theta_{t-}), \{\xi\}) \quad \forall \xi \in K. \tag{2.3}$$

The elements of the quadruple (A, B, Λ, μ) are called the *characteristics* of the process and under the above conditions define a regular PDMP uniquely (up to versions). Furthermore, under these conditions the following result characterising the extended generator of PDMPs is proven in [13, 42].

Theorem 2.1. *A function $f : H \times K \rightarrow \mathbb{R}$ is in the domain of the extended generator of a PDMP if the mapping $t \mapsto f(U_t, \Theta_t)$ is absolutely continuous almost surely and*

the mapping $(\xi, s, \omega) \mapsto f(U_{s-}, \xi) - f(U_{s-}(\omega), \Theta_{s-}(\omega))$ is integrable with respect to the random measure $\Lambda(U_{s-}, \Theta_{s-})\mu((U_{s-}, \Theta_{s-}), d\xi)ds$.

Moreover, if in addition f is continuously Fréchet-differentiable with respect to its first argument such that the Riesz Representation¹ $f_u \in H$ of the Fréchet derivative satisfies $f_u(u, \theta) \in X$ for $u \in X$ and is a locally bounded composition operator in $L^2((0, T), X)$,² then the extended generator $\mathcal{A}f$ is given by

$$\mathcal{A}f(u, \theta) = \langle A(\theta)u + B(\theta, u), f_u(u, \theta) \rangle_X + \Lambda(u, \theta) \int_K (f(u, \xi) - f(u, \theta)) \mu((u, \theta), d\xi). \quad (2.4)$$

2.2 An appropriate sequence of PDMPs

Let E denote another separable real Hilbert space. Further, for a certain $m \in \mathbb{N}$ (its significance is explained in the next paragraph) we denote by $\mathcal{H} = \times_{i=1}^m H$, $\mathcal{E} = \times_{i=1}^m E$ the direct products of the Hilbert spaces H and E which are Hilbert spaces themselves. Finally we set $\mathcal{E}^* = \times_{j=1}^m E^*$ which is the dual space to \mathcal{E} .

We now define the structure of the sequence of processes for which we derive the limit theorems. For all $n \in \mathbb{N}$ let $(\Omega^n, \mathcal{F}^n, (\mathcal{F}_t^n)_{t \geq 0}, \mathbb{P}^n)$ be a filtered probability space satisfying the usual conditions and the processes $(X_t^n)_{t \geq 0} = (U_t^n, \Theta_t^n)_{t \geq 0}$ defined thereon are regular PDMPs taking values in $H \times K_n$ with path properties as defined in Section 2.1. Correspondingly, the characteristics of the PDMPs are given by $(A^n, B^n, \Lambda^n, \mu^n)$. Note that the state space K_n for the piecewise constant component changes with varying index n whereas the state space H for the continuous component remains fixed. Therefore, in order for such a sequence of processes to allow for a limit we need to impose a special structure on the characteristics referring to the continuous component. To this end we assume there exists an $m \in \mathbb{N}$, introduced above, such that for each PDMP $(U_t^n, \Theta_t^n)_{t \geq 0}$ there exists a family of measurable coordinate functions $z_i^n : K_n \rightarrow E$, $i = 1, \dots, m$, such that the characteristics $A^n(\theta), B^n(\theta)$ depend on the piecewise constant component and on the index n only via the \mathcal{E} -valued coordinate process $z^n(\theta) = (z_1^n(\theta), \dots, z_m^n(\theta))$. That is, there exist measurable operators $A, B : \mathcal{E} \times X \rightarrow X^*$ such that for all $n \in \mathbb{N}$, all $u \in H$ and all $\theta \in K_n$

$$A^n(\theta)u = A(z^n(\theta))u, \quad B^n(\theta, u) = B(z^n(\theta), u). \quad (2.5)$$

The coordinates z^n can be interpreted as a ‘sufficient statistic’ of the piecewise constant component for the evolution of the continuous component. In statistics a sufficient statistic for a quantity of interest is a function of the observations that is sufficient to estimate this particular quantity. For example, the sample average of independently and identically distributed real random variables is a sufficient statistic for the mean of their distribution. In the present setting, this means that the coordinate functions contain all information about the vector θ that is needed to determine the continuous dynamics in between jumps. Further, the essence of the subsequent limit theorems is that the sequence of coordinate processes on the space \mathcal{E} allows for a limit under certain conditions. Typically, in applications one is interested in the dynamics of the continuous components only, thus a restriction of the attention to the coordinate functions is well justified. As \mathcal{E} is a (vector-valued) Hilbert space itself, no generality would be lost if instead of the family of coordinate functions we assumed the existence of Hilbert space-valued functions z^n taking values in the same Hilbert space for each n . However, we decided to use this more detailed notation since in examples one usually encounters that it is a set of coordinate functions that encodes the information necessary for defining

¹Note that the Fréchet derivative at a point $u \in H$ is a linear, bounded functional on H and thus an element of the dual H^* .

²An example of such a function f is $(u, \theta) \mapsto \|u\|_H^2$ in which case $f_u(u, \theta) = 2u$.

the dynamics of the continuous component.

In order to illustrate this set-up let us briefly discuss the Hodgkin-Huxley model as an example of the general excitable membrane model considered in Section 7. Here the sequence of abstract evolution equations (2.1) arises from parabolic partial differential equations modelling the space-time evolution of the membrane potential of the form

$$\dot{u}(t, x) = \Delta u(t, x) + \sum_{i=\text{Na,K,L}} \bar{g}_i p_i(t, x) (E_i - u(t, x)), \quad t \geq 0, x \in D \subset \mathbb{R}^d \quad (2.6)$$

with constants $\bar{g}_i > 0$ and $E_i \in \mathbb{R}$, cf. (7.6). The indices refer to electrical currents due to the movement of charged Sodium (Na) and Potassium (K) ions across the membrane and ohmic leakage (L) current mainly due to Chloride ions [26]. In hybrid versions of the Hodgkin-Huxley system the conductances $p_i(t, x)$ depend on the finite number of open ion channels distributed in the membrane which increases with n . Each individual channel is modelled stochastically opening or closing at random times with dynamics depending on u , cf. Section 7.2 for more details. In the case of constant potential $u(t, x) \equiv \bar{u}$ each channel were a continuous time Markov chain. The collection of channel states at any time instant t defines the discrete component Θ_t^n . Finally, the coordinate functions z^n relate channels in a specific state to their location in the physical space D , cf. their definition in (7.5). They map the channel configurations into piecewise constant space-time functions stating the local density of channels in the particular states, thus $p_i^n(t, x) := z_i^n(\Theta_t^n) \in L^2(D) = E$. Hence, equipped with suitable boundary conditions equation (2.6) is an abstract evolution equation of the type (2.1) where the Hilbert spaces H, X and E are spaces of real functions on $D \subset \mathbb{R}^d$.

3 The associated martingale process

For the limit theorems we derive in this paper, the main estimation procedures concern certain martingales associated with the PDMP. As these are of such central importance we discuss them in this separate section. The principle aim is, on the one hand, to derive conditions that imply the convergence in probability of the associated martingales as needed for the law of large numbers (cf. condition (4.5) in Theorem 4.1) and, on the other hand, we present some necessary structure for the central limit theorems. Therefore we define for all $j = 1, \dots, m$ the E -valued stochastic process M_j^n by

$$M_j^n(t) := z_j^n(\Theta_t^n) - z_j^n(\Theta_0^n) - \int_0^t [\mathcal{A}^n \langle \cdot, z_j^n(\cdot) \rangle_E](U_s^n, \Theta_s^n) ds, \quad (3.1)$$

where the integrand in the right hand side is given by

$$\begin{aligned} [\mathcal{A}^n \langle \cdot, z_j^n(\cdot) \rangle_E](U_s^n, \Theta_s^n) &= \Lambda^n(U_s^n, \Theta_s^n) \int_{K_n} \left(z_j^n(\xi) - z_j^n(\Theta_s^n) \right) \mu^n((U_s^n, \Theta_s^n), d\xi) \\ &= \Lambda^n(U_s^n, \Theta_s^n) \sum_{\xi \in K_n} \left(z_j^n(\xi) - z_j^n(\Theta_s^n) \right) \mu^n((U_s^n, \Theta_s^n), \{\xi\}). \end{aligned} \quad (3.2)$$

Hence the integrand is a countable convex combination of elements in E with time-dependent coefficients and in between jumps it depends continuously on s . Anticipating condition (3.4) below, which we generally assume to hold, we find that the integral in the right hand side of (3.1) almost surely exists in the sense of Bochner. For a brief discussion of the Bochner integral we refer to [39, App. A]. For an application of a

functional $\phi \in E^*$ to (3.1) we obtain

$$\langle \phi, M_j^n(t) \rangle_E = \langle \phi, z_j^n(\Theta_t^n) \rangle_E - \langle \phi, z_j^n(\Theta_0^n) \rangle_E - \int_0^t [\mathcal{A}^n \langle \phi, z_j^n(\cdot) \rangle_E](U_s^n, \Theta_s^n) ds, \quad (3.3)$$

where the integrand is

$$[\mathcal{A}^n \langle \phi, z_j^n(\cdot) \rangle_E](U_s^n, \Theta_s^n) = \Lambda^n(U_s^n, \Theta_s^n) \int_{K_n} \langle \phi, z_j^n(\xi) \rangle_E - \langle \phi, z_j^n(\Theta_s^n) \rangle_E \mu^n((U_s^n, \Theta_s^n), d\xi).$$

Thus the integral has the form of the extended generator, cf. Theorem 2.1, applied to the mapping $(u, \theta) \mapsto \langle \phi, z_j^n(\theta) \rangle_E$. This already suggests that the processes (3.3) are martingales under suitable boundedness conditions. In fact we are able to establish that the processes M_j^n are E -valued càdlàg martingales. We refer to [14, 39] for a brief discussion of martingales in infinite-dimensional spaces. The easiest way to validate the martingale property is due to the following result [39, Sec. 2.3]: If $\mathbb{E}^n \|M_j^n(t)\|_E < \infty$ for all $t \in [0, T]$, the Hilbert space-valued martingale property holds if and only if $\langle \phi, M_j^n(t) \rangle_E$ is a real-valued martingale for all $\phi \in E^*$. The following theorem gives a condition that guarantees that the processes (3.1) are square-integrable martingales and satisfy an Itô-isometry. The proof is rather technical and thus we have deferred it to the Appendix A.

Theorem 3.1. *Let $n \in \mathbb{N}$ be fixed and assume that for all $t > 0$ it holds that*

$$\mathbb{E} \int_0^t \left[\Lambda^n(U_s^n, \Theta_s^n) \int_{K_n} \|z_j^n(\xi) - z_j^n(\Theta_s^n)\|_E^2 \mu^n((U_s^n, \Theta_s^n), d\xi) \right] ds < \infty. \quad (3.4)$$

Then the process M_j^n is a square-integrable martingale and satisfies the Itô-isometry

$$\mathbb{E}^n \|M_j^n(t)\|_E^2 = \int_0^t \mathbb{E}^n \left[\Lambda^n(U_s^n, \Theta_s^n) \int_{K_n} \|z_j^n(\xi) - z_j^n(\Theta_s^n)\|_E^2 \mu^n((U_s^n, \Theta_s^n), d\xi) \right] ds. \quad (3.5)$$

We continue the investigation of the processes M_j^n as Hilbert space valued martingales. From now on we always assume that assumption (3.4) holds. Note that the finiteness of the second moments of the jump sizes is a standard condition in related fluid limit theorems [31, 37, 36]. We introduce a concept akin to the quadratic covariance operator in Euclidean finite dimensional spaces. This concept is important for the central limit theorems in, on the one hand, establishing weak convergence, and, on the other hand, characterising the limit. For further reference we refer to [35].

Definition 3.1. For the square-integrable, E -valued, càdlàg martingale M_j^n we denote by $(\ll M_j^n \gg_t)_{t \geq 0}$ its *predictable quadratic variation process*, i.e., the unique (up to indistinguishability), predictable $L_1(E^*, E)$ -valued³ process which satisfies that for all $\phi, \psi \in E^*$ the real-valued process

$$t \mapsto \langle \phi, M_j^n(t) \rangle_E \langle \psi, M_j^n(t) \rangle_E - \langle \phi, \ll M_j^n \gg_t \psi \rangle_E \quad (3.6)$$

is a local martingale.

The aim now is to obtain an explicit formula for the quadratic variation process of the individual martingales M_j^n as well as of the vector-valued process M^n of all martingales M_j^n , i.e., the \mathcal{E} -valued process

$$t \mapsto M^n(t) = (M_1^n(t), \dots, M_m^n(t)).$$

³ $L_1(E^*, E)$ denotes the space of trace class operators from the Hilbert space E^* into E .

To this end we define for all $i, j = 1, \dots, m$ operators $G_{ij}^n \in L(E^*, E)$ by

$$\begin{aligned} \psi &\mapsto G_{ij}^n(u, \theta^n)\psi := & (3.7) \\ &:= \Lambda^n(u, \theta^n) \int_{K_n} \langle \psi, z_i^n(\xi) - z_i^n(\theta^n) \rangle_E \left(z_j^n(\xi) - z_j^n(\theta^n) \right) \mu^n((u, \theta^n), d\xi). \end{aligned}$$

Clearly, these are linear, bounded operators mapping $E^* \rightarrow E$ and depend measurably on $(u, \theta^n) \in H \times K_n$. For $i = j$ each operator is *non-negative*, i.e., $\langle \phi, G_{jj}^n(u, \theta^n)\phi \rangle_E \geq 0$ for all $\phi \in E^*$, and *symmetric*, i.e., $\langle \psi, G_{jj}^n(u, \theta^n)\phi \rangle_E = \langle \phi, G_{jj}^n(u, \theta^n)\psi \rangle_E$ for all $\phi, \psi \in E^*$. Let $(\varphi_k)_{k \in \mathbb{N}}$ denote an orthonormal basis in E^* . We find due to the Riesz Representation Theorem and Parseval's identity that the trace of the operators G_{jj} satisfies

$$\begin{aligned} \text{Tr } G_{jj}^n(u, \theta^n) &= \Lambda^n(u, \theta^n) \int_{K_n} \sum_{k \in \mathbb{N}} \left(\langle \varphi_k, z_j^n(\xi) - z_j^n(\theta^n) \rangle_E \right)^2 \mu^n((u, \theta^n), d\xi) \\ &= \Lambda^n(u, \theta^n) \int_{K_n} \|z_j^n(\xi) - z_j^n(\theta^n)\|_E^2 \mu^n((u, \theta^n), d\xi). \end{aligned} \tag{3.8}$$

For arbitrary i, j the trace is bounded in terms of (3.8) as it follows from Young's inequality that $\text{Tr } G_{ij}^n(u, \theta^n) \leq \frac{1}{2} \text{Tr } G_{ii}^n(u, \theta^n) + \frac{1}{2} \text{Tr } G_{jj}^n(u, \theta^n)$.

Let $\Phi = (\phi_1, \dots, \phi_m)$ and $\Psi = (\psi_1, \dots, \psi_m)$ be elements of \mathcal{E}^* . Summing over all operators (3.7) applied to the components of Φ, Ψ as indicated by the indices, i.e.,

$$\langle \Phi, G^n(u, \theta^n) \Psi \rangle_{\mathcal{E}} := \sum_{i,j=1}^m \langle \phi_i, G_{ij}^n(u, \theta^n) \psi_j \rangle_E, \tag{3.9}$$

we obtain a linear, bounded operator $G^n(u, \theta^n)$ mapping \mathcal{E}^* to \mathcal{E} . This operator is symmetric as the operators G_{ij}^n satisfy $\langle \phi, G_{ij}^n(u, \theta^n)\psi \rangle_E = \langle \psi, G_{ji}^n(u, \theta^n)\phi \rangle_E$ for all i, j . Moreover, the operator $G^n(u, \theta^n)$ is non-negative as it holds that

$$\langle \Psi, G^n(u, \theta^n) \Psi \rangle_{\mathcal{E}} = \Lambda^n(u, \theta^n) \int_{K_n} \left(\sum_{i=1}^m \langle \psi_i, z_i^n(\xi) - z_i^n(\theta^n) \rangle_E \right)^2 \mu^n((u, \theta^n), d\xi).$$

Finally, the operator $G^n(u, \theta^n)$ is of trace class if the operators G_{jj} , $j = 1, \dots, m$, are of trace class and the trace satisfies

$$\text{Tr } G^n(u, \theta^n) = \Lambda^n(u, \theta^n) \int_{K_n} \|z^n(\xi) - z^n(\theta^n)\|_{\mathcal{E}}^2 \mu^n((u, \theta^n), d\xi). \tag{3.10}$$

We next prove that the operators (3.8) give the quadratic variations of the martingales (3.1).

Proposition 3.1. *The quadratic variation of the martingale M_j^n satisfies for all $t \geq 0$*

$$\langle\langle M_j^n \rangle\rangle_t = \int_0^t G_{jj}^n(U_s^n, \Theta_s^n) ds. \tag{3.11}$$

Remark 3.1. It is an immediate consequence of Proposition 3.1 that the quadratic variation of the \mathcal{E} -valued martingale M^n is given analogously to (3.11) by integrating the operator G^n .

Proof. First of all note that due to the characterisation of the trace (3.8) and condition (3.4) it holds that the process in the right hand side of (3.11) takes values in $L_1(E^*, E)$

almost surely. Further, it holds that $\ll M_j^n \gg_t$ satisfies for all $\phi, \psi \in E$ that

$$\begin{aligned} &\langle \phi, \ll M_j^n \gg_t \psi \rangle_E = \\ &= \int_0^t \Lambda^n(U_s^n, \Theta_s^n) \int_{K_n} \langle \psi, z_j^n(\xi) - z_j^n(\Theta_s^n) \rangle_E \langle \phi, z_j^n(\xi) - z_j^n(\Theta_s^n) \rangle_E \mu^n((U_s^n, \Theta_s^n), d\xi) ds \end{aligned}$$

as this right hand side is, due to [24, Prop. 4.6.2], the unique real-valued process such that $\langle \phi, M_j^n(t) \rangle_E \langle \psi, M_j^n(t) \rangle_E - \langle \phi, \ll M_j^n \gg_t \psi \rangle_E$ is a local martingale. Here $\langle \phi, M_j^n(t) \rangle_E$ and $\langle \psi, M_j^n(t) \rangle_E$ are understood as real-valued stochastic integrals with respect to the associated martingale measure of the PDMP. Thus we infer that for all $\phi, \psi \in E$ it holds

$$\langle \phi, \ll M_j^n \gg_t \psi \rangle_E = \int_0^t \langle \phi, G_{jj}^n(U_s^n, \Theta_s^n) \psi \rangle_E ds.$$

Finally, the linearity of the Bochner integral (note that $L_1(E^*, E)$ is a Banach space) implies (3.11). \square

A further second property of the quadratic variation is that the process

$$t \mapsto \|M_j^n(t)\|_E^2 - \text{Tr} \ll M_j^n \gg_t$$

is a local martingale. We note that the trace process $t \mapsto \text{Tr} \ll M_j^n \gg_t$ is the unique, predictable increasing process exhibiting this property. Using the characterisation (3.11) of the quadratic variation we thus obtain that the process

$$t \mapsto \|M_j^n(t)\|_E^2 - \text{Tr} \left(\int_0^t G_{jj}^n(U_s^n, \Theta_s^n) ds \right) = \|M_j^n(t)\|_E^2 - \int_0^t \text{Tr} G_{jj}^n(U_s^n, \Theta_s^n) ds \quad (3.12)$$

is a local martingale vanishing almost surely at $t = 0$ and analogously in the case of the \mathcal{E} -valued martingale M^n .

We are now in a position to state a lemma which establishes the convergence in probability (3.14) of the processes $(M_j^n)_{t \geq 0}$ necessary for the law of large numbers, cf. condition (4.5) in Theorem 4.1.

Lemma 3.1. *Assume that for all $T > 0$*

$$\lim_{n \rightarrow \infty} \mathbb{E} \int_0^T \left[\Lambda^n(U_s^n, \Theta_s^n) \int_{K_n} \|z_i^n(\xi) - z_i^n(\Theta_s^n)\|_E^2 \mu^n((U_s^n, \Theta_s^n), d\xi) \right] ds = 0. \quad (3.13)$$

Then the process (3.12) is a martingale and for all $T, \epsilon > 0$, it holds that

$$\lim_{n \rightarrow \infty} \mathbb{P}^n \left[\sup_{t \in [0, T]} \|M_j^n(t)\|_E > \epsilon \right] = 0. \quad (3.14)$$

Proof. As the process M_j^n is an E -valued càdlàg martingale, it holds that $\|M_j^n\|_E^2$ is a càdlàg submartingale. Thus an application of Markov's and Doob's inequalities yield the estimates

$$\mathbb{P}^n \left[\sup_{t \in [0, T]} \|M_j^n(t)\|_E^2 > \epsilon \right] \leq \frac{1}{\epsilon} \mathbb{E}^n \left[\sup_{t \in [0, T]} \|M_j^n(t)\|_E^2 \right] \leq \frac{4}{\epsilon} \mathbb{E}^n \|M_j^n(T)\|_E^2.$$

Now, the Itô-isometry (3.5) and condition (3.13) imply the convergence in probability (3.14). It remains to show that the process (3.12) is a martingale. A sufficient condition,

see, e.g., [24, Prop. B.0.13], is that for all $T > 0$ it holds

$$\mathbb{E}^n \left[\sup_{t \in [0, T]} \left| \|M_j^n(t)\|_E^2 - \int_0^t \text{Tr} G_{jj}^n(U_s^n, \Theta_s^n) ds \right| \right] < \infty. \tag{3.15}$$

Estimating the term inside the expectation we obtain

$$\begin{aligned} & \sup_{t \in [0, T]} \left| \|M_j^n(t)\|_E^2 - \int_0^t \text{Tr} G_{jj}^n(U_s^n, \Theta_s^n) ds \right| \\ & \leq \sup_{t \in [0, T]} \|M_j^n(t)\|_E^2 + \sup_{t \in [0, T]} \int_0^t \Lambda^n(u, \Theta^n) \int_{K_n} \|z_j^n(\xi) - z_j^n(\Theta^n)\|_E^2 \mu^n((u, \Theta^n), d\xi) ds. \end{aligned}$$

The expectation of the first supremum term in the right hand side is bounded due to Doob's inequality and the square-integrability of the martingale. The term inside the second supremum is increasing, thus its expectation is finite due to condition (3.13). \square

4 A weak law of large numbers

In order to propose a deterministic limit of the sequence of PDMPs we consider functions $F_j : \mathcal{E} \times H \rightarrow E, j = 1, \dots, m$. In combination with the operators A, B these functions are used to define a coupled system of deterministic abstract evolution equations

$$\begin{aligned} \dot{u} &= A(p)u + B(p, u), \\ \dot{p}_j &= F_j(p, u), \quad j = 1, \dots, m. \end{aligned} \tag{4.1}$$

We assume that to suitable initial condition $(u_0, p_0) \in H \times \mathcal{E}$ there exists a unique weak solution $(u(t), p(t))_{t \geq 0}$ in $C(\mathbb{R}_+, H \times \mathcal{E})$ of (4.1). Additionally, we assume that for all $i = 1, \dots, m$ the components p_i satisfy

$$\langle \phi, p_i(t) \rangle_E = \langle \phi, p_i(0) \rangle_E + \int_0^t \langle \phi, F_i(p(s), u(s)) \rangle_E ds \quad \forall t \in [0, T], \phi \in E^*. \tag{4.2}$$

That is, the components p_j satisfy the equation (4.1) in the sense of an Hilbert space valued integral equation. We note that in application one usually encounters deterministic limit systems that possess strong or classical solutions and hence the current weak framework is satisfied. Finally, we assume that the operators A, B and $F_j, j = 1, \dots, m$, satisfy Lipschitz-type conditions on $L^2((0, T), \mathcal{E} \times X)$ in the sense that for every $T > 0$ there exist constants L_1 and L_2 such that for all $u, v \in L^2((0, T), X)$ and all $p, q \in L^2((0, T), \mathcal{E})$ it holds that

$$\begin{aligned} & \int_0^T \langle A(q)v - A(p)u, v - u \rangle_X + \langle B(q, v) - B(p, u), v - u \rangle_X dt \\ & \leq L_1 \int_0^T \|v - u\|_H^2 + \sum_{i=1}^m \|q_i - p_i\|_E^2 dt. \end{aligned} \tag{4.3}$$

and

$$\left(\int_0^T \|F_j(q, v) - F_j(p, u)\|_E dt \right)^2 \leq L_2 \int_0^T \|v - u\|_H^2 + \sum_{i=1}^m \|q_i - p_i\|_E^2 dt, \tag{4.4}$$

where we have omitted the arguments t of the functions u, v, p and q .

Remark 4.1. In the proof of the law of large numbers, see Section 6.1, these Lipschitz conditions are applied such that one pairing (v, q) refers to a path segment of the con-

tinuous component of a PDMP and the coordinate process and the second (u, p) to the deterministic limit functions. Thus for the applications of (4.3) and (4.4) in the proof it is sufficient that these hold only for pairings (v, q) out of a set containing almost all paths of the sequence of PDMPs and (u, p) being the deterministic limit, i.e., one (!) distinguished pairing. This restriction of (4.3) and (4.4) to be satisfied only for particular pairings (v, q) and (u, p) out of the whole path space has a decisive advantage: In order to establish these conditions we are able to incorporate additional qualitative results on the trajectories of the PDMPs and the deterministic limit and the constants L_1, L_2 may depend on (u, p) . For example, in the application to excitable membrane models such an additional qualitative is that the components corresponding to u, v, p, q are pointwise bounded.

We now present a weak law of large numbers in Theorem 4.1 below. The proof of the theorem follows the lines of previously published limit theorems considering processes in finite dimensions [31, 37]. The main difficulties arising in infinite-dimensional phase space concerns the bounds on the martingale part, cf. condition (C1), which is rarely a problem in finite dimensions. However, using the appropriate martingale theory in Hilbert spaces these can be kept to a minimum. Then the difficulties are mainly of a technical nature as martingale theory in connection with PDMPs in infinite-dimensional spaces gets more involved and is not covered by previous results in [24]. We have established the necessary theory in the preceding Section 3 and addressed the question of the convergence of the martingale part (C1) within this framework. Most importantly, in Lemma 3.1 we have proven a sufficient condition for (C1) to be satisfied. In particular, this sufficient condition (3.13) is a natural extension of the condition employed in finite dimensions, cf. [31, 37].

A different approach to establishing condition (C1) which avoids using martingale theory in Hilbert spaces is exemplified in the law of large numbers proved in [5]. In infinite-dimensional space this approach encounters the problem of simultaneously controlling countably many real martingales compared to only finitely many in the case of its finite-dimensional counterpart. This problem can be overcome with an intricate compactness argument which relies on the assumption that the dual space E^* is compactly embedded in some additional normed space and all estimates – especially an estimate which also implies condition (3.13) – have to be derived in the norm of this additional space. Furthermore, the condition, that all martingales $(\langle \phi, M_j^n(t) \rangle_E)_{t \geq 0}$, $j = 1, \dots, m$ and $\phi \in E^*$, possess almost surely uniformly bounded paths, has to be introduced. We are of the opinion that our approach is more elegant, but, more importantly, it avoids the introduction of additional conditions.

Finally, consistently with the notation in Section 3 we use in the subsequent theorem and its proof the notation $[\mathcal{A}^n \langle \cdot, z_j^n(\cdot) \rangle_E]$ as defined in (3.2). Then, for given $(u, \theta^n) \in H \times K_n$ functionals $[\mathcal{A}^n \langle \cdot, z_j^n(\cdot) \rangle_E](u, \theta^n)$ on E^* are defined by the mappings $\phi \mapsto [\mathcal{A}^n \langle \phi, z_j^n(\cdot) \rangle_E](u, \theta^n)$. As usual we identify the bidual E^{**} with E and thus $[\mathcal{A}^n \langle \cdot, z_j^n(\cdot) \rangle_E](u, \theta^n) \in E$.

Theorem 4.1. *We assume that the following conditions hold:*

(C1) *For all $j = 1, \dots, m$ it holds that for fixed $T, \epsilon > 0$*

$$\lim_{n \rightarrow \infty} \mathbb{P}^n [\sup_{t \in [0, T]} \|M_j^n(t)\|_E > \epsilon] = 0. \tag{4.5}$$

(C2) *The functions $F_j, j = 1, \dots, m$, satisfy for all $\epsilon > 0$ that*

$$\lim_{n \rightarrow \infty} \mathbb{P}^n \left[\int_0^T \left\| [\mathcal{A}^n \langle \cdot, z_j^n(\cdot) \rangle_E](U_t^n, \Theta_t^n) - F_j(z^n(\Theta_t^n), U_t^n) \right\|_E dt > \epsilon \right] = 0, \tag{4.6}$$

where we have omitted the argument t of the functions u and θ .

(C3) The initial conditions (U_0^n, Θ_0^n) of the sequence of PDMPs converge in probability to the initial conditions of the deterministic limit in the sense that for all $\epsilon > 0$

$$\lim_{n \rightarrow \infty} \mathbb{P}^n \left[\|U_0^n - u_0\|_H + \sum_{i=1}^m \|z_i^n(\Theta_0^n) - p_i(0)\|_E > \epsilon \right] = 0.$$

Then, for every $\epsilon > 0$ and every fixed $T > 0$ it holds that

$$\lim_{n \rightarrow \infty} \mathbb{P}^n \left[\sup_{t \in [0, T]} \left(\|U_t^n - u(t)\|_H^2 + \sum_{j=1}^m \|z_j^n(\Theta_t^n) - p_j(t)\|_E^2 \right) > \epsilon \right] = 0. \quad (4.7)$$

Remark 4.2. The result (4.7) implies convergence in probability of the processes $(U_t^n, z^n(\Theta_t^n))_{t \geq 0}$ to the deterministic function $(u(t), p(t))_{t \in [0, T]}$ in the Hilbert space $L^2((0, T), H \times \mathcal{E})$. If the differences of the components are almost surely bounded independent of n the convergence even holds in the mean, cf. the application of the law of large numbers in Theorem 7.1. Further, the conditions (C1)–(C3) are generalisations from Euclidean space to infinite-dimensional Hilbert spaces of those employed in the corresponding theorems for PDMPs in Euclidean space [37] and, in particular, of the original formulation in case of pure jump processes in Euclidean space [31]. In these cases the conditions above reduce to the corresponding assumptions.

5 The central limit theorem and the Langevin approximation

We proceed to the presentation of the central limit theorem for associated martingales $(M_t^n)_{t \geq 0}$ defined in (3.1). The central limit theorem provides the theoretical basis for an approximation of spatio-temporal PDMPs by Hilbert-space valued diffusion processes where the latter can be represented by solutions of stochastic partial differential equations, see Section 5.2.

Proving central limit theorems usually involves two tasks: On the one hand, one has to show the existence of a limit and, on the other hand, one has to provide a characterisation of the limit as a certain stochastic process. The former is equivalent to the problem of tightness of the stochastic processes. In the case of martingales sufficient conditions for tightness depending on the quadratic variation process are stated in [36]. In order to characterise the limit there exist different approaches, showing either that the limit solves a given (local) martingale problem which is known to have a unique solution (cf. [25, 36]) or proving weak convergence of the finite dimensional distributions (cf. [27, 37]). We present two central limit theorems, Theorems 5.1 and 5.2, employing the two methods, respectively, however, to avoid repetition we state only the proof of the first in the present study and refer to the PhD thesis of one of the authors [42] for the proof of the second. The two theorems differ in a technical assumption which in each case arises in addition to the central condition of the convergence of the quadratic variations. We believe that for applications of the limit theorems it is advantageous to know both versions of the martingale central limit theorem, as it is easily conceivable that only one of these technical assumptions is satisfied. Hence the theorems are applicable in different situations.

Finally, we emphasise that in the following the space \mathcal{E} need not necessarily be the same space for which the law of large numbers is satisfied. However, clearly, the space \mathcal{E} in the present section contains the space in the law of large numbers as subspace. In applications, usually, the law of large numbers holds in a space with a stronger norm, for example, for the excitable membrane model considered in Section 7 the law of large

numbers holds in $L^2(D)$ whereas the central limit theorem holds in the space $H^{-2s}(D)$.⁴ This is a major difference to the corresponding results in finite-dimensional space where both limit theorems hold in the same space.⁵

5.1 A martingale central limit theorem

In this section we present central limit theorems for the scaled \mathcal{E} -valued martingales $(\sqrt{\alpha_n} M_t^n)_{t \geq 0}$ associated with a sequence of PDMPs where $\alpha_n \in \mathbb{R}_+$, $n \in \mathbb{N}$, is a suitable rescaling sequence such that $\lim_{n \rightarrow \infty} \alpha_n = \infty$. Clearly, the rescaling is necessary in order to be able to obtain a limit different from the trivial limit as (4.5) implies that $(M_t^n)_{t \geq 0}$ converges to zero in distribution. We note that the sequence α_n can also be interpreted as characterising the speed of convergence of the martingales $(M_t^n)_{t \geq 0}$.

In the following let $t \mapsto G(u(t), p(t)) \in L(\mathcal{E}^*, \mathcal{E})$ be a Bochner-integrable operator-valued map such that each $G(u(t), p(t))$ is a symmetric, positive trace class operator. Particularly this implies for all $\Phi \in \mathcal{E}^*$ and all $t > 0$, that it holds that

$$\int_0^t \langle \Phi, G(u(s), p(s)) \Phi \rangle_{\mathcal{E}} ds < \infty. \tag{5.1}$$

Here $(u(t), p(t))_{t \geq 0}$ is the deterministic limit obtained in Theorem 4.1 and the use of this notation for the – at this point – arbitrary time-dependent operator G only illustrates that in applications the time-dependence is due to a dependence on the deterministic limit system. These operator-valued functions are used to define a unique centred diffusion process on \mathcal{E} , i.e., an \mathcal{E} -valued centred Gaussian process with independent increments, continuous sample paths. Such a process is uniquely defined by its covariance operator and due to a theorem of Itô stated in [27] every family of trace class operators $C^*(t) \in L_1(\mathcal{E}, \mathcal{E})$ which are increasing and continuous in t define a centred diffusion process. In the present situation we define C^* in the following way. We denote by $\iota : \mathcal{E} \rightarrow \mathcal{E}^*$ the canonical identification of a Hilbert space with its dual, hence we can define for $x, y \in \mathcal{E}$,

$$(x, C^*(t) y)_{\mathcal{E}} = \int_0^t \langle \iota(x), G(u(s), p(s)) \iota(y) \rangle_{\mathcal{E}} ds$$

which is continuous and increasing for all $x \in \mathcal{E}$ and $C^*(t)$ is a trace class operator on \mathcal{E} . Moreover, for operators $C(t) \in L_1(\mathcal{E}^*, \mathcal{E})$, defined by

$$\langle \Phi, C(t) \Psi \rangle_{\mathcal{E}} = \int_0^t \langle \Phi, (G(u(s), p(s)) \Psi) \rangle_{\mathcal{E}} ds, \tag{5.2}$$

there is obviously a one-to-one relationship between C^* and C . Hence, we may say that also the latter defines a diffusion process on the space \mathcal{E} .

We proceed to the statement of the central limit theorem. The proof of the theorem employs a characterisation of the limit via the local martingale problem. The essential condition characterising the limit is the convergence of the quadratic variation processes (5.6). The second condition (5.7) is a technical condition on the jump heights which arises due to the method of proof and is usually satisfied in applications. The remaining conditions are such that (D1) guarantees tightness of the sequence of processes and in combination with (D2) that any limit is a continuous stochastic process. The proof of the following theorem is deferred to Section 6.2.

⁴Here and everywhere else $H^{-2s}(D)$ is the dual space to the Sobolev space $H^{2s}(D)$ where $D \subset \mathbb{R}^d$ and $s > d/2$.

⁵Note also that in finite-dimensional spaces all norms, and hence also all norms on subspaces, are equivalent which does not hold in the case of an infinite-dimensional Hilbert space.

Theorem 5.1. *We assume that the following conditions hold:*

(D1) *For all $t > 0$ it holds that*

$$\sup_{n \in \mathbb{N}} \alpha_n \mathbb{E}^n \int_0^t \left[\Lambda^n(U_s^n, \Theta_s^n) \int_{K_n} \|z^n(\xi) - z^n(\Theta_s^n)\|_{\mathcal{E}}^2 \mu^n((U_s^n, \Theta_s^n), d\xi) ds \right] < \infty, \quad (5.3)$$

and there exists an orthonormal basis $(\varphi_k)_{k \in \mathbb{N}}$ of \mathcal{E}^ such that for all $k \in \mathbb{N}$ and all $(u, \theta^n) \in H \times K_n$ except on a set of potential zero⁶*

$$\alpha_n \mathbb{E}^n \left[\int_0^t \langle \varphi_k, G^n(U_s^n, \Theta_s^n) \varphi_k \rangle_{\mathcal{E}} ds \mid (U_0^n, \Theta_0^n) = (u, \theta^n) \right] \leq \gamma_k C(t), \quad (5.4)$$

where the constants $\gamma_k > 0$, independent of n, t and (u, θ^n) , satisfy $\sum_{k \in \mathbb{N}} \gamma_k < \infty$, and the constant $C(t) > 0$, independent of n, k and (u, θ^n) , satisfies $\lim_{t \rightarrow 0} C(t) = 0$.

(D2) *For all $\beta > 0$ and every $\Phi \in \mathcal{E}^*$ it holds that*

$$\lim_{n \rightarrow \infty} \mathbb{E}^n \left[\int_0^t \Lambda^n(U_s^n, \Theta_s^n) \int_{\sqrt{\alpha_n} |\langle \Phi, z^n(\xi) - z^n(\Theta_s^n) \rangle_{\mathcal{E}}| > \beta} \mu^n((U_s^n, \Theta_s^n), d\xi) ds \right] = 0. \quad (5.5)$$

(D3) *Further, for all $\Phi \in \mathcal{E}^*$ and all $t > 0$ it holds that*

$$\lim_{n \rightarrow \infty} \int_0^t \mathbb{E}^n |\langle \Phi, G(u(s), p(s)) \Phi \rangle_{\mathcal{E}} - \alpha_n \langle \Phi, G^n(U_s^n, \Theta_s^n) \Phi \rangle_{\mathcal{E}}| ds = 0. \quad (5.6)$$

Finally, we assume that the jump heights of the rescaled martingales are almost surely uniformly bounded, i.e., there exists a constant $C < \infty$ such that it holds almost surely for all $n \in \mathbb{N}$ that

$$\sup_{t \geq 0} \sqrt{\alpha_n} \|z^n(\Theta_t^n) - z^n(\Theta_{t-}^n)\|_{\mathcal{E}} < C. \quad (5.7)$$

Then it follows that the process $(\sqrt{\alpha_n} M_t^n)_{t \geq 0}$ converges weakly to an \mathcal{E} -valued centred diffusion process characterised by the covariance operator (5.2).

We now state a second version of the martingale central limit theorem wherein the limiting process is characterised by the convergence of the characteristic functions.

Theorem 5.2. *Assume that the laws of the martingales $(\sqrt{\alpha_n} M_t^n)_{t \geq 0}$ form a tight sequence, e.g., condition (D1) is satisfied.*

(D3') *The convergence (5.6) holds and there exists a sequence $\beta_n > 0$ decreasing to zero such that for all $\Phi \in \mathcal{E}^*$*

$$\lim_{n \rightarrow \infty} \alpha_n \mathbb{E}^n \left[\int_0^t \Lambda^n(U_s^n, \Theta_s^n) \int_{\sqrt{\alpha_n} |\langle \Phi, z^n(\xi) - z^n(\Theta_s^n) \rangle_{\mathcal{E}}| > \beta_n} |\langle \Phi, z^n(\xi) - z^n(\Theta_s^n) \rangle_{\mathcal{E}}|^2 \mu^n((U_s^n, \Theta_s^n), d\xi) ds \right] = 0. \quad (5.8)$$

Then it follows that the process $(\sqrt{\alpha_n} M_t^n)_{t \geq 0}$ converges weakly to an \mathcal{E} -valued centred diffusion process characterised by the covariance operator (5.2).

⁶A set of potential zero is a subset of the state space of the process which the process almost surely never reaches.

The central condition of the convergence of the quadratic variation processes (5.6) is unchanged, however, the second, technical condition (5.7) in (D3) is changed due to the different method of proof. That is, condition (5.8) arises instead of (5.7) as an assumption on the distribution of the jump heights employing a characterisation of the limit process using convergence of characteristic functions instead of the local martingale problem. The significance for applications of condition (5.8) in contrast to (5.7) is that the former avoids the almost sure uniform bound on the jump heights in the latter. That is, arbitrarily large jumps are possible for each martingale in the sequence as long as their probability decreases sufficiently fast. Note that (5.8) is stronger than the similar condition (D2) in the preceding theorem. We omit the proof of the theorem which is an adaptation of the estimating procedures in [32, 37] to the infinite-dimensional setting. For details we refer to the PhD thesis of one of the present authors [42].

Remark 5.1. We remark without proof that the assumptions (D1) and (5.6) imply the convergence of the trace processes, i.e., for all $T > 0$

$$\lim_{n \rightarrow \infty} \alpha_n \int_0^T \mathbb{E}^n \text{Tr} G^n(U_s^n, \Theta_s^n) ds = \int_0^T \text{Tr} G(u(s), p(s)) ds.$$

5.2 Langevin approximation

Usually, e.g., in models of excitable membranes, one is ultimately interested in the dynamics of the continuous component. We have discussed in Section 2.2 that the coordinate functions $z_i^n, i = 1, \dots, m$, carry all the information needed for the dynamics of the continuous component $(U_t^n)_{t \geq 0}$. Therefore, the knowledge of the coordinate process $(z^n(\Theta_t^n))_{t \geq 0}$, or a close approximation thereof, is sufficient for many applications. From this point of view the significance of the law of large numbers and the martingale central limit theorem is that they provide a justification of an approximation of the processes $(U_t^n, z^n(\Theta_t^n))_{t \geq 0}$ for large enough n by a deterministic evolution equation, on the one hand, and, as we argue in this section, by a stochastic partial differential equation on the other hand.

To this end we first discuss representations of the limiting diffusion in Theorems 5.2 and 5.1 as a stochastic integral. By definition $G(u(s), p(s)) \circ \iota$ is a non-negative, self-adjoint trace class operator acting on \mathcal{E} , hence there exists a unique non-negative square root, i.e., a non-negative operator $\sqrt{G(u(s), p(s)) \circ \iota}$ such that $G(u(s), p(s)) \circ \iota = \sqrt{G(u(s), p(s)) \circ \iota} \circ \sqrt{G(u(s), p(s)) \circ \iota}$. Let $(W_t)_{t \geq 0}$ be a standard cylindrical Wiener process on \mathcal{E} with covariance operator given by the identity (cf. [14, 39]). Then, as

$$\mathbb{E} \int_0^t \text{Tr}(\sqrt{G(u(s), p(s)) \circ \iota} \sqrt{I})(\sqrt{G(u(s), p(s)) \circ \iota} \sqrt{I})^* ds = \int_0^t \text{Tr} G(u(s), p(s)) ds < \infty,$$

the mapping $t \mapsto \sqrt{G(u(s), p(s)) \circ \iota}$ is a valid integrand process for a stochastic integral with respect to $(W_t)_{t \geq 0}$. That is, the process $(Z_t)_{t \geq 0}$ defined for all $t \geq 0$ by

$$Z_t := \int_0^t \sqrt{G(u(s), p(s)) \circ \iota} dW_s \tag{5.9}$$

is an \mathcal{E} -valued Gaussian process with continuous sample paths and independent increments which, in addition, is also a square-integrable martingale. Moreover, the process has the covariance given by the operator $\int_0^t G(u(s), p(s)) ds$. Therefore, due to unique definition of Gaussian processes via their covariance operators, the process $(Z_t)_{t \geq 0}$ is a version of the limiting diffusion identified for the sequence of martingales $(\sqrt{\alpha_n} M_t^n)_{t \geq 0}$.

Hence, formally inserting the limits into the decomposition of the PDMP we obtain that the Langevin approximation $(\tilde{U}_t^n, \tilde{P}_t^n)_{t \geq 0}$ of $(U_t^n, z^n(\Theta_t^n))_{t \geq 0}$ is given by the solution

of the system of stochastic partial differential equations

$$\begin{aligned} d\tilde{U}_t^n &= (A(\tilde{P}_t^n)\tilde{U}_t^n + B(\tilde{P}_t^n, \tilde{U}_t^n)) dt \\ d\tilde{P}_t^n &= F(\tilde{P}_t^n, \tilde{U}_t^n) dt + \frac{1}{\sqrt{\alpha_n}} \sqrt{G(\tilde{U}_t^n, \tilde{P}_t^n)} dW_t. \end{aligned} \tag{5.10}$$

The sequence of Langevin approximations $(\tilde{U}_t^n, \tilde{P}_t^n)_{t \geq 0}$ possesses the same asymptotic behaviour as the sequence of processes $(U_t^n, z^n(\Theta_n))_{t \geq 0}$. It is obvious that for $n \rightarrow \infty$ and thus $\alpha_n \rightarrow \infty$ the noise term in (5.10) vanishes and the system approximates the deterministic solution $(u(t), p(t))_{t \geq 0}$ of the system (4.1), just as was proven in the law of large numbers Theorem 4.1 for the sequence of PDMPs. It poses no difficulties to make this statement precise in the form of a weak law of large numbers similar to Theorem 4.1. Thus for large enough α_n one might expect that equation (5.10) produces a similar behaviour than the PDMP with the major advantage of being analytically (and numerically) to a great extent less complex.

In order to analyse properties of the Langevin approximation, clearly, well-posedness of the system (5.10) has to be addressed first. This is suitably done within the *variational approach* to stochastic partial differential equations. That is, equation (5.10) is assumed to hold as an integral equation in $X^* \times \mathcal{E}^*$ in contrast to the *semigroup approach* which defines the solution via the semigroup generated by the linear part of (5.10) and the variation of constants formula. (Note that in its generic form (5.10) does not necessarily possess a fully linear part.) The variational approach reflects the approach of using weak solution to abstract evolution equations defining the deterministic inter-jump motion of PDMPs taken in this paper. We refer to [33, Sec. 1.3.1] for a concise introduction to the variational approach to SPDEs containing an existence and uniqueness theorem as well as further references. We do not pursue the issue of well-posedness of the Langevin approximation any further at this point, as we are of the opinion that this question is best addressed when analysing the Langevin approximation for particular models.

Remark 5.2. The process (5.9) is not necessarily the only stochastic integral process which is a version of the limiting diffusion. Let U be another separable, real Hilbert space, where $U = \mathcal{E}$ is possible, and assume there exists an operator $Q \in L_1(U, U)$ (or Q cylindrical) and a function⁷ $g \in L^2((0, T), L_2(U, \mathcal{E}))$ for all $T > 0$ such that $G(u(t), p(t)) \circ \iota = g(u(s), p(s)) \circ Q \circ g^*(u(s), p(s))$ for all $t \geq 0$. Then, the process $(Z_t^Q)_{t \geq 0}$ defined by the stochastic integral

$$Z_t^Q := \int_0^t g(u(s), p(s)) dW_s^Q, \tag{5.11}$$

where $(W_t^Q)_{t \geq 0}$ is an \mathcal{E} -valued Q -Wiener process, has the same quadratic variation as $(Z_t)_{t \geq 0}$ hence the processes coincide in distribution. Then starting from the representation (5.11) the Langevin approximation is given by (5.10) with the obvious changes in the diffusion term. We note that in finite dimensions the non-uniqueness, see, e.g., [3, Chap. 8], of a stochastic integral associated to a given covariance matrix can be exploited to improve the speed of numerical approximations in Monte-Carlo simulations of diffusion approximations by choosing an optimal diffusion coefficient structure, see [34]. In infinite-dimensions the question of a practical implication of choosing a diffusion approximation based on (5.11) over (5.10) needs, to the best of our knowledge, still to be addressed.

⁷Here, $L^2((0, T), L_2(U, \mathcal{E}))$ denotes the space of square-integrable functions on $(0, T)$ taking values in the Hilbert-space of Hilbert-Schmidt operators from U to \mathcal{E} .

6 Proofs of the main results

6.1 Proof of Theorem 4.1 (Law of large numbers)

The central argument of the subsequent proof is an appropriate application of Gronwall’s Lemma such that the upper bound satisfies the convergence in probability. Here the estimating procedure yielding the estimate to which Gronwall’s Lemma is applied necessitates careful attention due to more intricate regularity aspects of solutions to abstract evolution equations in contrast to solutions of ODEs in Euclidean space.

The continuous component U_t^n of each PDMP is in between successive jump times the weak solution of an abstract evolution equation. Similarly $u(t)$ is the weak solution of the abstract evolution equation (4.1). Therefore also the difference of the two paths is in between jump times the weak solution of an abstract evolution equation. It thus holds due to [19, Sec. 5.9, Thm. 3] for almost all t that

$$\begin{aligned} \frac{d}{dt} \|U_t^n - u(t)\|_H^2 = \\ 2\langle A(z^n(\Theta_t^n)) U_t^n + B(U_t^n, z^n(\Theta_t^n)) - A(p(t)) u(t) - B(u(t), p(t)), U_t^n - u(t) \rangle_X. \end{aligned}$$

Integrating this equation we obtain the integral equation

$$\begin{aligned} \|U_{t_1}^n - u(t_1)\|_H^2 &= \|U_{t_0}^n - u_{t_0}\|_H^2 \\ &+ 2 \int_{t_0}^{t_1} \langle A(z^n(\Theta_s^n)) U_s^n + B(U_s^n, z^n(\Theta_s^n)) - A(p(s)) u(s) - B(u(s), p(s)), U_s^n - u(s) \rangle_X ds, \end{aligned} \tag{6.1}$$

which is valid for almost all t_0, t_1 in between two successive jump times. Since both sides of equation (6.1) are continuous the equality (6.1) holds for all t_0, t_1 between successive jump times. Moreover, as U_t^n is continuous also at jump times it follows that equation (6.1) holds for all $t \in [0, T]$, i.e., we have

$$\begin{aligned} \|U_t^n - u(t)\|_H^2 &= \|U_0^n - u_0\|_H^2 \\ &+ 2 \int_0^t \langle A(z^n(\Theta_s^n)) U_s^n + B(U_s^n, z^n(\Theta_s^n)) - A(p(s)) u(s) - B(u(s), p(s)), U_s^n - u(s) \rangle_X ds. \end{aligned} \tag{6.2}$$

Next we employ the one-sided Lipschitz condition (4.3) to estimate the integral in the right hand side of equation (6.2). This yields the inequality

$$\|U_t^n - u_t\|_H^2 \leq \|U_0^n - u_0\|_H^2 + 2L_1 \int_0^t \|U_s^n - u(s)\|_H^2 ds + 2L_1 \sum_{j=1}^m \int_0^t \|z_j^n(\Theta_s^n) - p_j(s)\|_E^2 ds. \tag{6.3}$$

The overall aim is to apply Gronwall’s inequality to the growth inequality (6.3). Therefore, in the next step we derive a control on the terms $\|z_j^n(\Theta_s^n) - p_j(s)\|_E^2$ in the right hand side of inequality (6.3). As p is a solution of (4.1) satisfying (4.2) we obtain for every functional $\phi \in E^*$ a decomposition

$$\begin{aligned} \langle \phi, z_j^n(\Theta_t^n) - p_j(t) \rangle_E &= \langle \phi, z_j^n(\Theta_0^n) - p_j(0) \rangle_E \\ &+ \int_0^t [\mathcal{A}^n \langle \phi, z_j^n(\cdot) \rangle_E](U_s^n, \Theta_s^n) ds - \int_0^t \langle \phi, F_j(p(s), u(s)) \rangle_E ds + \langle \phi, M_j^n(t) \rangle_E, \end{aligned} \tag{6.4}$$

where the term $\langle \phi, M_j^n(t) \rangle_E$ has precisely the form (3.3) for all $t \in [0, T]$. Next we expand

the decomposition (6.4) to obtain

$$\begin{aligned} \langle \phi, z_j^n(\Theta_t^n) - p_j(t) \rangle_E &= \langle \phi, z_j^n(\Theta_0^n) - p_j(0) \rangle_E \\ &+ \langle \phi, M_j^n(t) \rangle_E + \int_0^t [\mathcal{A}^n \langle \phi, z_j^n(\cdot) \rangle_E](U_s^n, \Theta_s^n) - \langle \phi, F_j(z^n(\Theta_s^n), U_s^n) \rangle_E ds \\ &+ \int_0^t \langle \phi, F_j(z^n(\Theta_s^n), U_s^n) - F_j(p(s), u(s)) \rangle_E ds. \end{aligned}$$

We take the supremum over all $\phi \in E^*$ with $\|\phi\|_{E^*} \leq 1$ on both sides of this equation, square both sides and apply to the right hand side the inequality $|a_1 + \dots + a_k|^2 \leq k(|a_1|^2 + \dots + |a_k|^2)$ and the Cauchy-Schwarz inequality which yields

$$\begin{aligned} &\|z_j^n(\Theta_t^n) - p_j(t)\|_E^2 \\ &\leq 4 \|z_j^n(\Theta_0^n) - p_j(0)\|_E^2 + 4 \|M_j^n(t)\|_E^2 + 4 \left(\int_0^t \|F_j(z^n(\Theta_s^n), U_s^n) - F_j(p(s), u(s))\|_E ds \right)^2 \\ &\quad + 4 \left(\int_0^t \|[\mathcal{A}^n \langle \cdot, z_j^n(\cdot) \rangle_E](U_s^n, \Theta_s^n) - F_j(z^n(\Theta_s^n), U_s^n)\|_E ds \right)^2. \end{aligned}$$

We next apply the Lipschitz condition (4.4) on F and obtain the estimate

$$\begin{aligned} &\|z_j^n(\Theta_t^n) - p_j(t)\|_E^2 \\ &\leq 4 \|z_j^n(\Theta_0^n) - p_j(0)\|_E^2 + 4L_2 \int_0^t \|U_s^n - u(s)\|_H^2 ds + 4L_2 \sum_{i=1}^m \int_0^t \|z_i^n(\Theta_s^n) - p_i(s)\|_E^2 ds \\ &\quad + 4 \left(\int_0^t \|[\mathcal{A}^n \langle \cdot, z_j^n(\cdot) \rangle_E](U_s^n, \Theta_s^n) - F_j(z^n(\Theta_s^n), U_s^n)\|_E ds \right)^2 + \|M_j^n(t)\|_E^2. \quad (6.5) \end{aligned}$$

To further estimate this term we employ the convergence (4.5) of the term $\|M_j^n\|_E$ and the convergence (4.6) of the generator. It follows by the definition of these limits that for every $\epsilon_1 > 0$ and every $\delta > 0$ we can find an $N_{\epsilon_1, \delta}$ such that for all $n \geq N_{\epsilon_1, \delta}$ it holds due to (4.5) for all $j = 1, \dots, m$ and all $t \in [0, T]$ that

$$\|M_j^n(t)\|_E \leq \sqrt{\frac{\epsilon_1}{m}},$$

and due to (4.6) and the Continuous mapping Theorem that

$$\left(\int_0^T \|[\mathcal{A}^n \langle \cdot, z_j^n(\cdot) \rangle_E](U_s^n, \Theta_s^n) - F_j(z^n(\Theta_s^n), U_s^n)\|_E ds \right)^2 \leq \frac{\epsilon_1}{m}$$

on a set $\Omega_1 \subset \Omega$ satisfying $\mathbb{P}^n(\Omega \setminus \Omega_1) \leq \delta$ for all $n \geq N_{\epsilon_1, \delta}$. Thus continuing to estimate only for paths on the set Ω_1 we obtain from (6.5) the inequality

$$\begin{aligned} \|z_j^n(\Theta_t^n) - p_j(t)\|_E^2 &\leq 4 \|z_j^n(\Theta_0^n) - p_j(0)\|_E^2 + 5 \frac{\epsilon_1}{m} \\ &\quad + 4L_2 \int_0^t \|U_s^n - u(s)\|_H^2 ds + 4L_2 \sum_{i=1}^m \int_0^t \|z_i^n(\Theta_s^n) - p_i(s)\|_E^2 ds. \end{aligned} \quad (6.6)$$

In order to finally obtain the growth estimate suitable for an application of Gronwall's

inequality we add inequality (6.3) and inequalities (6.6) for all $j = 1, \dots, m$ which yields

$$\begin{aligned} \|U_t^n - u(t)\|_H^2 + \sum_{j=1}^m \|z_j^n(\Theta_t^n) - p_j(t)\|_E^2 &\leq \|U_0^n - u_0\|_H^2 + 4 \sum_{j=1}^m \|z_j^n(\Theta_0^n) - p_j(0)\|_E^2 \quad (6.7) \\ &+ 5\epsilon_1 + C \int_0^t \|U_s^n - u(s)\|_H^2 ds + C \sum_{j=1}^m \int_0^t \|z_j^n(\Theta_s^n) - p_j(s)\|_E^2 ds \end{aligned}$$

with constant $C = 2L_1 + 4L_2m$. An application of Gronwall's inequality to (6.7) yields

$$\sup_{t \in [0, T]} \left(\|U_t^n - u(t)\|_H^2 + \sum_{j=1}^m \|z_j^n(\Theta_t^n) - p_j(t)\|_E^2 \right) \leq K_1 e^{CT} \quad (6.8)$$

where

$$K_1 = \|U_0^n - u_0\|_H^2 + 4 \sum_{j=1}^m \|z_j^n(\Theta_0^n) - p_j(0)\|_E^2 + 5\epsilon_1.$$

Finally, due to (C3), i.e., the convergence in probability of the initial conditions, it holds that for every $\epsilon_2 > 0$ we can find to every $\delta > 0$ an $N_{\epsilon_2, \delta}$ such that on a set $\Omega_2 \subset \Omega$ with $\mathbb{P}^n(\Omega \setminus \Omega_2) < \delta$ it holds for all $n \geq N_{\epsilon_2, \delta}$ that

$$\|U_0^n - u_0\|_H^2 \leq \frac{\epsilon_2}{m+2}, \quad \|z_j^n(\Theta_0^n) - p_j(0)\|_E^2 \leq \frac{\epsilon_2}{4(m+2)} \quad \forall j = 1, \dots, m. \quad (6.9)$$

Let $\epsilon, \delta > 0$ be arbitrary. Then we obtain choosing $\epsilon_2 = \epsilon e^{-CT}$ and $\epsilon_1 = \frac{\epsilon_2}{5(m+2)}$, thus $K_1 = \epsilon_2$, that for all $n \geq N_{\epsilon, \delta} := N_{\epsilon_1, \delta} \vee N_{\epsilon_2, \delta}$ it holds that

$$\sup_{t \in [0, T]} \left(\|U_t^n - u(t)\|_H^2 + \sum_{j=1}^m \|z_j^n(\Theta_t^n) - p_j(t)\|_E^2 \right) \leq \epsilon$$

on the set $\Omega_1 \cap \Omega_2$. Therefore it holds for all $n \geq N_{\epsilon, \delta}$ that

$$\mathbb{P}^n \left[\sup_{t \in [0, T]} \left(\|U_t^n - u(t)\|_H^2 + \sum_{j=1}^m \|z_j^n(\Theta_t^n) - p_j(t)\|_E^2 \right) > \epsilon \right] \leq 2\delta.$$

As δ and ϵ are arbitrary the statement (4.7) follows.

6.2 Proof of Theorem 5.1 (Central limit theorem)

The proof of Theorem 5.1 is split into three successive steps. In the first step we prove tightness of the sequence of martingales which guarantees the existence of a limit. Secondly, we show that any limit is a continuous process. Finally, in the last step we prove that the limit is the specific diffusion process as stated in the theorem. The conditions (D1)–(D3) in Theorem 5.1 are such that each, in addition, to the preceding is needed in the successive steps of the proof.

Tightness

In order to prove tightness of the sequence of \mathcal{E} -valued martingales $(\sqrt{\alpha_n} M_t^n)_{t \geq 0}$ it suffices to show that the following conditions are satisfied, cf. [36] wherein general conditions for tightness of sequences of Hilbert space valued processes and, in particular, martingales are considered:

(T1) The sequence of initial conditions $(\sqrt{\alpha_n} M_0^n)_{n \geq 0}$ is tight.

(T2) For all $t \geq 0$ it holds that

$$\lim_{\delta \rightarrow \infty} \sup_{n \in \mathbb{N}} \mathbb{P}^n [\text{Tr} \llcorner \sqrt{\alpha_n} M^n \ggcorner_t > \delta] = 0, \tag{6.10}$$

and there exists an orthonormal basis $(\varphi_k)_{k \in \mathbb{N}}$ of E^* such that for each $\epsilon > 0$

$$\lim_{m \rightarrow \infty} \sup_{n \in \mathbb{N}} \mathbb{P}^n \left[\sum_{k > m} \langle \varphi_k, \llcorner \sqrt{\alpha_n} M^n \ggcorner_t \varphi_k \rangle_E > \epsilon \right] = 0. \tag{6.11}$$

(A) The sequence of the real-valued trace processes $(\text{Tr} \llcorner \sqrt{\alpha_n} M^n \ggcorner_t)_{t \geq 0}$, $n \geq \mathbb{N}$, satisfies the *Aldous condition*: For every $T, \epsilon, \delta > 0$ there exists a $h > 0$ and an $N > 0$ such that for any sequence of stopping times⁸ $(\sigma^n)_{n \geq 0}$ with $\sigma^n \leq T$ it is valid that

$$\sup_{n \geq N} \sup_{0 \leq s \leq h} \mathbb{P}^n [|\text{Tr} \llcorner \sqrt{\alpha_n} M^n \ggcorner_{\sigma^n + s} - \text{Tr} \llcorner \sqrt{\alpha_n} M^n \ggcorner_{\sigma^n}| \geq \delta] \leq \epsilon. \tag{6.12}$$

We next establish the above conditions. First note that condition (T1) is trivially satisfied as $M_0^n = 0$ for all $n > 0$. Hence we proceed to condition (T2). In order to establish the first condition (6.10) we use Markov's inequality to obtain the estimate

$$\mathbb{P}^n [\text{Tr} \llcorner \sqrt{\alpha_n} M^n \ggcorner_t > \delta] \leq \frac{\alpha_n}{\delta} \mathbb{E}^n \left[\int_0^t \text{Tr} G^n(Y_s^n, \theta_s^n) ds \right],$$

where the right hand side is finite due to assumption (5.3). Taking the supremum on both sides the same assumption implies (6.10).

Next, in order to show the second condition (6.11) we employ Markov's inequality, the monotone convergence theorem (in order to change the order of expectation and the countable summation over all $k > m$), the form of the quadratic variation (3.11) and inequality (5.4) to obtain for the term in the left hand side the estimates

$$\begin{aligned} \mathbb{P}^n \left[\sum_{k > m} \langle \varphi_k, \llcorner \sqrt{\alpha_n} M^n \ggcorner_t \varphi_k \rangle_{\mathcal{E}} > \delta \right] &\leq \frac{\alpha_n}{\delta} \mathbb{E}^n \left[\sum_{k > m} \langle \varphi_k, \llcorner M^n \ggcorner_t \varphi_k \rangle_{\mathcal{E}} \right] \\ &\leq \frac{1}{\delta} \left(\sum_{k > m} \gamma_k \right) C(t), \end{aligned}$$

where the upper bound is independent of $n \in \mathbb{N}$. Moreover, the property $\sum_{k \in \mathbb{N}} \gamma_k < \infty$ implies that $\lim_{m \rightarrow \infty} \sum_{k > m} \gamma_k = 0$ and hence (6.11) holds for all $t \geq 0$.

Finally, it remains to show (A). Let $T, \delta > 0$ and $\sigma^n < T$ be an arbitrary sequence of stopping times, then for all for all $h > 0$ it holds that for $s \leq h$

$$\begin{aligned} \mathbb{P}^n \left[\left| \alpha_n \text{Tr} \llcorner M^n \ggcorner_{\sigma^n + s} - \alpha_n \text{Tr} \llcorner M^n \ggcorner_{\sigma^n} \right| \geq \delta \right] &\leq \frac{\alpha_n}{\delta} \mathbb{E}^n \left[\int_{\sigma^n}^{\sigma^n + h} \text{Tr} G^n(Y_r^n, \theta_r^n) dr \right] \\ &= \frac{1}{\delta} \mathbb{E}^n \left[\sum_{k \in \mathbb{N}} \alpha_n \mathbb{E}^{X_{\sigma^n}^n} \int_0^h \langle \varphi_k, G^n(Y_r^n, \theta_r^n) \varphi_k \rangle_{\mathcal{E}} dr \right] \\ &\leq \frac{C(h)}{\delta} \sum_{k \in \mathbb{N}} \gamma_k. \end{aligned}$$

Here we have used Markov's inequality, the strong Markov property of the PDMP and the assumption (5.4). As the final upper bound is independent of s and n and converges

⁸Here every σ^n is a stopping time on the respective probability space $(\Omega^n, \mathcal{F}^n, (\mathcal{F}_t^n)_{t \geq 0}, \mathbb{P}^n)$.

to zero for $h \rightarrow 0$ condition (A) follows.

Any Limit is a continuous process

In the preceding part of the proof we have established that the laws of the sequence of martingales $(\sqrt{\alpha_n} M_t^n)_{t \geq 0}$ are tight which is equivalent to there existence of a weakly convergent subsequence. We now prove that under the additional condition (D2) any cluster point of the sequence is a measure supported on $C(\mathbb{R}_+, \mathcal{E})$. The method of proof follows the outline of [25, Lemma 3.2] adapted for the stochastic processes being PDMPs on Hilbert spaces, the general setup in this study and the particular conditions (D1) and (D2) in Theorem 5.1 which differ from [25]. Furthermore, we have extended the result in [25, Lemma 3.2], which only considers convergence on finite time intervals $[0, T]$, to convergence on $D(\mathbb{R}_+, \mathcal{E})$. In the following we employ the abbreviations $Z_t^n := \sqrt{\alpha_n} M_t^n$ and $\Delta_t Z^n := Z_t^n - Z_{t-}^n$, i.e., $(\Delta_t Z^n)_{t \geq 0}$ denotes the process of jump heights. Note that $\Delta_t Z^n = \sqrt{\alpha_n} \Delta_t z^n(\theta^n)$.

Further, let \mathbb{P}^* denote an accumulation point of the sequence $(\mathbb{P}^n)_{n \in \mathbb{N}}$. Without loss of generality we use \mathbb{P}^n , $n \geq 1$, to also denote the subsequence converging weakly to \mathbb{P}^* . Furthermore, here \mathbb{P}^n is understood as a law on the Skorokhod space $D(\mathbb{R}_+, \mathcal{E})$ given by the pushforward measure of the process $(\sqrt{\alpha_n} M_t^n)_{t \geq 0}$. Then due to the Skorokhod Representation Theorem, e.g., [18, Chap. 3, Thm. 1.8], there exists a probability space $(\Omega^o, \mathcal{F}^o, \mathbb{P}^o)$ supporting $D(\mathbb{R}_+, \mathcal{E})$ -valued random variables ζ^n , $n \geq 1$, and ζ^* with distributions \mathbb{P}^n and \mathbb{P}^* , respectively, such that ζ^n converges to ζ^* almost surely with respect to \mathbb{P}^o . Further, it clearly holds that $\mathbb{E}^n f(Z^n) = \mathbb{E}^o f(\zeta^n)$ for suitable functionals f .

We begin the proof with preliminary estimates on functions evaluated along the path of the PDMPs. These ultimately allow to infer that the process of jumps vanishes in the limit. Let g be a measurable, bounded, non-negative function $g : \mathbb{R} \rightarrow \mathbb{R}$, that vanishes in a neighbourhood of 0 and of ∞ , that is, there exists a finite constant $C_g := \sup_{x \in \mathbb{R}} g(x)/x^2 < \infty$. For such a function g and any $\Phi \in \mathcal{E}^*$ we define the process

$$\begin{aligned} G_t^n(\langle \Phi, Z^n \rangle_{\mathcal{E}}) &:= \sum_{s \in (0, t]} g(\langle \Phi, \Delta_s Z^n \rangle_{\mathcal{E}}) \\ &\quad - \int_0^t \Lambda^n(U_s^n, \Theta_s^n) \int_{K_n} g(\sqrt{\alpha_n} \langle \Phi, z^n(\xi) - z^n(\Theta_s^n) \rangle_{\mathcal{E}}) \mu^n((U_s^n, \Theta_s^n)) \, ds \\ &= \int_0^t \int_{K_n} g(\sqrt{\alpha_n} \langle \Phi, z^n(\xi) - z^n(\Theta_{s-}^n) \rangle_{\mathcal{E}}) M^n(d\xi, ds), \end{aligned}$$

where M^n is the martingale measure associated with the PDMP, and we infer that $G_t^n(\langle \Phi, Z^n \rangle_{\mathcal{E}})$ is a martingale. Note that the above summation over all $s \in (0, t]$ is well-defined as the PDMPs are regular and thus $g(\langle \Phi, \Delta_s Z^n \rangle_{\mathcal{E}})$ is non-zero for only finitely many $s \leq t$.

The proof now proceeds as follows. We first show (a) that for all $t \geq 0$ the random variables $G_t^n(\langle \Phi, \zeta^n \rangle_{\mathcal{E}})$, $n \in \mathbb{N}$, are uniformly integrable for all t and (b) that they converge to $\sum_{s \in (0, t]} g(\langle \Phi, \Delta_s \zeta^n \rangle_{\mathcal{E}})$ in probability. This allows to infer that the convergence result also holds as convergence in mean. In part (c) we then use these results to show that the jump heights of the canonical process of the law \mathbb{P}^* are constantly zero almost surely. This implies that $\mathbb{P}^*(C([0, t], \mathcal{E})) = 1$ for every $t > 0$ where $C([0, t], \mathcal{E})$ is understood as the subset of $D(\mathbb{R}_+, \mathcal{E})$ consisting of those càdlàg functions which are continuous up to and including time t . The proof is completed by (d) extending this result to $\mathbb{P}^*(C(\mathbb{R}_+, \mathcal{E})) = 1$.

(a) To show that the sequence of random variables $G_t^n(\langle \Phi, \zeta^n \rangle_{\mathcal{E}})$, $n \in \mathbb{N}$, is uniformly integrable in the space $(\Omega^o, \mathcal{F}^o, \mathbb{P}^o)$ it is sufficient that the second moments are uni-

formly bounded, cf. [18, Appendix, Prop. 2.2]. The Itô-isometry for real-valued stochastic integrals with respect to the associated martingale measures, which is implied by taking the expectation of the processes in [24, Prop. 4.5.3], yields

$$\begin{aligned} \sup_{n \in \mathbb{N}} \mathbb{E}^o |G_t^n(\langle \Phi, \zeta^n \rangle_{\mathcal{E}})|^2 &= \sup_{n \in \mathbb{N}} \mathbb{E}^n |G_t^n(\langle \Phi, Z^n \rangle_{\mathcal{E}})|^2 \\ &= \sup_{n \in \mathbb{N}} \mathbb{E}^n \left[\int_0^t \Lambda^n(U_s^n, \Theta_s^n) \int_{K_n} g(\sqrt{\alpha_n} \langle \Phi, z^n(\xi) - z^n(\Theta_s^n) \rangle_{\mathcal{E}})^2 \mu^n((U_s^n, \theta_s^n), d\xi) ds \right]. \end{aligned}$$

Therefore, employing the special structure of the map g we obtain the estimate

$$\begin{aligned} \sup_{n \in \mathbb{N}} \mathbb{E}^o |G_t^n(\langle \Phi, \zeta^n \rangle_{\mathcal{E}})|^2 \\ \leq C_g \sup_{n \in \mathbb{N}} \alpha_n \mathbb{E}^n \left[\int_0^t \Lambda^n(U_s^n, \Theta_s^n) \int_{K_n} |\langle \Phi, z^n(\xi) - z^n(\Theta_s^n) \rangle_{\mathcal{E}}|^2 \mu^n((U_s^n, \Theta_s^n), d\xi) ds \right], \end{aligned}$$

where the right hand side is finite for every $t > 0$ due to condition (5.3) in (D1).

(b) In this part of the proof we establish convergence in probability of the random variables $G_t^n(\langle \Phi, \Delta \zeta^n \rangle_{\mathcal{E}})$. Let $\beta > 0$ be such that $g(x) = 0$ for $|x| \leq \beta$, i.e., the interval $(-\beta, \beta)$ is contained in the neighbourhood of 0 whereon g vanishes. Then we obtain using Markov's inequality and due to the boundedness of g the estimates

$$\begin{aligned} \mathbb{P}^o \left[\sum_{s \in (0,t]} g(\langle \Phi, \Delta_s \zeta^n \rangle_{\mathcal{E}}) - G_t^n(\langle \Phi, \zeta^n \rangle_{\mathcal{E}}) > \delta \right] \\ = \mathbb{P}^n \left[\int_0^t \Lambda^n(U_s^n, \Theta_s^n) \int_{K_n} g(\sqrt{\alpha_n} \langle \Phi, z^n(\xi) - z^n(\Theta_s^n) \rangle_E) \mu^n((U_s^n, \Theta_s^n), d\xi) ds > \delta \right] \\ \leq \frac{1}{\delta} \mathbb{E}^n \left[\int_0^t \Lambda^n(U_s^n, \Theta_s^n) \int_{K_n} g(\sqrt{\alpha_n} \langle \Phi, z^n(\xi) - z^n(\Theta_s^n) \rangle_E) \mu^n((U_s^n, \Theta_s^n), d\xi) ds \right] \\ \leq \frac{\sup_{x \in \mathbb{R}} |g(x)|}{\delta} \mathbb{E}^n \left[\int_0^t \Lambda^n(U_s^n, \Theta_s^n) \int_{\sqrt{\alpha_n} |\langle \Phi, z^n(\xi) - z^n(\Theta_s^n) \rangle_E| > \beta} \mu^n((U_s^n, \Theta_s^n), d\xi) ds \right]. \end{aligned}$$

Thus due to condition (5.5) in (D2) it holds that

$$\lim_{n \rightarrow \infty} \mathbb{P}^o \left[\sum_{s \in (0,t]} g(\langle \Phi, \Delta_s \zeta^n \rangle_{\mathcal{E}}) - G_t^n(\langle \Phi, \zeta^n \rangle_{\mathcal{E}}) > \delta \right] = 0.$$

Moreover, it holds on $(\Omega^o, \mathcal{F}^o, \mathbb{P}^o)$ almost surely that

$$\lim_{n \rightarrow \infty} \sum_{s \in (0,t]} g(\langle \Phi, \Delta_s \zeta^n \rangle_{\mathcal{E}}) = \sum_{s \in (0,t]} g(\langle \Phi, \Delta_s \zeta^* \rangle_{\mathcal{E}}).$$

Therefore, combining these two convergence results we obtain that

$$G_t^n(\langle \Phi, \zeta^n \rangle_{\mathcal{E}}) \longrightarrow \sum_{s \in (0,t]} g(\langle \Phi, \Delta_s \zeta^* \rangle_{\mathcal{E}}) \tag{6.13}$$

holds as convergence in probability.

(c) From parts (a) and (b) we infer that (6.13) also holds as convergence in mean.

Together with Jensen's inequality this implies

$$\begin{aligned} \lim_{n \rightarrow \infty} \left| \mathbb{E}^o \left(G_t^n(\langle \Phi, \Delta_s \zeta^n \rangle_{\mathcal{E}}) - \sum_{s \in (0, T]} g(\langle \Phi, \Delta_s \zeta^* \rangle_{\mathcal{E}}) \right) \right| \\ \leq \lim_{n \rightarrow \infty} \mathbb{E}^o \left| G_t^n(\langle \Phi, \Delta_s \zeta^n \rangle_{\mathcal{E}}) - \sum_{s \in (0, t]} g(\langle \Phi, \Delta_s \zeta^* \rangle_{\mathcal{E}}) \right| = 0, \end{aligned}$$

and hence we infer that

$$\mathbb{E}^o \sum_{s \in (0, t]} g(\langle \Phi, \Delta_s \zeta^* \rangle_{\mathcal{E}}) = \lim_{n \rightarrow \infty} \mathbb{E}^o G_t^n(\langle \Phi, \zeta^n \rangle_{\mathcal{E}}). \tag{6.14}$$

Furthermore, $G_t^n(\langle \Phi, Z^n \rangle_{\mathcal{E}})$ is a martingale which satisfies $G_0^n(\langle \Phi, Z^n \rangle_{\mathcal{E}}) = 0$. This, in turn, implies that $\mathbb{E}^n G_t^n(\langle \Phi, Z^n \rangle_{\mathcal{E}}) = 0$ for every $n \in \mathbb{N}$. Therefore we obtain due to (6.14)

$$\begin{aligned} \mathbb{E}^* \sum_{s \in (0, t]} g(\langle \Phi, \Delta_s Z \rangle_{\mathcal{E}}) &= \mathbb{E}^o \sum_{s \in (0, t]} g(\langle \Phi, \Delta_s \zeta^* \rangle_{\mathcal{E}}) \\ &= \lim_{n \rightarrow \infty} \mathbb{E}^o G_t^n(\langle \Phi, \zeta^n \rangle_{\mathcal{E}}) = \lim_{n \rightarrow \infty} \mathbb{E}^n G_t^n(\langle \Phi, Z^n \rangle_{\mathcal{E}}) = 0. \end{aligned} \tag{6.15}$$

In a next step, let g_m be a sequence of functions satisfying the properties for functions g proposed above. Further we assume that the functions $g_m(x)$ increase pointwise to x^2 for $m \rightarrow \infty$ (for an example of such functions we refer to [25]). Then due to the monotone convergence theorem it holds that

$$\lim_{m \rightarrow \infty} \mathbb{E}^* \sum_{s \in (0, t]} g_m(\langle \Phi, \Delta_s Z \rangle_{\mathcal{E}}) = \mathbb{E}^* \sum_{s \in (0, t]} |\langle \Phi, \Delta_s Z \rangle_{\mathcal{E}}|^2.$$

Furthermore, the limiting expectation in the right hand side is zero as each element of the sequence of expectations in the left hand side is zero due to (6.15). Next we choose Φ to be an element of an orthonormal basis $(\varphi_k)_{k \in \mathbb{N}}$ of \mathcal{E} and sum the expectations over all elements of the basis yielding

$$\sum_{k \in \mathbb{N}} \mathbb{E}^* \sum_{s \in (0, t]} |\langle \varphi_k, \Delta_s Z \rangle_{\mathcal{E}}|^2.$$

Due to the dominated convergence theorem we can interchange the countable summation and the expectation and, as the PDMP is regular, we afterwards interchange the resulting two summation inside the expectation. Then Parseval's identity yields

$$\mathbb{E}^* \sum_{s \in (0, t]} \|\Delta_s Z\|_{\mathcal{E}}^2 = 0.$$

As the non-negative random variable inside the expectation is zero only for continuous paths of the process $(Z_s)_{s \in [0, t]}$ we infer that almost all paths are continuous, i.e., $\mathbb{P}^*(C([0, t], \mathcal{E})) = 1$.

(d) To conclude the proof let $t_k, k \in \mathbb{N}$, be a sequence of times increasing to infinity then

$$C(\mathbb{R}_+, \mathcal{E}) = \bigcap_{k \in \mathbb{N}} C([0, t_k], \mathcal{E}),$$

and the events in the right hand side satisfy $C([0, t_{k+1}], \mathcal{E}) \subseteq C([0, t_k], \mathcal{E})$. The properties

of a probability measure thus yield

$$\mathbb{P}^*(C(\mathbb{R}_+, \mathcal{E})) = \lim_{k \rightarrow \infty} \mathbb{P}^*(C([0, t_k], \mathcal{E})) = 1,$$

that is a process with distribution given by the limit \mathbb{P}^* possesses almost surely continuous paths.

Limit is a diffusion process

In the final part of the proof we uniquely characterise the limit of the sequence of martingales $(\sqrt{\alpha_n} M_t^n)_{t \geq 0}$ under the additional assumptions (D3). The method of proof is via the local martingale problem motivated by a proof presented in [36], i.e., the limiting probability measure is the unique solution to a particular martingale problem. The author in [36] considers Hilbert space valued stochastic integral equations driven by Hilbert space valued martingales with state dependent quadratic variation. A central limit theorem for the martingales is presented. The arguments of the subsequent proof are closely related to [36]. This is as the general result on martingales associated with PDMPs, which we have proven in Section 3, result in the problem in this part of the proof to be of the same underlying structure as in [36]. One difference, however, is that the present conditions (D1)–(D3) are more general than the conditions in [36] and adapted to the PDMP setup, hence some estimates differ.

As in the preceding part of the proof we interpret the sequence of martingales $(\sqrt{\alpha_n} M_t^n)_{t \geq 0}$ defined on the probability spaces $(\Omega^n, \mathcal{F}^n, (\mathcal{F}_t^n)_{t \geq 0}, \mathbb{P}^n)$ as random variables on the space $D(\mathbb{R}_+, \mathcal{E})$ equipped with its natural σ -field \mathcal{D} . Further, laws on the canonical space are given by the pushforward measure. In order to simplify the notation we denote the laws on the canonical space also by \mathbb{P}^n . Due to results in the preceding two parts of the proof we know the sequence \mathbb{P}^n , $n \in \mathbb{N}$, admits a limit \mathbb{P}^* supported on $C(\mathbb{R}_+, \mathcal{E})$. We use $(\zeta_t)_{t \geq 0}$ to denote the canonical process on $D(\mathbb{R}_+, \mathcal{E})$ which is a version of the martingale $(\sqrt{\alpha_n} M_t^n)_{t \geq 0}$ under the push-forward measure \mathbb{P}^n for all $n \in \mathbb{N}$ or of the weak limit under the measure \mathbb{P}^* .

In the following we prove that the limit \mathbb{P}^* is a solution to a local martingale problem the unique solution of which is an \mathcal{E} -valued centered diffusion process with covariance operator $C(t) \in L_1(\mathcal{E}^*, \mathcal{E})$ as given in (5.2). For any twice continuously differentiable function $f : \mathcal{E} \rightarrow \mathbb{R}$ the extended generator $\mathcal{A}f$ of such a diffusion is given by

$$\mathcal{A}f(x, t) = \frac{1}{2} \text{Tr}(D^2 f(x) \circ G(t)).$$

Then, in order to uniquely characterise the solution to the local martingale problem connected with this generator and supported on the space $C(\mathbb{R}, \mathcal{E})$ it suffices to consider mappings f of the form $\langle \Phi, \cdot \rangle_{\mathcal{E}}$ and $\langle \Phi, \cdot \rangle_{\mathcal{E}}^2$ for all $\Phi \in \mathcal{E}^*$, cf. [36]. That is, we have to show that the canonical process ζ_t is such that for all $\Phi \in \mathcal{E}^*$ the processes $\langle \Phi, \zeta_t \rangle_{\mathcal{E}}$ and

$$\langle \Phi, \zeta_t \rangle_{\mathcal{E}}^2 - \int_0^t \langle \Phi, G(u_s, p_s) \Phi \rangle_{\mathcal{E}} ds \tag{6.16}$$

are \mathbb{P}^* -local martingales. We start introducing some notation and then show in parts (a) and (b) the local martingale properties of the two indicated processes on the canonical space $D(\mathbb{R}_+, \mathcal{E})$.

As before we use $Z_t^n := \sqrt{\alpha_n} M_t^n$ and $\Delta_t Z^n := Z_t^n - Z_{t-}^n$. Further, as indicated above the notation is such that we use \mathbb{P}^n and \mathbb{E}^n to denote probabilities and expectations on the original given measurable spaces $(\Omega^n, \mathcal{F}^n)$ as well as on the canonical space $(D(\mathbb{R}_+, \mathcal{E}), \mathcal{D})$. That is, e.g., $\mathbb{E}^n f(Z_t^n) = \mathbb{E}^n f(\zeta_t)$ for any bounded function f , where the

former is the expectation taken on the original space $(\Omega^n, \mathcal{F}^n, \mathbb{P}^n)$ and the latter the expectation on the canonical space of càdlàg processes with respect to the pushforward measure. Furthermore, we employ the Itô-formula [35, Thm. 25.7] for smooth functions $f \in C_c^\infty(\mathbb{R})$ applied to semi-martingales. For the particular choice of the semi-martingales being the real martingales $\langle \Phi, Z_t^n \rangle_\mathcal{E}$ the Itô-formula reads

$$\begin{aligned} f(\langle \Phi, Z_t^n \rangle_\mathcal{E}) &= \frac{1}{2} \int_0^t f''(\langle \Phi, Z_{s-}^n \rangle_\mathcal{E}) (\langle \Phi, \alpha_n \ll M^n \gg_t \Phi \rangle_\mathcal{E}) ds \\ &+ \sum_{s \leq t} \left[f(\langle \Phi, Z_s^n \rangle_\mathcal{E}) - f(\langle \Phi, Z_{s-}^n \rangle_\mathcal{E}) - \langle \Phi, \Delta_s Z^n \rangle_\mathcal{E} f'(\langle \Phi, Z_{s-}^n \rangle_\mathcal{E}) \right] \\ &- \frac{1}{2} \sum_{s \leq t} \left[\langle \Phi, \Delta_s Z^n \rangle_\mathcal{E}^2 f''(\langle \Phi, Z_t^n \rangle_\mathcal{E}) \right] + M_t^{f,n} \end{aligned} \tag{6.17}$$

where $(M_t^{f,n})_{t \geq 0}$ is some martingale on $(\Omega^n, \mathcal{F}^n, (\mathcal{F}_t^n)_{t \geq 0}, \mathbb{P}^n)$ depending on Z^n and f .

Next, we introduce on the canonical space for all positive ρ the stopping times $\tau_\rho := \inf\{t \in \mathbb{R}_+ \mid \|\zeta_t\|_\mathcal{E} > \rho\}$ and note that due to the bound (5.7) in (D3) on the jump heights we have that for any law \mathbb{P}^n , $n \geq 1$, it holds almost surely

$$\|\zeta_{\tau_\rho}\|_\mathcal{E} \leq \rho + C. \tag{6.18}$$

Analogously we define the stopping times $\tau_\rho^n := \inf\{t \in \mathbb{R}_+ \mid \|Z_t^n\|_\mathcal{E} > \rho\}$ on the spaces $(\Omega^n, \mathcal{F}^n, (\mathcal{F}_t^n)_{t \geq 0}, \mathbb{P}^n)$.

Finally, as already mentioned $(\mathcal{D}_t)_{t \geq 0}$ denotes the natural filtration on the canonical space. Then for $A \in \mathcal{D}_t$ we define $A^n := (Z^n)^{-1}A \in \mathcal{F}_t^n$ its preimage with respect to the random variable Z^n . We now proceed to show that the two processes $\langle \Phi, \zeta_t \rangle_\mathcal{E}$ and (6.16) are indeed local martingales with respect to the limit measure \mathbb{P}^* .

(a) Let $\Phi \in \mathcal{E}^*$ be fixed and we choose for every ρ a smooth function $f_\rho \in C_c^\infty(\mathbb{R})$ which satisfies $f_\rho(x) = x$ if $|x| \leq \|\Phi\|_{\mathcal{E}^*}(\rho + C)$ and thus $f'_\rho(x) = 1$ and $f''_\rho(x) = 0$ for $|x| \leq \|\Phi\|_{\mathcal{E}^*}(\rho + C)$. Therefore it holds for $t < \tau_\rho^n$, which implies the estimate $|\langle \Phi, Z_{t-}^n \rangle_\mathcal{E}| \leq \|\Phi\|_{\mathcal{E}^*}(\rho + C)$, that

$$f''_\rho(\langle \Phi, Z_{t-}^n \rangle_\mathcal{E}) = 0$$

and

$$f_\rho(\langle \Phi, Z_t^n \rangle_\mathcal{E}) - f_\rho(\langle \Phi, Z_{t-}^n \rangle_\mathcal{E}) - \langle \Phi, \Delta_t Z^n \rangle_\mathcal{E} f'_\rho(\langle \Phi, Z_{s-}^n \rangle_\mathcal{E}) = 0.$$

It follows that applying the Itô-formula (6.17) to the function f_ρ and the martingale $Z_{t \wedge \tau_\rho^n}^n$ all terms besides the martingale M^{n, f_ρ} vanish in the the right hand side. Therefore we obtain for $t_2 \geq t_1$ and all $A \in \mathcal{D}_{t_1}$ that

$$\begin{aligned} \mathbb{E}^n \left[\mathbb{I}_A \left(\langle \Phi, \zeta_{t_2 \wedge \tau_\rho} \rangle_\mathcal{E} - \langle \Phi, \zeta_{t_1 \wedge \tau_\rho} \rangle_\mathcal{E} \right) \right] &= \mathbb{E}^n \left[\mathbb{I}_A \left(f_\rho(\langle \Phi, \zeta_{t_2 \wedge \tau_\rho} \rangle_\mathcal{E}) - f_\rho(\langle \Phi, \zeta_{t_1 \wedge \tau_\rho} \rangle_\mathcal{E}) \right) \right] \\ &= \mathbb{E}^n \left[\mathbb{I}_{A^n} \left(f_\rho(\langle \Phi, Z_{t_2 \wedge \tau_\rho^n}^n \rangle_\mathcal{E}) - f_\rho(\langle \Phi, Z_{t_1 \wedge \tau_\rho^n}^n \rangle_\mathcal{E}) \right) \right] \\ &= 0. \end{aligned} \tag{6.19}$$

The proof of the first martingale property is concluded as in [36]: The mapping $\zeta \rightarrow f_\rho(\langle \Phi, \zeta_{t_2 \wedge \tau_\rho} \rangle_\mathcal{E})$ is almost surely (with respect to the probability \mathbb{P}^*) continuous and as \mathbb{P}^n converges weakly to \mathbb{P}^* it holds due to (6.19) that

$$\mathbb{E}^* \left[\mathbb{I}_A \left(\langle \Phi, \zeta_{t_2 \wedge \tau_\rho} \rangle_\mathcal{E} - \langle \Phi, \zeta_{t_1 \wedge \tau_\rho} \rangle_\mathcal{E} \right) \right] = 0.$$

Here we have employed a weaker version of the continuous mapping theorem, see, e.g.,

[6, Thm. 2.7].

We infer from the definition of the conditional expectation that the stopped processes are martingales. Furthermore, as ζ_t possesses continuous paths almost surely under the measure \mathbb{P}^* it holds that τ_ρ diverges to ∞ almost surely for $\rho \rightarrow \infty$. Hence, we can find a sequence of stopping times τ_{ρ_k} , $k \in \mathbb{N}$, such that $\tau_{\rho_k} \rightarrow \infty$ almost surely for $k \rightarrow \infty$. Thus, the process $\langle \Phi, \zeta_t \rangle$ is a local martingale with respect to \mathbb{P}^* .

(b) For the second class of processes we consider smooth functions $g_\rho \in C_c^\infty(\mathbb{R})$ such that $g_\rho(x) = x^2$ for all $|x| \leq \|\Phi\|_{\mathcal{E}^*}(\rho + C)$. Starting from the definition of the conditional expectation as in (6.19) we obtain

$$\begin{aligned} & \mathbb{E}^n \left[\mathbb{I}_A \left(\langle \Phi, \zeta_{t_2 \wedge \tau_\rho} \rangle_{\mathcal{E}}^2 - \int_0^{t_2 \wedge \tau_\rho} \langle \Phi, G(u(s), p(s)) \Phi \rangle_{\mathcal{E}} ds - \langle \Phi, \zeta_{t_1 \wedge \tau_\rho} \rangle_{\mathcal{E}}^2 \right. \right. \\ & \qquad \qquad \qquad \left. \left. + \int_0^{t_1 \wedge \tau_\rho} \langle \Phi, G(u(s), p(s)) \Phi \rangle_{\mathcal{E}} ds \right) \right] \\ &= \mathbb{E}^n \left[\mathbb{I}_A \left(\langle \Phi, \zeta_{t_2 \wedge \tau_\rho} \rangle_{\mathcal{E}}^2 - \langle \Phi, \zeta_{t_1 \wedge \tau_\rho} \rangle_{\mathcal{E}}^2 - \int_{t_1 \wedge \tau_\rho}^{t_2 \wedge \tau_\rho} \langle \Phi, G(u(s), p(s)) \Phi \rangle_{\mathcal{E}} ds \right) \right] \\ &= \mathbb{E}^n \left[\mathbb{I}_{A^n} \left(\langle \Phi, Z_{t_2 \wedge \tau_\rho}^n \rangle_{\mathcal{E}}^2 - \langle \Phi, Z_{t_1 \wedge \tau_\rho}^n \rangle_{\mathcal{E}}^2 - \int_{t_1 \wedge \tau_\rho}^{t_2 \wedge \tau_\rho} \alpha_n \langle \Phi, G^n(Y_s^n, \theta_s^n) \Phi \rangle_{\mathcal{E}} ds \right) \right] \\ & \quad + \mathbb{E}^n \left[\mathbb{I}_{A^n} \left(\int_{t_1 \wedge \tau_\rho}^{t_2 \wedge \tau_\rho} \alpha_n \langle \Phi, G^n(Y_s^n, \theta_s^n) \Phi \rangle_{\mathcal{E}} - \langle \Phi, G^n(u(s), p(s)) \Phi \rangle_{\mathcal{E}} ds \right) \right]. \end{aligned}$$

Here the first expectation in the final right hand side vanishes due to the Itô-formula (6.17): We apply the Itô-formula for the function g_ρ and the martingales $Z_{t \wedge \tau_\rho}^n$ to the terms $\langle \Phi, Z_{t_2 \wedge \tau_\rho}^n \rangle_{\mathcal{E}}^2$ and $\langle \Phi, Z_{t_1 \wedge \tau_\rho}^n \rangle_{\mathcal{E}}^2$. Then we find – similarly to part (a) – that the summands in the right hand side of the Itô-formula vanish. Therefore we are left with only the martingale M^{n, g_ρ} and the integral term, wherein $g_\rho''(\langle \phi, Z_{t-}^n \rangle_{\mathcal{E}}) = 2$ for all $t < \tau_\rho^n$. The martingale term vanishes due to the martingale property and the remaining integral is cancelled by the integral in the above expectation. Overall this shows that the first expectation vanishes.

Next we take the absolute value on both sides of the above equality and obtain, estimating the second expectation and extending the integration interval to $[0, T]$, the inequality

$$\begin{aligned} & \left| \mathbb{E}^n \left[\mathbb{I}_A \left(\langle \Phi, \zeta_{t_2 \wedge \tau_\rho} \rangle_{\mathcal{E}}^2 - \int_0^{t_2 \wedge \tau_\rho} \langle \Phi, G(u(s), p(s)) \Phi \rangle_{\mathcal{E}} ds - \langle \Phi, \zeta_{t_1 \wedge \tau_\rho} \rangle_{\mathcal{E}}^2 \right. \right. \right. \\ & \qquad \qquad \qquad \left. \left. \left. + \int_0^{t_1 \wedge \tau_\rho} \langle \Phi, G(u(s), p(s)) \Phi \rangle_{\mathcal{E}} ds \right) \right] \right| \\ & \leq \int_0^T \mathbb{E}^n \left| \alpha_n \langle \Phi, G^n(Y_s^n, \theta_s^n) \Phi \rangle_{\mathcal{E}} - \langle \Phi, G(u(s), p(s)) \Phi \rangle_{\mathcal{E}} \right| ds. \end{aligned}$$

The convergence of the upper bound to zero for $n \rightarrow \infty$ follows by assumption (5.6). Hence we have proven an analogous result to (6.19) in part (a). The same line of argument that concluded part (a) also concludes part (b). The proof is completed.

7 Application to models of excitable membranes

The primary motivation for the present work stems from the study of stochastic version of the Hodgkin-Huxley model [23] describing action potential generation and propagation in spatially extended neurons in a PDMP formulation. This model is analogous in structure to hybrid models that are used for the modelling of Calcium dynamics,

cf. [20, 26, 44], or models of cardiac tissue, cf. [21]. Therefore, we consider as an example of the application of the presented limit theorems a general *compartmental-type hybrid stochastic model* for spatially extended excitable membranes introduced in [42, Sec. 3.2] which subsumes the above mentioned applications. (Another example for the application of Theorem 4.1 is the law of large numbers that is presented in [5] for a particular one-dimensional hybrid model.) We refrain from discussing the physiological derivations of this type of model and the implications and interpretations of the limit theorems in this setting. These aspects will be subject to a forthcoming publication. We now fix some notation for the remainder of the section. The set $D \subset \mathbb{R}^d$ denotes bounded spatial domain with the physically reasonable dimensions $d \leq 3$. That is, the set D is a bounded interval when $d = 1$ and when $d \in \{2, 3\}$ we assume it possesses a C^3 -boundary. Further, for a given dimension d , let s denote the smallest integer such that $s > d/2$. Finally, let $m \in \mathbb{N}$ be the fixed number of states ion channels can be in.

7.1 Deterministic limit system

The deterministic limit is the solution to the membrane equation

$$\dot{u} = \sum_{i,j=1}^d a_{ij}(x)u_{x_i x_j} + \sum_{i=1}^m g_i(x)p_i(E_i - u) \tag{7.1}$$

with $p_j, j = 1 \dots, m$ given by solutions of the coupled equations

$$\dot{p}_j = F_j(p, u) := \sum_{i \neq j} q_{ij}(u)p_i - q_{ij}(u)p_j. \tag{7.2}$$

We choose Dirichlet boundary conditions for the component u , i.e., $u(t, x) = 0$ for all $x \in \partial D$ and all $t \in [0, T]$, which, however, is of no particular importance for the considerations that follow and can be readily changed. Here the coefficient functions a_{ij} and g_i are smooth on \bar{D} , with g_i non-negative, and the differential operator is strongly elliptic. Further, the rate functions q_{ij} are sufficiently smooth.⁹ Finally, the initial conditions satisfy $u_0 \in H_0^1(D) \cap H^s(D)$ and $p_i(0) \in H^s(D)$ and, in addition, the pointwise bounds $u(0, x) \in [\bar{u}_-, \bar{u}_+]$ and $p_i(0, x) \in [0, 1], i = 1, \dots, m$, such that $\sum_{i=1}^m p_i(0, x) = 1$, hold for all $x \in D$. Then, the deterministic system (7.1), (7.2) is well-posed, that is, there exists a unique global solution depending continuously on the initial condition, which also satisfies (4.2) [42, Sec. 3.3.1]. In particular, the solution $(u(t), p(t))_{t \in [0, T]}$ is in $C([0, T], H^s(D))$ componentwise for every $T > 0$ and pointwise bounded, i.e., $u(t, x) \in [\bar{u}_-, \bar{u}_+]$ and $p_i(t, x) \in [0, 1]$ for all $(t, x) \in [0, T] \times \bar{D}$ and all $i = 1, \dots, m$.

7.2 Compartmental-type membrane models

We briefly summarise the essential features of PDMPs $(U_t^n, \Theta_t^n)_{t \geq 0}, n \in \mathbb{N}$, constituting compartmental-type membrane models.

Firstly, an integral component of the sequence of models is a sequence of compartmentalisation of the spatial domain D . Thus, for each $n \in \mathbb{N}$ let \mathcal{P}_n be a convex partition of the domain D , i.e., \mathcal{P}_n is a finite collection of mutually disjoint convex¹⁰ subsets of D , called *compartments*, such that their union equals D .

The second fundamental aspect is the channel distribution across the compartments

⁹In detail the conditions are [42, Sec. 3.3.1]: The functions q_{ij} are bounded and bounded away from zero on the interval $[\bar{u}_-, \bar{u}_+]$. Further, on this interval they satisfy a Lipschitz and polynomial growth condition and are twice continuously differentiable with bounded derivatives.

¹⁰The convexity of the compartments is a technical assumption which allows to employ Poincaré's inequality in the proof of the limit theorems with a known optimal Poincaré constant [1, 38].

yielding the stochastic jump dynamics and the coordinate functions z^n . We assume that each compartment either does not contain channels or a fixed deterministic number. Let $p(n)$ denote the number of non-empty compartments of the n th model denoted by $D_{1,n}, \dots, D_{p(n),n}$ and $l(k, n)$ be for $k \leq p(n)$ the total number of channels in the k th non-empty compartment of the n th model. Then the piecewise constant components of the PDMPs are given by $mp(n)$ -dimensional vectors $\Theta_t^n = (\Theta_i^{k,n}(t))_{i=1, \dots, m, k=1, \dots, p(n)}$ with finite state spaces K_n . Each component $\Theta_i^{k,n}(t)$ counts the number of channels located in the domain $D_{k,n}$ which are in state i at time t . and it holds that

$$\sum_{i=1}^m \Theta_i^{k,n}(t) = l(k, n).$$

as channels can neither be destroyed nor created. We proceed to define the stochastic jump dynamics. As two channel switching do not occur simultaneously, the only jumps in the configuration $\theta^n \in K_n$ with non-zero probability are transitions concerning one single channel. That is, events for which in one particular compartment one particular channel changes its state. Let $q_{ij} : \mathbb{R} \rightarrow \mathbb{R}_+$ denote the u -dependent instantaneous rate of one channel switching from state i to j . Then given a specific configuration $\theta^n \in K_n$ the rate that one channel in compartment $D_{k,n}$ switches from state i to j is given by

$$\theta_i^{k,n} Q_{ij}^{k,n}(u) \in \mathbb{R}_+, \tag{7.3}$$

where $Q_{ij}^n(u)$ is a functional of the membrane variable $u \in L^2(D)$ defined as

$$Q_{ij}^{k,n}(u) := q_{ij} \left(\frac{1}{|D_{k,n}|} \int_{D_{k,n}} u(x) dx \right).$$

That is, $Q_{ij}^{k,n}(u)$ is the instantaneous rate q_{ij} evaluated at the average value of the membrane variable over the compartment $D_{k,n}$. Hence the rate (7.3) is the number of channels in state i in domain $D_{k,n}$ times the rate of one channel switching from i to j . This definition yields by summing over all events the total instantaneous rate

$$\Lambda^n(u, \theta^n) := \sum_{i,j=1}^m \sum_{k=1}^{p(n)} \theta_i^{k,n} Q_{ij}^{k,n}(u). \tag{7.4}$$

Note that for each n the total instantaneous rate is bounded and as expected proportional to the total number of channels which implies that the PDMPs are regular. Finally, we define on the set K_n for $i = 1, \dots, m$ the coordinate functions

$$z_i^n(\theta^n) := \sum_{k=1}^{p(n)} \frac{\theta_i^{k,n}}{l(k, n)} \mathbb{I}_{D_{k,n}} \in L^2(D). \tag{7.5}$$

The coordinate process $z^n(\Theta_t^n)$ is càdlàg with each component taking values in $L^2(D)$. Clearly, the coordinate process is zero on those compartments which do not contain channels. Moreover, each $z_i^n(\Theta_t^n)$ is for every $t \geq 0$ a piecewise constant function on the spatial domain D which takes values in $[0, 1]$.

Thirdly, the family of abstract evolution equations (2.1) defining the dynamics of the PDMP's continuous component U^n are given by the parabolic, linear, inhomogeneous

second order partial differential equations

$$\dot{u} = \sum_{i,j=1}^d a_{ij}(x)u_{x_i x_j} + \sum_{i=1}^m g_i(x) z_i^n(\theta^n) (E_i - u). \tag{7.6}$$

Consistently with the deterministic limit system we equip equation (7.6) with Dirichlet boundary conditions. Finally, we define the operators A, B depending on θ^n only via suitable coordinate functions, cf. (2.5), by

$$A(z^n(\theta^n)) u := \sum_{i,j=1}^d a_{ij}(x)u_{x_i x_j}, \quad B(z^n(\theta^n), u) := \sum_{i=1}^m g_i(x) z_i^n(\theta^n) (E_i - u). \tag{7.7}$$

To conclude, it is easy to see that the characteristics defined via the individual rates (7.3), the total jump rate (7.4) and the evolution equation (7.7) define a sequence of $L^2(D) \times K_n$ -valued infinite-dimensional PDMPs $(U_t^n, \Theta_t^n)_{t \geq 0}$. Moreover, the membrane component $(U_t^n)_{t \geq 0}$ is almost everywhere pointwise bounded, i.e., $U_t^n(x) \in [\bar{u}_-, \bar{u}_+]$ for almost all $x \in D$ and all $t \geq 0$, where $\bar{u}_- := \min_i E_i \leq 0$ and $\bar{u}_+ := \max_i E_i \geq 0$, for initial conditions U_0^n satisfying these bounds, cf. [42, Sec. 3.2], which we always assume.

7.3 Limit theorems for compartmental-type models

Applying the limit theorems derived in Sections 4 and 5 to compartmental models we find that the conditions therein translate into assumptions on the behaviour of the sequence of partitions \mathcal{P}_n and the number of ion channels in the membrane, see Appendix B. Thus, we denote by $\delta(n)$ the maximal diameter of the non-empty compartments in the n th model, i.e.,

$$\delta_+(n) := \max_{k=1, \dots, p(n)} \text{diam}(D_{k,n}),$$

and by $\ell_+(n)$ and ℓ_- the maximal and minimal number of channels in non-empty compartments, i.e.,

$$\ell_+(n) := \max_{k=1, \dots, p(n)} l(k, n), \quad \ell_-(n) := \min_{k=1, \dots, p(n)} l(k, n).$$

Then the law of large numbers takes the following form.

Theorem 7.1. *Assume that the sequence of partitions satisfies that*

$$\lim_{n \rightarrow \infty} \delta_+(n) = 0, \quad \lim_{n \rightarrow \infty} \ell_-(n) = \infty, \tag{7.8}$$

and that the initial conditions $(U_0^n, z^n(\Theta_0^n))$ converge in probability to (u_0, p_0) in the space $L^2(D)^{m+1}$. Then the compartmental-type models converge in probability to the deterministic solution of the excitable media system (7.1), (7.2) in the sense that it holds for all $\epsilon > 0$ that

$$\lim_{n \rightarrow \infty} \mathbb{P} \left[\sup_{t \in [0, T]} \|U_t^n - u(t)\|_{L^2} + \sum_{i=1}^m \sup_{t \in [0, T]} \|z_i^n(\Theta_t^n) - p(t)\|_{L^2} > \epsilon \right] = 0. \tag{7.9}$$

Moreover, the convergence also holds in the mean in the space $L^2((0, T), L^2(D))$, i.e.,

$$\lim_{n \rightarrow \infty} \mathbb{E}^n \left[\|U_t^n - u(t)\|_{L^2((0, T), L^2)} + \sum_{i=1}^m \|z_i^n(\Theta_t^n) - p(t)\|_{L^2((0, T), L^2)} \right] = 0. \tag{7.10}$$

Next we present the appropriate quadratic variation process for the martingale cen-

tral limit theorem. For the definition of the limiting diffusion we consider for $u, p_i \in C(\overline{D})$ the bilinear form

$$\begin{aligned}
 (\Psi, \Phi) \mapsto (G(u, p) \Psi, \Phi)_{L^2} &= \sum_{j=1}^m \sum_{i \neq j} \int_D p_i(x) q_{ij}(u(x)) \psi_j(x) \phi_j(x) dx \\
 &+ \sum_{j=1}^m \sum_{i \neq j} \int_D p_j(x) q_{ji}(u(x)) \psi_j(x) \phi_j(x) dx \\
 &- \sum_{j=1}^m \sum_{i \neq j} \int_D p_j(x) q_{ji}(u(x)) \psi_i(x) \phi_j(x) dx \\
 &- \sum_{j=1}^m \sum_{i \neq j} \int_D p_i(x) q_{ij}(u(x)) \psi_i(x) \phi_j(x) dx.
 \end{aligned} \tag{7.11}$$

Note that the right hand side is finite for all $\phi_i, \psi_i \in L^2(D)$ as p_i and $q_{ij}(u)$ are bounded functions. Hence, for every given $\Psi \in L^2(D)^m$ the mapping $\Phi \mapsto (G(u, p)\Psi, \Phi)$ is a linear, bounded functional on $L^2(D)^m$ and, conversely, for every given $\Phi \in L^2(D)^m$ the mapping $\Psi \mapsto (G(u, p)\Psi, \Phi)$ is a linear, bounded functional on $L^2(D)^m$.

Proposition 7.1. *The operator $G(u, p)$ defined via (7.11) is for $s > d/2$ a trace class operator mapping $H^s(D)$ into its dual $H^{-s}(D)$. Moreover, the operator-valued map $t \mapsto G(u(t), p(t))$ defines a unique centred diffusion process on $H^{-s}(D)$.*

Proof. As stated in [27] it is sufficient for the statement of the proposition that the operator $G(u(t), p(t))$ is self-adjoint, positive and of trace class. These properties are easily verified and for a detailed proof we refer to [42]. □

In order to state the conditions on the partitions in the central limit theorem we define $\nu_+(n)$ and $\nu_-(n)$ to be the maximum and minimum Lebesgue measure of non-empty compartments, i.e.,

$$\nu_+(n) := \max_{k=1, \dots, p(n)} |D_{k,n}|, \quad \nu_-(n) := \min_{k=1, \dots, p(n)} |D_{k,n}|.$$

Finally, note that in the following the coordinate functions z^n are considered as maps from K_n into the space $H^{-2s}(D)$.

Theorem 7.2. *Let s be the smallest integer such that $s > d/2$. If in addition to (7.8) and the convergence of the initial conditions the sequence of partitions satisfies*

$$\lim_{n \rightarrow \infty} \frac{\ell_-(n) \nu_-(n)}{\ell_+(n) \nu_+(n)} = 1, \tag{7.12}$$

then the sequence of $H^{-2s}(D)$ -valued martingales $\left(\sqrt{\frac{\ell_-(n)}{\nu_+(n)}} M_t^n\right)_{t \geq 0}$ converges weakly to the $(H(D)^{-2s})^m$ -valued diffusion defined by (7.11).

Remark 7.1. We note that for all reasonable physical domains D and all initial conditions (u_0, p_0) sequences of partitions \mathcal{P}_n and initial conditions (U_0^n, Θ_0^n) for the PDMPs can be found satisfying the conditions of Theorems 7.1 and 7.2. For example, a suitable sequence of partitions is obtained by grids of uniform cubes with decreasing edge length covering the domain D and putting channels only into these cubes which are fully contained in D . For a more detailed discussion of these aspects we refer to the PhD thesis of one of the present authors [42].

8 Conclusions

As a general theoretical results for PDMPs we have derived a law of large numbers and martingale central limit theorem in Sections 4 and 5 of this study. The former establishes a connection of stochastic hybrid models to deterministic models given, e.g., by systems of partial differential equations. Whereas the latter connects the stochastic fluctuations in the hybrid models to diffusion processes. As a prerequisite to these limit theorems we carried out a thorough discussion of Hilbert space valued martingales associated to the PDMPs. Furthermore, these limit theorems provide the basis for a general Langevin approximation to PDMPs, i.e., certain stochastic partial differential equations that are expected to be similar in their dynamics to PDMPs. We have applied these results to compartmental-type models of spatially extended excitable membranes. Ultimately this yields a system of SPDEs which models the internal noise of a biological excitable membrane based on a theoretical derivation from exact stochastic hybrid models.

Topics for further research are motivated by corresponding results in finite-dimensions [32, 37] and for spatially inhomogeneous chemical reaction systems converging to reaction diffusion equations, cf. [27]. In these studies limit theorems are derived for the fluctuations around the deterministic limit identified by the law of large numbers. Using the notation of Section 5 we conjecture that the sequence of processes, $(\sqrt{\alpha_n}(U_t^n - u(t), z^n(\Theta_t^n) - p(t)))_{t \geq 0}$, $n \in \mathbb{N}$, converges in distribution to a suitable diffusion process. Moreover, we further conjecture that this limit is closely related to the asymptotic linearisation of the Langevin approximation around the solution of the deterministic limit, cf. [37] wherein this result is proven for finite-dimensional PDMPs.

Further, on the applications side we believe that the Langevin approximation to spatio-temporal PDMP models of excitable membranes poses an important object for further investigation. Its derivation was the initial motivation of the study of the limit theorems in the present study and it is their main application herein which enables to write down the system of SPDEs that constitute a Langevin approximation. This system now demands for further analysis, particularly, first of all the question of existence and uniqueness of the Langevin approximation has to be addressed. Subsequently, as SPDEs are analytically more accessible than PDMPs a theoretical analysis of qualitative and quantitative properties of the models may be possible.

Finally, we want to mention that the limit theorems presented also find applications beyond excitable membrane models. In current work in progress by one of the present authors the limit theorems derived in Sections 4 and 5 are applied to stochastic neural field equations, based on a model presented in [12], cf. a preliminary account in [41]. We also plan to investigate the connection to similar limits derived for reaction-diffusion models, cf. the series of results on variations of the model in [27, 28, 29, 30] and [7, 8, 9, 10, 11]. An answer to this question would contribute to a more complete picture of limit-theorems for spatio-temporal stochastic models.

Acknowledgments. During the time the presented work was accomplished M. Riedler was a PhD student at Heriot-Watt University supported by the EPSRC grant EP/E03635X/1. M. Riedler further acknowledges support from a joint UK Mathematical Neuroscience Network (MNN) and the Cell Signalling Network (SIGNET) travel grant. Finally, the authors thank one anonymous reviewer for a fast reviewing of our work.

A Proof of Theorem 3.1 (Itô-isometry)

In this proof we show that under condition (3.4) the processes M_j^n , $j = 1, \dots, m$, $n \in \mathbb{N}$, defined in (3.1) are square-integrable, càdlàg martingales which satisfy the Itô-

isometry (3.5). Throughout the proof we fix a $j = 1, \dots, m$ and $n \in \mathbb{N}$ and the results holds for any such j and n . Therefore, speaking of a PDMP in the following always refers to the PDMP $(U_t^n, \Theta_t^n)_{t \geq 0}$ corresponding to the fixed n . Further, for notational simplicity we omit the indices n and j discriminating processes and characteristics of PDMPs, i.e., M_j^n and $z_j^n(\Theta^n)$ are denoted simply by M and $z(\Theta)$. Finally, recall that τ_k , $k = 1, 2, \dots$, denotes the sequence of increasing random jump times of the PDMP which are stopping times satisfying $\lim_{k \rightarrow \infty} \tau_k = \infty$ almost surely.

First of all, note that the process M is càdlàg by definition. The proof of the remaining open results is split into three parts. In the first, part (a), we prove the martingale property for the real process $(\langle \phi, M(t) \rangle_E)_{t \geq 0}$ for every $\phi \in E^*$. Then, the first main statement of Theorem 3.1, the square-integrability of the process $M(t)$, is proved in part (b). Moreover, as square-integrability implies integrability, the Hilbert space martingale property follows. Finally, the second main statement, the Itô-Isometry (3.5), is established in part (c). The proof we present in part (b) is motivated by the proof of [24, Prop. 4.5.3] which states the corresponding results for real-valued martingales associated with PDMPs. In extending to the present setup the method of proof employed therein one has to ensure, on the one hand, that the employed results and estimation procedures all have corresponding analoga in the infinite-dimensional setting. On the other hand, one has to carefully make sure that only the weaker regularity results available in infinite-dimensions are used. Finally, the introduction of random initial conditions, not considered in [24], also necessitates some adaptations.

(a) First note that for all $\phi \in E^*$ the real-valued processes $\langle \phi, M(t) \rangle_E$ satisfy

$$\begin{aligned} \langle \phi, M(t) \rangle_E &= \langle \phi, z(\Theta_t) \rangle_E - \langle \phi, z(\Theta_0) \rangle_E \\ &\quad - \int_0^t \Lambda(U_{s-}, \Theta_{s-}) \int_K \langle \phi, z(\xi) \rangle_E - \langle \phi, z(\Theta_{s-}) \rangle_E \mu((U_{s-}, \Theta_{s-}), d\xi) ds. \end{aligned} \tag{A.1}$$

Equation (A.1) is obtained from (3.3) due to the regularity of the PDMP as the set of jump times in $[0, t]$ is almost surely finite for all $t \geq 0$. Therefore the integrands in the right hand sides of (3.3) and (A.1) differ only on a set of Lebesgue measure zero almost surely. Moreover, the integrand in the right hand side of (A.1) has the form of the extended generator, cf. Theorem 2.1, applied to the map

$$(u, \xi) \mapsto \langle \phi, z(\xi) \rangle_E, \tag{A.2}$$

which is independent of u . It follows that the process $\langle \phi, M(t) \rangle_E$ is a local martingale if the map (A.2) is in the domain of the extended generator, cf. Theorem 2.1. Obviously, path-differentiability almost everywhere is trivially satisfied as the map $t \mapsto \langle \phi, z(\Theta_t) \rangle_E$ is piecewise constant. Hence, it remains to consider the integrability condition for which it is a sufficient that

$$\mathbb{E} \int_0^t \Lambda(U_{s-}, \Theta_{s-}) \int_K |\langle \phi, z(\xi) - z(\Theta_{s-}) \rangle_E| \mu((U_{s-}, \Theta_{s-}), d\xi) ds < \infty \quad \forall t \geq 0, \tag{A.3}$$

cf. [13, 17, 24]. Using Young's inequality we obtain an upper bound to (A.3) by

$$\frac{1}{2} \mathbb{E} \int_0^t \Lambda(U_{s-}, \Theta_{s-}) ds + \frac{1}{2} \mathbb{E} \int_0^t \Lambda(U_{s-}, \Theta_{s-}) \int_K |\langle \phi, z(\xi) - z(\Theta_{s-}) \rangle_E|^2 \mu((U_{s-}, \Theta_{s-}), d\xi) ds.$$

Here the first expectation is finite due to the PDMP being regular and the second is finite by an immediate consequence of assumption (3.4).

Next, we show that the process is not only a local martingale but even a martingale.

As mentioned above the process $\langle \phi, M(t) \rangle_E$ satisfies

$$\langle \phi, M(t) \rangle_E = \int_0^t \int_K \langle \phi, z(\xi) - z(\Theta_{s-}) \rangle_E \widetilde{M}(ds, d\xi)$$

where $\widetilde{M} := N - \widehat{N}$ is the random martingale measure associated with the PDMP with counting measure N and compensator $\widehat{N}(d\xi, ds) = \Lambda(U_{s-}, \Theta_{s-}) \mu((U_{s-}, \Theta_{s-}), d\xi) ds$. The validity of this formula follows as (A.2) is in the extended generator. Thus the process $\langle \phi, M(t) \rangle_E$ has the form of a stochastic integral with respect to the martingale measure associated with the PDMP. Furthermore, due to [24, Thm. 4.6.1] it holds that the process is a martingale if (A.3) is finite for all $t \geq 0$. But we have already shown that this holds due to the regularity of the PDMP and assumption (3.4).

(b) We now prove the square-integrability of the process M . In a first step we prove in (b.1) that M stopped at the first jump τ_i is square-integrable. Subsequently in part (b.2) this result is extended to M stopped at any jump time τ_k , $k \in \mathbb{N}$. Then we are able to infer square-integrability of the process M . As square-integrability implies integrability it follows from part (a) that M is a Hilbert space valued martingale.

(b.1) Note that prior to τ_1 the jump component Θ of the PDMP remains constant. We introduce the notation

$$\widetilde{N}(s) := \int_0^s \Lambda(U_r, \Theta_0) \int_K z(\xi) - z(\theta_0) \mu((U_r, \Theta_0), d\xi) dr$$

which implies that $s \mapsto \|\widetilde{N}(s)\|_E^2$ is almost surely absolutely continuous with derivative

$$\begin{aligned} \frac{d}{ds} \|\widetilde{N}(s)\|_E^2 &= 2 \left(\frac{d}{ds} \widetilde{N}(s), \widetilde{N}(s) \right)_E \\ &= 2 \Lambda(U_s, \Theta_0) \int_K (z(\xi) - z(\Theta_0), \widetilde{N}(s))_E \mu((U_s, \Theta_s), d\xi). \end{aligned} \quad (\text{A.4})$$

Due to the structure of a PDMP we obtain for the conditional expectation with respect to the initial condition

$$\begin{aligned} \mathbb{E}[\|M(\tau_1 \wedge t)\|_E^2 | \mathcal{F}_0] &= \|\widetilde{N}(t)\|_E^2 \exp\left(-\int_0^t \Lambda(U_r, \Theta_0) dr\right) \\ &+ \int_0^t \left[\int_K \|z(\vartheta) - z(\Theta_0) - \widetilde{N}(s)\|_E^2 \mu((U_s, \Theta_0), d\vartheta) \right] \Lambda(U_s, \Theta_0) \exp\left(-\int_0^s \Lambda(U_r, \Theta_0) dr\right) ds. \end{aligned}$$

That is, the first term in the right hand side is the position of the stopped process $\|M(\tau_1 \wedge t)\|_E^2$ at time t if $t < \tau_1$ times the conditional probability that the first jump does not occur before t . The second term is its position after the jump integrated over the conditional density that a jump occurs in $[0, t]$. We apply integration by parts to the first term (note that $\widetilde{N}(0) = 0$) and find that

$$\begin{aligned} \|\widetilde{N}(t)\|_E^2 \exp\left(-\int_0^t \Lambda(U_s, \Theta_0) ds\right) &= \int_0^t \left[2 \left(\frac{d}{ds} \widetilde{N}(s), \widetilde{N}(s) \right)_E \exp\left(-\int_0^s \Lambda(U_r, \Theta_0) dr\right) \right. \\ &\quad \left. - \|\widetilde{N}(s)\|_E^2 \Lambda(U_s, \Theta_0) \exp\left(-\int_0^s \Lambda(U_r, \Theta_0) dr\right) \right] ds. \end{aligned}$$

Therefore we obtain

$$\begin{aligned} \mathbb{E}[\|M(\tau_1 \wedge t)\|_E^2 | \mathcal{F}_0] &= \int_0^t \left[2 \left(\frac{d}{dt} \tilde{N}(s), \tilde{N}(s) \right)_E \exp\left(-\int_0^s \Lambda(U_r, \Theta_0) dr\right) \right] ds \\ &\quad + \int_0^t \left[\left(\int_K \|z(\vartheta) - z(\Theta_0) - \tilde{N}(s)\|_E^2 - \|\tilde{N}(s)\|_E^2 \mu((U_s, \Theta_0), d\vartheta) \right) \right. \\ &\quad \left. \Lambda(U_s, \Theta_0) \exp\left(-\int_0^s \Lambda(U_r, \Theta_0) dr\right) \right] ds. \end{aligned}$$

Note that $\|z(\vartheta) - z(\Theta_0) - \tilde{N}(s)\|_E^2 = \|z(\vartheta) - z(\Theta_0)\|_E^2 + \|\tilde{N}(s)\|_E^2 - 2(z(\vartheta) - z(\Theta_0), \tilde{N}(s))_E$ and thus

$$\begin{aligned} \mathbb{E}[\|M(\tau_1 \wedge t)\|_E^2 | \mathcal{F}_0] &= \\ &\int_0^t \left[2 \left(\frac{d}{dt} \tilde{N}(s), \tilde{N}(s) \right)_E \exp\left(-\int_0^s \Lambda(U_r, \Theta_0) dr\right) \right] ds \\ &- 2 \int_0^t \left[\left(\int_K (z(\vartheta) - z(\Theta_0), \tilde{N}(s))_E \mu((U_s, \Theta_0), d\vartheta) \right) \Lambda(U_s, \Theta_0) \exp\left(-\int_0^s \Lambda(U_r, \Theta_0) dr\right) \right] ds \\ &+ \int_0^t \left[\left(\int_K \|z(\vartheta) - z(\Theta_0)\|_E^2 \mu((U_s, \Theta_0), d\vartheta) \right) \Lambda(U_s, \Theta_0) \exp\left(-\int_0^s \Lambda(U_r, \Theta_0) dr\right) \right] ds. \end{aligned}$$

Due to form of the derivative (A.4) the first two terms cancel and we are left with the equality

$$\begin{aligned} \mathbb{E}[\|M(\tau_1 \wedge t)\|_E^2 | \mathcal{F}_0] & \tag{A.5} \\ &= \int_0^t \Lambda(U_s, \Theta_0) \int_K \|z(\vartheta) - z(\Theta_0)\|_E^2 \mu((U_s, \Theta_0), d\vartheta) \exp\left(-\int_0^s \Lambda(U_r, \Theta_0) dr\right) ds. \end{aligned}$$

Next we calculate the expectation of the real-valued process

$$\tilde{N}_2(s) := \int_0^s \Lambda(U_r, \Theta_0) \int_K \|z(\vartheta) - z(\Theta_0)\|_E^2 \mu((U_r, \Theta_0), d\vartheta) dr$$

stopped at τ_1 . The process \tilde{N}_2 is connected to the process \tilde{N} defined at the beginning of part (b.1) inasmuch as the integrand of the former is the squared norm of the latter. Furthermore note that \tilde{N}_2 is the term inside the expectation in the right hand side of the Itô-isometry (3.5). Thus the aim is now to show that the conditional expectation of $\tilde{N}_2(t \wedge \tau_1)$ equals the conditional expectation of $\|M(t \wedge \tau_1)\|_E^2$. Again due to the particular structure of the PDMP we obtain for the conditional expectation

$$\begin{aligned} \mathbb{E}[\tilde{N}_2(\tau_1 \wedge t) | \mathcal{F}_0] &= \tilde{N}_2(s) \exp\left(-\int_0^t \Lambda(U_r, \Theta_0) dr\right) + \int_0^t \left[\tilde{N}_2(s) \Lambda(U_s, \Theta_0) \exp\left(-\int_0^s \Lambda(U_r, \Theta_0) dr\right) \right] ds. \end{aligned}$$

Integration by parts applied to the integral term yields

$$\begin{aligned} \int_0^t \left[\tilde{N}_2(s) \Lambda(U_s, \Theta_0) \exp\left(-\int_0^s \Lambda(U_r, \Theta_0) dr\right) \right] ds &= -\tilde{N}_2(t) \exp\left(-\int_0^t \Lambda(U_r, \Theta_0) dr\right) \\ &+ \int_0^t \left[\Lambda(U_s, \Theta_0) \int_K \|z(\vartheta) - z(\Theta_0)\|_E^2 \mu((U_s, \Theta_0), d\vartheta) \exp\left(-\int_0^s \Lambda(U_r, \Theta_0) dr\right) \right] ds. \end{aligned}$$

Therefore we obtain that

$$\begin{aligned} & \mathbb{E}[\tilde{N}_2(\tau_1 \wedge t) | \mathcal{F}_0] & (A.6) \\ &= \int_0^t \Lambda(U_s, \Theta_0) \int_K \|z(\vartheta) - z(\Theta_0)\|_E^2 \mu((U_s, \Theta_0), d\vartheta) \exp\left(-\int_0^s \Lambda(U_r, \Theta_0) dr\right) ds. \end{aligned}$$

A comparison of the right hand sides in equalities (A.5) and (A.6) shows that they are equal and thus we obtain after taking the expectation of both conditional expectations that

$$\mathbb{E}\|M(\tau_1 \wedge t)\|_E^2 = \mathbb{E}\tilde{N}_2(\tau_1 \wedge t). \tag{A.7}$$

As \tilde{N}_2 is increasing and thus $\tilde{N}_2(\tau_1 \wedge t) \leq \tilde{N}_2(t)$ almost surely, we obtain that the right hand side in this equation is finite due to condition (3.4). Note that (A.7) is the Itô-isometry (3.5) for the stopped process $M(t \wedge \tau_1)$.

(b.2) In this part of the proof we show the square-integrability for the process M stopped at an arbitrary jump time τ_k , $k \in \mathbb{N}$, and finally for the non-stopped process M . To this end we first note that Analogously to part (b.1) we find that

$$\mathbb{E}\left[\|M(\tau_{k+1} \wedge t) - M(\tau_k \wedge t)\|_E^2 | \mathcal{F}_{\tau_k}\right] = \mathbb{E}\left[\tilde{N}_2(\tau_{k+1} \wedge t) - \tilde{N}_2(\tau_k \wedge t) | \mathcal{F}_{\tau_k}\right].$$

Thus taking expectations on both sides of this equality yields

$$\mathbb{E}\|M(\tau_{k+1} \wedge t) - M(\tau_k \wedge t)\|_E^2 = \mathbb{E}\tilde{N}_2(\tau_{k+1} \wedge t) - \mathbb{E}\tilde{N}_2(\tau_k \wedge t) < \infty, \tag{A.8}$$

where the right hand side is finite as due to(3.4) both expectations are finite.

By induction we next show that each $M(\tau_k \wedge t)$ is square-integrable. Assume that $\mathbb{E}\|M(\tau_k \wedge t)\|_E^2 < \infty$, where the induction basis for $k = 1$ holds due to part (b.1). Then the reverse triangle inequality yields that

$$\begin{aligned} & \mathbb{E}\|M(\tau_{k+1} \wedge t)\|_E^2 + \mathbb{E}\|M(\tau_k \wedge t)\|_E^2 - 2\mathbb{E}(\|M(\tau_{k+1} \wedge t)\|_E \|M(\tau_k \wedge t)\|_E) \\ & \leq \mathbb{E}\|M(\tau_{k+1} \wedge t) - M(\tau_k \wedge t)\|_E^2. \end{aligned}$$

Here the right hand side is finite due to (A.8) and an application of Young’s inequality to the product in the left hand side yields that for all $\epsilon > 0$

$$(1 - 2\epsilon) \mathbb{E}\|M(\tau_{k+1} \wedge t)\|_E^2 + \left(1 - \frac{1}{2\epsilon}\right) \mathbb{E}\|M(\tau_k \wedge t)\|_E^2 < \infty$$

Assume that $\mathbb{E}\|M(\tau_{k+1} \wedge t)\|_E^2 = \infty$. Then choosing $\epsilon < 1/2$ we obtain a contradiction due to the induction hypotheses.

In a final step of this part of the proof we show square-integrability for the non-stopped process. Using Fatou’s Lemma and monotone convergence for interchanging limits and expectation we obtain the following upper estimate

$$\begin{aligned} \mathbb{E}\|M(t)\|_E^2 &= \mathbb{E} \liminf_{k \rightarrow \infty} \|M(\tau_k \wedge t)\|_E^2 \\ &\leq \liminf_{k \rightarrow \infty} \mathbb{E}\|M(\tau_k \wedge t)\|_E^2 = \lim_{k \rightarrow \infty} \mathbb{E}\tilde{N}_2(\tau_k \wedge t) = \mathbb{E}\tilde{N}_2(t), \end{aligned} \tag{A.9}$$

where the final term is finite due to condition (3.4). Moreover, as square-integrability implies integrability, the martingale property for the Hilbert space valued process M now follows due to part (a).

(c) Finally, in the last part of the proof we establish the Itô-isometry. To this end we first show that equality (A.7) holds for all $\tau_k \wedge t$, $k \in \mathbb{N}$. Again we proceed by induction

with the induction basis given by (A.7). We observe that

$$\begin{aligned} \|M(\tau_{k+1} \wedge t) - M(\tau_k \wedge t)\|_E^2 &= \|M(\tau_{k+1} \wedge t)\|_E^2 - \|M(\tau_k \wedge t)\|_E^2 \\ &\quad - 2(M(\tau_k \wedge t), M(\tau_{k+1} \wedge t) - M(\tau_k \wedge t))_E. \end{aligned} \quad (\text{A.10})$$

Taking the conditional expectation with respect to the stopped σ -field $\mathcal{F}_{\tau_k \wedge t}$ we find that the second term in the right hand side of (A.10) vanishes as it holds

$$\begin{aligned} \mathbb{E}[(M(\tau_k \wedge t), M(\tau_{k+1} \wedge t) - M(\tau_k \wedge t))_E | \mathcal{F}_{\tau_k \wedge t}] &= \\ &= (M(\tau_k \wedge t), \mathbb{E}[M(\tau_{k+1} \wedge t) - M(\tau_k \wedge t) | \mathcal{F}_{\tau_k \wedge t}])_E = 0 \end{aligned}$$

due to the following properties of the conditional expectation: Firstly, for E -valued random variables X, Y such that $\mathbb{E}\|X\|_E \|Y\|_E < \infty$ it holds for \mathcal{G} -measurable X that $\mathbb{E}[(X, Y)_E | \mathcal{G}] = (X, \mathbb{E}[Y | \mathcal{G}])_E$ [43, Lemma 2.1.2]. Secondly, the Optional Sampling Theorem, i.e., $\mathbb{E}[M(\tau_{k+1} \wedge t) | \mathcal{F}_{\tau_k \wedge t}] = M(\tau_k \wedge t)$ in the above application, also holds for Hilbert space-valued martingales¹¹. Thus we obtain

$$\begin{aligned} \mathbb{E}[\|M(\tau_{k+1} \wedge t)\|_E^2 | \mathcal{F}_{\tau_k \wedge t}] - \mathbb{E}[\|M(\tau_k \wedge t)\|_E^2 | \mathcal{F}_{\tau_k \wedge t}] \\ = \mathbb{E}[\tilde{N}_2(\tau_{k+1} \wedge t) | \mathcal{F}_{\tau_k \wedge t}] - \mathbb{E}[\tilde{N}_2(\tau_k \wedge t) | \mathcal{F}_{\tau_k \wedge t}]. \end{aligned}$$

Taking the expectation on both sides of this equality and using the induction hypotheses, i.e., the second expectations on both sides of the above equality equate, yields

$$\mathbb{E}\|M(\tau_{k+1} \wedge t)\|_E^2 = \mathbb{E}\tilde{N}_2(\tau_{k+1} \wedge t). \quad (\text{A.11})$$

We conclude the proof extending the Itô-isometry (A.11) from the stopped processes to the non-stopped process. We have already obtained the upper estimate $\mathbb{E}\|M(t)\|_E^2 \leq \mathbb{E}\tilde{N}_2(t)$, cf. (A.9). Hence it remains to prove that a lower bound is given by the same term. As $\|M(t)\|_E^2$ is a real-valued submartingale it holds for all $k \geq 1$ due to the standard Optional Sampling Theorem for càdlàg submartingales, see, e.g., [24, App. B], that

$$\mathbb{E}\|M(t)\|_E^2 \geq \mathbb{E}\|M(\tau_k \wedge t)\|_E^2 = \mathbb{E}\tilde{N}_2(\tau_k \wedge t).$$

Hence, for $k \rightarrow \infty$ we obtain by monotone convergence $\mathbb{E}\|M(t)\|_E^2 \geq \mathbb{E}\tilde{N}_2(t)$ which, combined with the upper bound (A.9), yields the Itô-isometry (3.5). The proof is completed.

B Proofs for the excitable membrane models

B.1 Proof of Theorem 7.1 (Conditions for the LLN)

We apply Theorem 4.1 for the choice of spaces $X = H_0^1(D)$, $H = L^2(D)$ and $E = L^2(D)$. Hence, we have to prove in the following that the assumptions therein are satisfied, i.e., (i) the one-sided Lipschitz condition (4.3) on the operators A and B defined by (7.7), (ii) the Lipschitz condition on the right hand side of the gating system (7.2), (iii) the uniform convergence of the generator and (iv) the martingale convergence. Finally, in (v) we extend the convergence in probability due to Theorem 4.1 to convergence in the mean (7.10). In the following we use \cdot to denote the pointwise product of real functions on D .

¹¹The Optional Sampling Theorem can be proved similarly to the methods employed for [43, Lemma 2.1.2] relying on the linearity properties of the Bochner integral and the monotone convergence theorem.

(i) For the non-linear operator B we find that the left hand side in the Lipschitz condition is for almost all t given by a finite sum of terms

$$\langle p_i \cdot (E_i - u) - \widehat{p}_i \cdot (E_i - v), u - v \rangle_{H^1}, \quad i = 1, \dots, m, \tag{B.1}$$

with $u, v \in H_0^1(D)$ and $p_i, \widehat{p}_i \in L^2(D)$. Hence, the duality pairing corresponds to the inner product in $L^2(D)$. We estimate each of the summand of the type (B.1) separately. Using the triangle inequality we obtain

$$|\langle p_i \cdot (E_i - u) - \widehat{p}_i \cdot (E_i - v), u - v \rangle_{H^1}| \leq |E_i| |(p_i - \widehat{p}_i, u - v)_{L^2}| + |(p_i \cdot u - \widehat{p}_i \cdot v, u - v)_{L^2}|.$$

Here, the first term in this right hand side is further estimated using Cauchy-Schwarz and Young's inequality, which yields

$$|(p_i - \widehat{p}_i, u - v)_{L^2}| \leq \frac{1}{2} \|p_i - \widehat{p}_i\|_{L^2}^2 + \frac{1}{2} \|u - v\|_{L^2}^2.$$

For the second term we obtain, making use of the triangle inequality, Cauchy-Schwarz and Young's inequality and the pointwise bounds on p_i and v , the sequence of estimates

$$\begin{aligned} |(p_i \cdot u - \widehat{p}_i \cdot v, u - v)_{L^2}| &\leq |(p_i \cdot (u - v), u - v)_{L^2}| + |(p_i - \widehat{p}_i, v \cdot (u - v))_{L^2}| \\ &\leq \|p_i \cdot (u - v)\|_{L^2} \|u - v\|_{L^2} + \|p_i - \widehat{p}_i\|_{L^2} \|v \cdot (u - v)\|_{L^2} \\ &\leq \|u - v\|_{L^2}^2 + \frac{\bar{u}^2}{2} \|u - v\|_{L^2}^2 + \frac{1}{2} \|p_i - \widehat{p}_i\|_{L^2}^2. \end{aligned}$$

A summation over all these estimates for $i = 1, \dots, m$ yields

$$\langle B(p, u) - B(\widehat{p}, v), u - v \rangle_{H^1} \leq m(1 + \frac{\bar{u} + \bar{u}^2}{2}) \|u - v\|_{L^2}^2 + \frac{1 + \bar{u}}{2} \sum_{i=1}^m \|p_i - \widehat{p}_i\|_{L^2}^2.$$

Adding the estimate

$$\langle A(u - v), u - v \rangle_{H^1} \leq -\gamma_1 \|u - v\|_{H^1}^2 + \gamma_2 \|u - v\|_{L^2}^2 \leq \gamma_2 \|u - v\|_{L^2}^2$$

for some $\gamma_1, \gamma_2 > 0$, which holds as the linear operator A is coercive and independent of p , we obtain

$$\langle A(u - v), u - v \rangle_{H^1} + \langle B(p, u) - B(\widehat{p}, v), u - v \rangle_{H^1} \leq C \left(\|u - v\|_{L^2}^2 + \sum_{i=1}^m \|p_i - \widehat{p}_i\|_{L^2}^2 \right)$$

for a suitable constant C . Finally, integrating over $(0, T)$ we find the one-sided Lipschitz condition (4.3) is satisfied.

(ii) Due to the triangle inequality it suffices to consider differences of the form $\|p_i \cdot q(u) - \widehat{p}_i \cdot q(v)\|_{L^2}$, where q substitutes for an arbitrary rate function q_{jk} . Using the triangle inequality, the pointwise boundedness of \widehat{p}_i and q by 1 and \bar{q} , respectively, and the Lipschitz condition on the rate functions q (with common Lipschitz constant L) we obtain

$$\begin{aligned} \|p_i \cdot q(u) - \widehat{p}_i \cdot q(v)\|_{L^2} &\leq \|p_i \cdot q(u) - \widehat{p}_i \cdot q(u)\|_{L^2} + \|\widehat{p}_i \cdot q(u) - \widehat{p}_i \cdot q(v)\|_{L^2} \\ &\leq \bar{q} \|p_i - \widehat{p}_i\|_{L^2} + L \|u - v\|_{L^2}. \end{aligned}$$

A summation over all such separate estimates, integrating and squaring both resulting sides yield the Lipschitz condition (4.4).

(iii) In order to prove the convergence of the generators (4.6) we employ in the following two technical results which we collect in a separate proposition. Firstly, the purpose of the formula (B.2) is to transform the generator of the PDMP into a form that allows comparison with the deterministic limit system (7.2). Secondly, the inequality (B.3), which bounds the norm $\|U^n\|_{L^2((0,T),H^1)}$ by a deterministic constant uniformly over $n \in \mathbb{N}$, is used repeatedly in the subsequent estimation procedures.

Proposition B.1. (a) *The generator of the PDMP satisfies*

$$\Lambda^n(u, \theta^n) \int_{K_n} \left(z_i^n(\xi) - z_i^n(\theta^n) \right) \mu^n((u, \theta^n), d\xi) = \sum_{j \neq i} \left(z_j^n(\theta^n) \cdot q_{ji}^n(u) - z_i^n(\theta^n) \cdot q_{ij}^n(u) \right) \tag{B.2}$$

where

$$q_{ij}^n(u) = \sum_{k=1}^{p(n)} Q_{ij}^{k,n}(u) \mathbb{I}_{D_{k,n}} \in L^2(D).$$

(b) *For all $n \in \mathbb{N}$ and all $T > 0$ it holds that*

$$\int_0^T \|U_t^n\|_{H^1}^2 dt \leq C_1(1+T)e^{2C_2T}, \tag{B.3}$$

where the constants C_1, C_2 are deterministic and independent of $n \in \mathbb{N}$.

Proof. (a) We denote by $\theta_{k,i \rightarrow j}^n$ for all $k = 1, \dots, p(n)$ and all $i \neq j, i, j = 1, \dots, m$ the configuration in K_n that arises from the configuration θ^n through the event that a channel in state i located in the compartment $D_{k,n}$ switches to state j . Then simple reorganisation of finite sums yields

$$\begin{aligned} & \Lambda^n(u, \theta^n) \int_{K_n} \left(z_i^n(\xi) - z_i^n(\theta^n) \right) \mu^n((u, \theta^n), d\xi) \\ &= \sum_{k=1}^{p(n)} \sum_{j \neq i} \left(z_i^n(\theta_{k,i \rightarrow j}^n) - z_i^n(\theta^n) \right) \theta_j^{k,n} Q_{ji}^{k,n}(u) + \sum_{k=1}^{p(n)} \sum_{j \neq i} \left(z_i^n(\theta_{k,i \rightarrow j}^n) - z_i^n(\theta^n) \right) \theta_i^{k,n} Q_{ij}^{k,n}(u) \\ &= \sum_{k=1}^{p(n)} \sum_{j \neq i} \left(\frac{1}{l(k,n)} \mathbb{I}_{D_{k,n}} \right) \theta_j^{k,n} Q_{ji}^{k,n}(u) + \sum_{k=1}^{p(n)} \sum_{j \neq i} \left(-\frac{1}{l(k,n)} \mathbb{I}_{D_{k,n}} \right) \theta_i^{k,n} Q_{ij}^{k,n}(u) \\ &= \sum_{j \neq i} z_j^n(\theta^n) \cdot \left(\sum_{k=1}^{p(n)} Q_{ji}^{k,n}(u) \mathbb{I}_{D_{k,n}} \right) - \sum_{j \neq i} z_i^n(\theta^n) \cdot \left(\sum_{k=1}^{p(n)} Q_{ij}^{k,n}(u) \mathbb{I}_{D_{k,n}} \right). \end{aligned}$$

Thus we obtain that the generator satisfies (B.2).

(b) By definition of a PDMP it holds that the component $(U_t^n)_{t \geq 0}$ is the weak solution of the evolution equation

$$\dot{U}_t^n = AU_t^n + \sum_{i=1}^m g_i z_i^n(\Theta_t^n) (E_i - U_t^n)$$

with initial condition U_0^n . We consider the reaction term in this equation as a given inhomogeneity. Then standard estimation procedures from the theory of linear parabolic partial differential equations, cf. [19, Sec. 7], yield, after appropriately estimating the inhomogeneous term,

$$\int_0^T \|U_t^n\|_{H^1}^2 dt \leq K_1 e^{2K_2T} \left(\|U_0^n\|_{L^2}^2 + 2\bar{u}^2 \sum_i \|g_i\|_{L^\infty} \int_0^T \|z_i^n(\Theta_t^n)\|_{L^1} dt \right),$$

where the constants K_1, K_2 are deterministic and depend only on the domain D and the coefficients of A . Further, it holds that $\|z_i^n(\Theta_t^n)\|_{L^1} \leq |D|$ and the sequence of initial conditions is bounded by assumption as $U_0^n(x) \in [\bar{u}_-, \bar{u}_+]$ for all $x \in \bar{D}$ almost surely. The inequality (B.3) follows. \square

We now proceed to the actual proof of the convergence (4.6). To this end we need to consider for almost every t and all $i = 1, \dots, m$, the convergence in $L^2(D)$ of (B.2) to $F_i(z^n(\Theta_t^n), U_t^n)$ where F_i is as defined in (7.2). That is, we have to estimate

$$\left\| \sum_{j \neq i} \left(z_j^n(\Theta_t^n) \cdot q_{ji}^n(U_t^n) - z_i^n(\Theta_t^n) \cdot q_{ij}^n(U_t^n) \right) - \sum_{j \neq i} \left(z_j^n(\Theta_t^n) \cdot q_{ji}(U_t^n) - z_i^n(\Theta_t^n) \cdot q_{ij}(U_t^n) \right) \right\|_{L^2}. \tag{B.4}$$

We find that the single summands in the two summations match up and thus it suffices to consider each of them separately. Employing the boundedness of the coordinate functions, i.e., $\|z_j^n(\Theta_t^n)\|_{L^\infty} \leq 1$ we obtain the estimates

$$\begin{aligned} \left\| z_j^n(\Theta_t^n) \cdot q_{ji}^n(U_t^n) - z_j^n(\Theta_t^n) \cdot q_{ji}(U_t^n) \right\|_{L^2}^2 &= \|z_j^n(\Theta_t^n)\|_{L^\infty}^2 \|q_{ji}^n(U_t^n) - q_{ji}(U_t^n)\|_{L^2}^2 \\ &\leq \left\| \sum_{k=1}^{p(n)} \left(\mathbb{I}_{D_{k,n}} Q_{ij}^{k,n}(U_t^n) \right) - q_{ij}(U_t^n) \right\|_{L^2}^2 \\ &= \sum_{k=1}^{p(n)} \int_{D_{k,n}} |Q_{ij}^{k,n}(U_t^n) - q_{ij}(U_t^n(x))|^2 dx. \end{aligned} \tag{B.5}$$

For the last equality we have used that the summands are mutually orthogonal in $L^2(D)$. Next we estimate each of the remaining integrals in (B.5) using the Lipschitz continuity of q_{ij} and Poincaré’s inequality in $L^2(D_{k,n})$, i.e.,

$$\begin{aligned} \int_{D_{k,n}} \left| q_{ij} \left(\frac{1}{|D_{k,n}|} \int_{D_{k,n}} U_t^n(y) dy \right) - q_{ij}(U_t^n(x)) \right|^2 dx \\ \leq L^2 \int_{D_{k,n}} \left| \frac{1}{|D_{k,n}|} \int_{D_{k,n}} U_t^n(y) dy - U_t^n(x) \right|^2 dx \\ \leq L^2 \pi^{-2} \text{diam}(D_{k,n})^2 \|\nabla U_t^n\|_{L^2(D_{k,n})}^2, \end{aligned}$$

where $\|\nabla U_t^n\|_{L^2(D_{k,n})}$ is the norm in $L^2(D_{k,n})$ of the Euclidean norm of the gradient vector ∇U_t^n . Here we have employed that for convex domains the optimal Poincaré constant is given by $\pi^{-1} \text{diam}(D_{k,n})$ [38]. Hence, a summation over all $k = 1, \dots, p(n)$ and employing the estimate $\|\nabla U_t^n\|_{L^2}^2 \leq \|U_t^n\|_{H^1}^2$ yields

$$\|q_{ji}^n(U_t^n) - q_{ji}(U_t^n)\|_{L^2}^2 \leq \delta_+(n)^2 L^2 \pi^{-2} \|U_t^n\|_{H^1}^2.$$

Integrating over $(0, T)$ we therefore obtain for (B.4) the estimate

$$\begin{aligned} \int_0^T \left\| [\mathcal{A}(\phi, z_j^n(\cdot))_{L^2}](U_t^n, \Theta_t^n) - F_j(z^n(\Theta_t^n), U_t^n) \right\|_{L^2}^2 dt \\ \leq \delta_+(n)^2 L^2 \pi^{-2} 2(m-1) \int_0^T \|U_t^n\|_{H^1}^2 dt. \end{aligned}$$

Finally, the norm $\|U_t^n\|_{L^2((0,T), H^1)}$ is bounded independently of $n \in \mathbb{N}$ by a deterministic constant due to Proposition B.1(b). This upper bound holds for almost all paths of the

PDMPs $(U_t^n, \theta_t^n)_{t \geq 0}$ and thus there exists a constant $C > 0$ independent of n such that

$$\int_0^T \left\| [\mathcal{A}(\phi, z_j^n(\cdot))_{L^2}](U_t^n, \Theta_t^n) - F_j(z^n(\Theta_t^n), U_t^n) \right\|_{L^2}^2 dt \leq \delta_+(n)^2 C \quad (\text{B.6})$$

almost surely. Due to the assumption (7.8) the estimate in the right hand side converges to zero for $n \rightarrow \infty$ and the convergence (4.6) follows.

(iv) Next we consider convergence in probability of the martingale part. To this end we employ Lemma 3.1. As before we denote by $\theta_{k,i \rightarrow j}^n$ the channel configuration that arises from the configuration θ^n if a channel in compartment $D_{k,n}$ switches from state i to state j . Then it holds that

$$\begin{aligned} & \Lambda^n(U_s^n, \Theta_s^n) \int_{K_n} \|z_i^n(\xi) - z_i^n(\Theta_s^n)\|_{L^2}^2 \mu^n((U_s^n, \Theta_s^n), d\xi) \\ &= \sum_{k=1}^{p(n)} \sum_{j \neq i} \left(\|z_i^n(\theta_{k,i \rightarrow j}^n(s)) - z_i^n(\Theta_s^n)\|_{L^2}^2 Q_{ij}^{k,n}(U_s^n) \Theta_i^n(s) \right. \\ & \quad \left. + \|z_i^n(\theta_{k,j \rightarrow i}^n(s)) - z_i^n(\Theta_s^n)\|_{L^2}^2 Q_{ji}^{k,n}(U_s^n) \Theta_j^n(s) \right) \\ &\leq \bar{q} \sum_{k=1}^{p(n)} \frac{|D_{k,n}|}{l(k,n)^2} \sum_{j \neq i} (\Theta_i^n(s) + \Theta_j^n(s)). \end{aligned}$$

This implies that

$$\mathbb{E}^n \int_0^t \left[\Lambda^n(U_s^n, \Theta_s^n) \int_{K_n} \|z_i^n(\xi) - z_i^n(\Theta_s^n)\|_{L^2}^2 \mu^n((U_s^n, \Theta_s^n), d\xi) \right] ds = \mathcal{O}(\ell_-(n)^{-1}).$$

Hence, under condition (7.8) the assumption of Lemma 3.1 is satisfied.

(v) Finally, we extend the convergence in probability to convergence in the mean for the individual components being in the space $L^2((0, T), L^2)$, see the remark following Theorem 4.1. First of all note that the components are bounded, i.e.,

$$\|U_t^n - u(t)\|_{L^2} \leq 2\bar{u}|D|, \quad \|z_i^n(\Theta_t^n) - p_i(t)\|_{L^2} \leq 2|D|.$$

Therefore it holds that

$$\|X^n - X\| := \|U^n - u\|_{L^2((0,T),L^2)} + \sum_{i=1}^m \|z_i^n(\Theta^n) - p_i\|_{L^2((0,T),L^2)} \leq C$$

for a suitable deterministic bound $C < \infty$ independent of $n \in \mathbb{N}$. Then for all $\epsilon_0 > 0$ it holds that

$$\begin{aligned} \mathbb{E}^n \|X_n - X\| &= \mathbb{E}^n [\|X_n - X\| \mathbb{I}_{[\|X_n - X\| \leq \epsilon_0]}] + \mathbb{E}^n [\|X_n - X\| \mathbb{I}_{[\|X_n - X\| > \epsilon_0]}] \\ &\leq \epsilon_0 + M \mathbb{P}^n [\|X_n - X\| > \epsilon_0]. \end{aligned}$$

Next choose $\epsilon_0 < \epsilon/2$ and note that due to the convergence in probability there exists an N_ϵ such that $M \mathbb{P}^n [\|X_n - X\| > \epsilon_0] \leq \epsilon/2$ for all $n > N_\epsilon$. Hence, for every $\epsilon > 0$ there exists an N_ϵ such that $\mathbb{E}^n \|X_n - X\| < \epsilon$ for all $n > N_\epsilon$. Convergence in the mean is proven.

B.2 Proof of Theorem 7.2 (Conditions for the CLT)

In order to prove Theorem 7.2 we employ Theorem 5.1 for the space $E = H^{-2s}(D)$ where s is the smallest integer such that $s > d/2$. We usually employ the simpler notation E and $\mathcal{E} = E^m$ throughout the proof, however occasionally switch to $H^{-2s}(D)$ if we want to emphasise the specific choice of the Hilbert space. The reason choosing this particular integer s is that it is the smallest integer such that the embedding of $H^{2s}(D)$ into $H^s(D)$ is of Hilbert-Schmidt type¹² due to Maurin’s Theorem [2, Thm. 6.61] and $H^s(D)$ is embedded in $C(\bar{D})$ due to the Sobolev Embedding Theorem. These two properties are essential in order to prove the conditions (5.3) – (5.7) of Theorem 5.1. All conditions except (5.6), which establishes the convergence of the quadratic variation, are straightforward consequences of the assumptions of the theorem. These are shown in part (i) of the subsequent proof. For condition (5.6) more involved estimation procedures are necessary which are presented in part (ii).

(i) We first show condition (5.3). As in the preceding section $\theta_{k,i \rightarrow j}^n$ denotes the element of K_n that differs from θ^n by one channel in the k th compartment being in state i instead of state j . Then, the Sobolev Embedding Theorem yields the estimate

$$\|z_i^n(\theta_{k,i \rightarrow j}^n(t)) - z^n(\Theta_t^n)\|_E = \sup_{\|\phi\|_{H^{2s}}=1} |l(k, n)^{-1} \langle \phi, \mathbb{I}_{D^{k,n}} \rangle_{H^{2s}}| \leq \frac{C}{l(k, n)} |D^{k,n}|, \quad (\text{B.7})$$

where C is a constant resulting from the continuous embedding of $H^{2s}(D)$ into $C(\bar{D})$. Using this estimate for the jump heights in the space $H^{-2s}(D)$ we find similarly to part (iv) of the proof of Theorem 7.1 that it holds

$$\alpha_n \mathbb{E}^n \int_0^T \left[\Lambda^n(U_t^n, \Theta_t^n) \int_{K_n} \|z^n(\xi) - z^n(\Theta_t^n)\|_{\mathcal{E}}^2 \mu^n((U_t^n, \Theta_t^n), d\xi) dt \right] = \mathcal{O}(1).$$

Hence, condition (5.3) is satisfied. Moreover, we infer from (B.7) that the rescaled jump sizes are bounded almost surely uniformly, i.e., condition (5.7) is satisfied. Particularly, it holds that $\sqrt{\alpha_n} \|z^n(\theta_{k,i \rightarrow j}^n) - z^n(\theta^n)\|_E = \mathcal{O}(\ell_-(n)^{-1/2})$. This implies that for arbitrary $\beta > 0$ and any $\Phi \in (H^{2s}(D))^m$ there exists N_β such that for all $n \geq N_\beta$

$$\int_{\sqrt{\alpha_n} |\langle \Phi, z_i^n(\xi) - z_i^n(\theta^n) \rangle_{\mathcal{E}}| > \beta} \mu^n((u, \theta^n), d\xi) = 0$$

holds for all values (u, θ^n) the PDMP attains. Therefore, by dominated convergence we infer that also condition (5.5) is satisfied. It remains to consider condition (5.4). To this end let $(\varphi_k)_{k \in \mathbb{N}}$ be an orthonormal basis in $(H^{2s}(D))^m$, where $\varphi_k = (\varphi_k^1, \dots, \varphi_k^m)$ and hence $(\varphi_k^i)_{k \in \mathbb{N}}$ is an orthonormal basis in $H^{2s}(D)$ for all $i = 1, \dots, m$. Then we obtain the estimate

$$\begin{aligned} & \langle \varphi_k, G^n(U_t^n, \Theta_t^n) \varphi_k \rangle_{\mathcal{E}} \\ &= \Lambda^n(U_t^n, \Theta_t^n) \int_{K_n} \left(\sum_{i=1}^m \langle \varphi_k^i, z_i^n(\xi) - z_i^n(\Theta_t^n) \rangle_{H^{-2s}} \right)^2 \mu^n((U_t^n, \Theta_t^n), d\xi) \\ &\leq m \sum_{i=1}^m \|\varphi_k^i\|_{H^s}^2 \left(\Lambda^n(U_t^n, \Theta_t^n) \int_{K_n} \|z_i^n(\xi) - z_i^n(\Theta_t^n)\|_{H^{-s}}^2 \mu^n((U_t^n, \Theta_t^n), d\xi) \right). \end{aligned}$$

Here we have employed for the the individual summands in the right hand side that for $z_i^n(\xi) - z_i^n(\Theta_t^n) \in L^2(D)$ the duality pairing in $H^{2s}(D)$ equals the duality pairing in

¹²The embedding of a Hilbert space X into another Hilbert space H is of *Hilbert-Schmidt type* if $\sum_{k \in \mathbb{N}} \|\varphi_k\|_H^2 < \infty$ for every orthonormal basis $(\varphi_k)_{k \in \mathbb{N}}$ of X .

$H^s(D)$. Further, note that $\|z_i^n(\xi) - z_i^n(\theta_t^n)\|_{H^{-s}}$ satisfies an estimate analogous to (B.7) due to the continuous embedding of $H^s(D)$ in $C(\overline{D})$. Therefore we overall obtain that

$$\alpha_n \langle \varphi_k, G^n(U_t^n, \Theta_t^n) \varphi_k \rangle_{\mathcal{E}} \leq C \sum_{i=1}^m \|\varphi_k^i\|_{H^s}^2$$

for a suitable non-random constant C independent of n . Finally, set $\gamma_k := \sum_{i=1}^m \|\varphi_k^i\|_{H^s}^2$ then it holds that $\sum_{k \in \mathbb{N}} \gamma_k < \infty$ as the embedding $H^{2s}(D) \hookrightarrow H^s(D)$ is of Hilbert-Schmidt type. We infer that condition (5.4) is satisfied.

(ii) In the second part of the proof we establish the central condition (5.6) of the convergence of the quadratic variation. For simplicity of notation we omit the time argument of the PDMP paths and the deterministic solution as the following estimates hold for almost all t . First of all we expand the quadratic variation of the martingales into the finite sum

$$\begin{aligned} & \Lambda^n(U^n, \Theta^n) \int_{K_n} \langle \Phi, z^n(\xi) - z^n(\Theta^n) \rangle_{\mathcal{E}}^2 \mu^n((U^n, \Theta^n), d\xi) \\ &= \sum_{j=1}^m \sum_{\substack{i=1 \\ i \neq j}}^m \sum_{k=1}^{p(n)} \frac{\Theta_j^{k,n}}{l(k,n)^2} Q_{ji}^{k,n}(U^n) \langle \phi_j, \mathbb{I}_{D_{k,n}} \rangle_E^2 + \sum_{j=1}^m \sum_{\substack{i=1 \\ i \neq j}}^m \sum_{k=1}^{p(n)} \frac{\Theta_i^{k,n}}{l(k,n)^2} Q_{ij}^{k,n}(U^n) \langle \phi_j, \mathbb{I}_{D_{k,n}} \rangle_E^2 \\ & \quad - \sum_{\substack{i,j=1 \\ i \neq j}}^m \sum_{k=1}^{p(n)} \left(\frac{\Theta_i^{k,n}}{l(k,n)^2} Q_{ij}^{k,n}(U^n) + \frac{\Theta_j^{k,n}}{l(k,n)^2} Q_{ji}^{k,n}(U^n) \right) \langle \phi_i, \mathbb{I}_{D_{k,n}} \rangle_E \langle \phi_j, \mathbb{I}_{D_{k,n}} \rangle_E. \end{aligned}$$

We find that the terms in this summation match with the integral terms in the definition of the operator $G(u, p)$ in (7.11). Thus, due to the triangle inequality it suffices to consider the convergence of the single summands separately, i.e., we have to consider, on the one hand, for all $j = 1, \dots, m$ and $i \neq j$ the differences

$$\left| \int_D p_j(x) q_{ji}(u(x)) \phi_j^2(x) dx - \alpha_n \sum_{k=1}^{p(n)} \frac{\Theta_j^{k,n}}{l(k,n)^2} Q_{ji}^{k,n}(U^n) \langle \phi_j, \mathbb{I}_{D_{k,n}} \rangle_E^2 \right| \quad (\text{B.8})$$

and, on the other hand, for all $i, j = 1, \dots, m$ such that $i \neq j$ the differences

$$\left| \int_D p_i(x) q_{ij}(u) \phi_i(x) \phi_j(x) dx - \alpha_n \sum_{k=1}^{p(n)} \frac{\Theta_i^{k,n}}{l(k,n)^2} Q_{ij}^{k,n}(U^n) \langle \phi_i, \mathbb{I}_{D_{k,n}} \rangle_E \langle \phi_j, \mathbb{I}_{D_{k,n}} \rangle_E \right|. \quad (\text{B.9})$$

We next estimate these terms separately in parts (ii.1) and (ii.2). Finally, in part (ii.3) the estimates are combined to prove the convergence of the quadratic variation.

(ii.1) A further application of the triangle inequality yields

$$\begin{aligned} (\text{B.8}) &= \left| \int_D p_j(x) q_{ji}(u(x)) \phi_j^2(x) dx - \int_D z_j^n(\Theta^n)(x) q_{ji}(U^n(x)) \phi_j^2(x) dx \right| \\ & \quad + \left| \sum_{k=1}^{p(n)} \frac{\Theta_j^{k,n}}{l(k,n)} \int_{D_{k,n}} q_{ji}(U^n(x)) \phi_j^2(x) dx - \alpha_n \sum_{k=1}^{p(n)} \frac{\Theta_j^{k,n}}{l(k,n)^2} Q_{ji}^{k,n}(U^n) \langle \phi_j, \mathbb{I}_{D_{k,n}} \rangle_E^2 \right|. \end{aligned} \quad (\text{B.10})$$

We estimate the two resulting differences separately and obtain for the first term in the

right hand side of (B.10) the estimate

$$\begin{aligned} & \left| \int_D p_j(x) q_{ji}(u(x)) \phi_j^2(x) dx - \int_D z_j^n(\Theta^n)(x) q_{ji}(U^n(x)) \phi^2(x) dx \right| \\ & \leq \left| \int_D p_j(x) q_{ji}(u(x)) \phi_j^2(x) dx - \int_D z_j^n(\Theta^n)(x) q_{ji}(u(x)) \phi^2(x) dx \right| \\ & \quad + \left| \int_D z_j^n(\Theta^n)(x) q_{ji}(u(x)) \phi^2(x) dx - \int_D z_j^n(\Theta^n)(x) q_{ji}(U^n(x)) \phi^2(x) dx \right| \\ & \leq \bar{q} \|\phi_j\|_{L^\infty}^2 \|p_j - z_j^n(\Theta^n)\|_{L^1} + L \|\phi_j\|_{L^\infty}^2 \|u - U^n\|_{L^1}. \end{aligned} \tag{B.11}$$

For the second term in the right hand side of (B.10) we obtain by employing $\Theta_j^{k,n}/l(k,n) \leq 1$ the estimate

$$\sum_{k=1}^{p(n)} \left| \int_{D_{k,n}} q_{ji}(U^n(x)) \phi_j^2(x) dx - \frac{\alpha_n}{l(k,n)} q_{ji} \left(\frac{1}{|D_{k,n}|} \int_{D_{k,n}} U^n(x) dx \right) \left(\int_{D_{k,n}} \phi_j(x) dx \right)^2 \right| \tag{B.12}$$

and we continue estimating each summand therein separately. We begin employing the Mean Value Theorem to expand the rate function q_{ji} in the integral in the left hand side such that

$$q_{ji}(U^n(x)) = q_{ij} \left(\frac{1}{|D_{k,n}|} \int_{D_{k,n}} U^n(y) dy \right) + q'_{ji}(\vartheta^{k,n}(x)) \left(U^n(x) - \frac{1}{|D_{k,n}|} \int_{D_{k,n}} U^n(y) dy \right), \tag{B.13}$$

where $\vartheta^{k,n}(x)$ denotes an appropriate mean value. For now we omit the remainder term, i.e., the second term in the right hand side of (B.13), a consideration of which is deferred. Hence, we obtain for the absolute value in each summand in (B.12) the estimate

$$q_{ji} \left(\frac{1}{|D_{k,n}|} \int_{D_{k,n}} U^n(y) dy \right) \left| \int_{D_{k,n}} \phi_j^2(x) dx - \frac{\alpha_n}{l(k,n)} \left(\int_{D_{k,n}} \phi_j(x) dx \right)^2 \right|.$$

We note that q_{ji} is bounded by \bar{q} and continue estimating which yields

$$\begin{aligned} & \leq \bar{q} |D_{k,n}| \left| \frac{1}{|D_{k,n}|} \int_{D_{k,n}} \phi_j^2(x) dx - \frac{\alpha_n |D_{k,n}|^2}{l(k,n) |D_{k,n}|} \left(\frac{1}{|D_{k,n}|} \int_{D_{k,n}} \phi_j(x) dx \right)^2 \right| \\ & \leq \bar{q} \int_{D_{k,n}} \left(\phi_j(x) - \frac{1}{|D_{k,n}|} \int_{D_{k,n}} \phi_j(y) dy \right)^2 dx \end{aligned} \tag{B.14}$$

$$+ \bar{q} |D_{k,n}| \left| \left(1 - \frac{\alpha_n |D_{k,n}|^2}{l(k,n) |D_{k,n}|} \right) \left(\frac{1}{|D_{k,n}|} \int_{D_{k,n}} \phi_j(x) dx \right)^2 \right| \tag{B.15}$$

The term (B.14) is estimated using Poincaré’s inequality which yields an upper bound by $\bar{q} \pi^{-2} \text{diam}^2(D_{k,n}) \|\nabla \phi\|_{L^2(D_{k,n})}^2$. For the terms (B.15) a summation over all $k = 1, \dots, p(n)$ yields

$$\bar{q} \sum_{k=1}^{p(n)} |D_{k,n}| \left| 1 - \frac{\alpha_n |D_{k,n}|^2}{l(k,n) |D_{k,n}|} \right| \left(\frac{1}{|D_{k,n}|} \int_{D_{k,n}} \phi_j(x) dx \right)^2 \leq \bar{q} \left| 1 - \frac{\ell_-(n) \nu_-(n)}{\ell_+(n) \nu_+(n)} \right| \|\phi_j^n\|_{L^2}^2, \tag{B.16}$$

where ϕ_j^n is a piecewise constant approximation to ϕ_j defined by

$$\phi_j^n := \sum_{k=1}^{p(n)} \left(\frac{1}{|D_{k,n}|} \int_{D_{k,n}} \phi_j(x) \, dx \right) \mathbb{I}_{D_{k,n}}.$$

As ϕ_j^n converges to ϕ_j in $L^2(D)$ it holds that the sequence of norms converge, hence $\|\phi_j^n\|_{L^2}$ is a bounded sequence. Therefore the right hand side in (B.16) is a component-wise product of convergent sequences. The sequence $|1 - (\ell_-(n) \nu_-(n) / (\ell_+(n) \nu_+(n)))|$ converges to zero, cf. condition (7.12), thus the right hand side in (B.16) converges to zero for $n \rightarrow \infty$.

Finally, it remains to consider the term arising from the remainder in the expansion of q_{ji} , see (B.13), inserted into (B.12). By assumption q'_{ji} is bounded (by a constant \bar{q}). Therefore we obtain an upper bound on the respective term by

$$\begin{aligned} \bar{q} \|\phi_j\|_{L^\infty}^2 \sum_{k=1}^{p(n)} \int_{D_{n,k}} \left| U^n(x) - \frac{1}{|D_{k,n}|} \int_{D_{k,n}} U^n(y) \, dy \right| dx &\leq \bar{q} \|\phi_j\|_{L^\infty}^2 \sum_{k=1}^{p(n)} \frac{\delta_+(n)}{2} \|\nabla U^n\|_{L^1(D_{n,k})} \\ &\leq \bar{q} \|\phi_j\|_{L^\infty}^2 \frac{\delta_+(n)}{2} \|\nabla U^n\|_{L^1}. \end{aligned}$$

Here we have employed the Poincaré inequality in L^1 with optimal Poincaré constant given by $\text{diam}(D_{k,n})/2$ [1].

A combination of these estimates yields an upper bound to (B.8) by

$$(B.8) \leq C_\Phi \left(\|p_j - z_j^n(\Theta^n)\|_{L^1} + \|u - U^n\|_{L^1} + \delta_+(n)^2 \|\nabla U^n\|_{L^1} + \delta_+^2(n) + \delta_+(n) + R(n) \right), \tag{B.17}$$

where the term $R(n)$ is given by the right hand side of (B.16) and converges to zero for $n \rightarrow \infty$. The constant $C_\Phi < \infty$ is a suitable deterministic constant independent of $n \in \mathbb{N}$ which depends on $\Phi \in (H^{2s}(D))^m$ via the norm in $H^s(D)$ of the components of Φ .

(ii.2) Next we consider the mixed terms (B.9). Analogously to part (ii.1) we apply the triangle inequality and obtain

$$\begin{aligned} (B.9) \leq & \left| \int_D p_j(x) q_{ji}(u(x)) \phi_j(x) \phi_i(x) \, dx - \int_D z_j^n(\Theta^n)(x) q_{ji}(U^n(x)) \phi_j(x) \phi_i(x) \, dx \right| \\ & + \left| \sum_{k=1}^{p(n)} \frac{\Theta_j^{k,n}}{l(k,n)} \int_{D_{k,n}} q_{ji}(U^n(x)) \phi_j(x) \phi_i(x) \, dx \right. \\ & \quad \left. - \alpha_n \sum_{k=1}^{p(n)} \frac{\Theta_j^{k,n}}{l(k,n)^2} Q_{ji}^{k,n}(U^n) \langle \phi_j, \mathbb{I}_{D_{k,n}} \rangle_E \langle \phi_i, \mathbb{I}_{D_{k,n}} \rangle_E \right| \end{aligned}$$

As in (ii.1) we obtain for the first term in this right hand side an upper bound by

$$\bar{q} \|\phi_i\|_{L^\infty} \|\phi_j\|_{L^\infty} \|p_j - z_j^n(\Theta^n)\|_{L^1} + L \|\phi_i\|_{L^\infty} \|\phi_j\|_{L^\infty} \|u - U^n\|_{L^1}.$$

Also the second term is treated as in (ii.1), i.e., applying the Mean Value Theorem and estimating the resulting terms accordingly. In particular the remainder term is

estimated completely analogously. Therefore, the only term we are left to estimate is

$$\bar{q} |D_{k,n}| \left| \frac{1}{|D_{k,n}|} \int_{D_{k,n}} \phi_i(x) \phi_j(x) dx - \left(\frac{1}{|D_{k,n}|} \int_{D_{k,n}} \phi_i(x) dx \right) \left(\frac{1}{|D_{k,n}|} \int_{D_{k,n}} \phi_j(x) dx \right) \right| \tag{B.18}$$

$$+ \bar{q} |D_{k,n}| \left| \left(1 - \frac{\alpha_n |D_{k,n}|^2}{l(k,n) |D_{k,n}|} \right) \left(\frac{1}{|D_{k,n}|} \int_{D_{k,n}} \phi_i(x) dx \right) \left(\frac{1}{|D_{k,n}|} \int_{D_{k,n}} \phi_j(x) dx \right) \right|. \tag{B.19}$$

First of all, using Young’s inequality we obtain for the second term the estimate

$$(B.19) \leq \frac{\bar{q}}{2} \left| 1 - \frac{\ell_-(n) \nu_-(n)}{\ell_+(n) \nu_+(n)} \right| (\|\phi_i^n\|_{L^2}^2 + \|\phi_j^n\|_{L^2}^2), \tag{B.20}$$

which converges to zero for $n \rightarrow \infty$.

We next estimate the term (B.18). Firstly, we note that as in part (a) we find using Poincaré’s inequality an upper bound to the term

$$|D_{k,n}| \left| \frac{1}{|D_{k,n}|} \int_{D_{k,n}} (\phi_i(x) - \phi_j(x))^2 dx - \left(\frac{1}{|D_{k,n}|} \int_{D_{k,n}} \phi_i(x) - \phi_j(x) dx \right)^2 \right| \tag{B.21}$$

and the upper bound is proportional to $\delta_+(n)^2$. Next, expanding the two squared terms in (B.21) we find using the reverse triangle inequality that the term (B.21) is an upper bound to

$$\begin{aligned} |D_{k,n}| \left| \left| \frac{1}{|D_{k,n}|} \int_{D_{k,n}} \phi_i(x)^2 dx + \frac{1}{|D_{k,n}|} \int_{D_{k,n}} \phi_j(x)^2 dx \right. \right. \\ \left. \left. - \left(\frac{1}{|D_{k,n}|} \int_{D_{k,n}} \phi_i(x) dx \right)^2 - \left(\frac{1}{|D_{k,n}|} \int_{D_{k,n}} \phi_j(x) dx \right)^2 \right| \\ - 2 \left| \frac{1}{|D_{k,n}|} \int_{D_{k,n}} \phi_i(x) \phi_j(x) dx - \left(\frac{1}{|D_{k,n}|} \int_{D_{k,n}} \phi_i(x) dx \right) \left(\frac{1}{|D_{k,n}|} \int_{D_{k,n}} \phi_j(x) dx \right) \right|. \end{aligned}$$

Thus also this term possesses an upper bound which is proportional to $\delta_+(n)^2$. For $n \rightarrow \infty$ the upper bound converges to zero. As for $\delta_+(n) \rightarrow 0$ also the term spanning the first and second line converges to zero which was established in (ii.1), necessarily also the term in the third line converges to zero. Therefore we infer that the term (B.18) converges to zero proportional to $\delta_+(n)^2$.

Now, a combination of these estimates yields analogously to (B.17) in (ii.1) that

$$(B.9) \leq C_\Phi \left(\|p_j - z_j^n(\theta^n)\|_{L^2} + \|u - U^n\|_{L^2} + \delta(n)^2 \|\nabla U^n\|_{L^1} + \delta^2(n) + \delta(n) + R(n) \right). \tag{B.22}$$

Here $R(n)$ is a term converging to zero for $n \rightarrow \infty$ arising from (B.20) and it is of the same type as the term $R(n)$ in (ii.1). The deterministic constant C_Φ is independent of $n \in \mathbb{N}$ and depends on Φ via the norm in $H^s(D)$ of the components of Φ .

(ii.3) A combination of the final results (B.17) and (B.22) in (ii.1) and (ii.2) yields that there exists a constant $C_\Phi < \infty$ such that for almost all t it holds that

$$\begin{aligned} \left| \langle \Phi, G(u(t), p(t)) \Phi \rangle_{\mathcal{E}} - \alpha_n \langle \Phi, G^n(U_t^n, \Theta_t^n) \Phi \rangle_{\mathcal{E}} \right| \\ \leq C_\Phi \left(\sum_{i=1}^m \|p_i(t) - z_i^n(\Theta_t^n)\|_{L^2} + \|u(t) - U^n(t)\|_{L^2} + \delta(n)^2 \|\nabla U^n\|_{L^2} + \delta^2(n) + \delta(n) + R(n) \right) \end{aligned}$$

Here we have also employed the continuous embedding of $L^2(D) \hookrightarrow L^1(D)$. We next square both sides of this inequality and integrate over $(0, T)$. Afterwards we take the square root of the integral terms and further take the expectation of the resulting inequality. Finally, appropriate applications of Jensen's inequality yields that

$$\begin{aligned} & \int_0^T \mathbb{E}^n \left| \langle \Phi, G(u(t), p(t)) \Phi \rangle_{\mathcal{E}} - \alpha_n \langle \Phi, G^n(U_t^n, \Theta_t^n) \Phi \rangle_{\mathcal{E}} \right| dt \\ & \leq C_{\Phi, T} \left(\delta^2(n) + \delta(n) + R(n) + \mathbb{E}^n \left[\|u - U^n\|_{L^2((0, T), L^2)} + \sum_{i=1}^m \|p_i - z_i^n(\Theta^n)\|_{L^2((0, T), L^2)} \right] \right) \end{aligned} \quad (\text{B.23})$$

for an appropriate constant $C_{T, \Phi} < \infty$. Note that in order to arrive at the estimate (B.23) we have further employed that the random term $\|\nabla U^n\|_{L^2((0, T), L^2)}$ can be estimated by a deterministic bound independent of $n \in \mathbb{N}$ due to Proposition B.1 (b). Finally, due to the law of large numbers, i.e., Theorem 7.1, the sequence of PDMPs converges to the deterministic limit in the mean. Hence the expectation in the right hand side in (B.23) converges to zero for $n \rightarrow \infty$. Furthermore, $\delta_+(n)$ converges to zero by assumption (7.8), as does the term $R(n)$. Thus, overall the right hand side in (B.23) converges to zero. The convergence of the quadratic variation is proved which completes the proof of Theorem 7.1.

References

- [1] G. Acosta and R. G. Durán. An optimal Poincaré inequality in L_1 for convex domains. *Proc. Amer. Math. Soc.*, 132(1):195–202, 2004. MR-2021262
- [2] R. A. Adams and J. F. J. Fournier. *Sobolev Spaces, 2nd Ed.* Academic Press, 2003. MR-2424078
- [3] L. Arnold. *Stochastische Differentialgleichungen.* Oldenbourg, München, 1973. MR-0443082
- [4] L. Arnold and M. Theodosopulu. Deterministic limit of the stochastic model of chemical reactions with diffusion. *Adv. Appl. Prob.*, 12(2):367–379, 1980. MR-0569433
- [5] T. D. Austin. The emergence of the deterministic Hodgkin-Huxley equations as a limit from the underlying stochastic ion-channel mechanism. *Ann. Appl. Prob.*, 18(4):1279–1325, 2008. MR-2434172
- [6] P. Billingsley. *Convergence of Probability Measures, 2nd Ed.* Wiley, New York, 1999. MR-1700749
- [7] D. Blount. Comparison of stochastic and deterministic models of a linear chemical reaction with diffusion. *Ann. Probab.*, 19(4):1440–1462, 1991. MR-1127711
- [8] D. Blount. Law of large numbers in the supremum norm for a chemical reaction with diffusion. *Ann. Appl. Probab.*, 2(1):131–141, 1992. MR-1143396
- [9] D. Blount. Limit theorems for a sequence of nonlinear reaction-diffusion systems. *Stoch. Proc. Appl.*, 45(2):193–207, 1993. MR-1208868
- [10] D. Blount. Density-dependent limits for a nonlinear reaction-diffusion model. *Ann. Probab.*, 22(4):2040–2070, 1994. MR-1331215
- [11] D. Blount. Diffusion limits for a nonlinear density dependent space-time population model. *Ann. Probab.*, 24(2):639–659, 1996. MR-1404523
- [12] P. C. Bressloff. Stochastic neural field theory and the system-size expansion. *SIAM J. Appl. Math.*, 70:1488–1521, 2009. MR-2578680
- [13] E. Buckwar and M. G. Riedler. Exact modelling of neuronal membranes including spatio-temporal evolution. *J. Math. Bio.*, 63(6):1053–1091, 2011. MR-2855804
- [14] G. Da Prato and J. Zabczyk. *Stochastic Equations in Infinite Dimensions.* Cambridge University Press, Cambridge, 1992. MR-1207136
- [15] M. H. A. Davis. *Lectures on Stochastic Control and Nonlinear Filtering.* Springer, 1984.

- [16] M. H. A. Davis. Piecewise-deterministic Markov processes: a general class of non-diffusion stochastic models (with discussion). *J. R. Stat. Soc., Ser. B*, 46:353–388, 1984. MR-0790622
- [17] M. H. A. Davis. *Markov Models and Optimisation*. Chapman and Hall, London, 1993. MR-1283589
- [18] S. N. Ethier and Kurtz T. G. *Markov Processes: Characterization and Convergence*. Wiley, New York, 1986. MR-0838085
- [19] L. C. Evans. *Partial Differential Equations, 2nd Ed*. AMS, 2010. MR-2597943
- [20] M. Falcke. On the role of stochastic channel behaviour in intracellular Ca^{2+} dynamics. *Biophys. J.*, 84:42–56, 2002.
- [21] F. H. Fenton and E. M. Cherry. Models of cardiac cell. *Scholarpedia*, 8(3):1868, 2008.
- [22] A. Genadot and M. Thieullen. Averaging for a fully coupled Piecewise Deterministic Markov Process in infinite dimension. arXiv:1109.6581, 2011.
- [23] A. L. Hodgkin and A. F. Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. *J. Physiol.*, 117:500–544, 1952.
- [24] M. Jacobsen. *Point Process Theory and Applications: Marked Point and Piecewise Deterministic Processes*. Birkhäuser, Boston, 2006. MR-2189574
- [25] G. Kallianpur and J. Xiong. Diffusion approximation of nuclear space-valued stochastic differential equations driven by Poisson random measures. *Ann. Appl. Probab.*, 5(2):493–517, 1995. MR-1336880
- [26] C. Koch. *Biophysics of Computation*. Oxford University Press, New York, 1999.
- [27] P. Kotelenez. Law of large numbers and central limit theorem for linear chemical reactions with diffusion. *Ann. Probab.*, 14(1):173–193, 1986. MR-0815964
- [28] P. Kotelenez. Linear parabolic differential equations as limits of space-time jump Markov processes. *J. Math. Anal. Appl.*, 116(1):42–76, 1986. MR-0837340
- [29] P. Kotelenez. Fluctuations near homogeneous states of chemical reactions with diffusion. *Adv. Appl. Probab.*, 19(2):352–370, 1987. MR-0889941
- [30] P. Kotelenez. High density limit theorems for nonlinear chemical reactions with diffusion. *Prob. Th. Rel. Fields*, 78(1):11–37, 1988. MR-0940864
- [31] T. G. Kurtz. Solutions of ordinary differential equations as limits of pure jump Markov processes. *J. Appl. Probab.*, 7:49–58, 1970. MR-0254917
- [32] T. G. Kurtz. Limit theorems for a sequence of jump Markov processes approximating ordinary differential equations. *J. Appl. Probab.*, 8:344–356, 1971. MR-0287609
- [33] K. Liu. *Stability of Infinite Dimensional Stochastic Differential Equations with Applications*. Chapman & Hall/CRC, 2006. MR-2165651
- [34] B. Mélykúti, K. Burrage, and K. C. Zygalakis. Fast stochastic simulation of biochemical reaction systems by alternative formulations of the chemical langevin equation. *J. Chem. Phys.*, 132:164109, 2010.
- [35] M. Métivier. *Semimartingales*. deGruyter, Berlin, 1982. MR-0688144
- [36] M. Métivier. Convergence faible et principe d’invariance pour des martingales à valeurs dans des espaces de sobolev. *Ann. Inst. Henri Poincaré, Sec. B*, 20(4):329–348, 1984. MR-0771893
- [37] K. Pakdaman, M. Thieullen, and G. Wainrib. Fluid limit theorems for stochastic hybrid systems with application to neuron models. *Adv. in Appl. Probab.*, 42(3):761–794, 2010. MR-2779558
- [38] L. E. Payne and H. F. Weinberger. An optimal Poincaré inequality for convex domains. *Arch. Rat. Mech. Anal.*, 5:286–292, 1960. MR-0117419
- [39] C. Prevot and M. Roeckner. *A Concise Course on Stochastic Partial Differential Equations*. Springer, Berlin, 2007. MR-2329435
- [40] M. Renardy and R. C. Rogers. *Introduction to Partial Differential Equations, 2nd Ed*. Springer, 2004. MR-2028503

- [41] M. G. Riedler. Approximation of stochastic hybrid systems. In *Oberwolfach Reports, Report Nr. 40: Mini-workshop: Dynamics of Stochastic Systems and their Approximation*. European Mathematical Society, 2011.
- [42] M. G. Riedler. *Spatio-temporal Stochastic Hybrid Models of Excitable Biological Membranes*. PhD thesis, Heriot-Watt University, 2011.
- [43] S. Stolze. Stochastic equations in Hilbert space with Lévy noise and their applications in finance. Master's thesis, Universität Bielefeld, 2005.
- [44] S. Swillens, P. Campeil, L. Combettes, and G. Dupont. Stochastic simulation of a single inositol 1,4,5-triphosphate-sensitive Ca^{2+} channel reveals repetitive openings during 'blip-like' Ca^{2+} transients. *Cell Calcium*, 23(5):291–302, 1998.
- [45] D. Vermes. Optimal control of piecewise deterministic Markov processes. *Stochastics*, 14:165–207, 1985. MR-0800243
- [46] E. Zeidler. *Nonlinear Functional Analysis and its Applications, Vol. 2*. Springer, New York, 1990.