

# LARGE DEVIATIONS FOR WHITE-NOISE DRIVEN, NONLINEAR STOCHASTIC PDES IN TWO AND THREE DIMENSIONS

MARTIN HAIRER AND HENDRIK WEBER

ABSTRACT. We study the stochastic Allen-Cahn equation driven by a noise term with intensity  $\sqrt{\varepsilon}$  and correlation length  $\delta$  in two and three spatial dimensions. We study diagonal limits  $\delta, \varepsilon \rightarrow 0$  and describe fully the large deviation behaviour depending on the relationship between  $\delta$  and  $\varepsilon$ .

The recently developed theory of regularity structures allows to fully analyse the behaviour of solutions for vanishing correlation length  $\delta$  and fixed noise intensity  $\varepsilon$ . One key fact is that in order to get non-trivial limits as  $\delta \rightarrow 0$ , it is necessary to introduce diverging counterterms. The theory of regularity structures allows to rigorously analyse this renormalisation procedure for a number of interesting equations.

Our main result is a large deviation principle for these renormalised solutions. One interesting feature of this result is that the diverging renormalisation constants disappear at the level of the large deviations rate function. We apply this result to derive a sharp condition on  $\delta, \varepsilon$  that guarantees a large deviation principle for diagonal schemes  $\varepsilon, \delta \rightarrow 0$  for the equation without renormalisation.

RÉSUMÉ. Nous étudions l'équation d'Allen-Cahn stochastique conduite par un bruit d'intensité  $\sqrt{\varepsilon}$  et de longueur de corrélation  $\delta$  en dimensions spatiales deux et trois. Nous considérons la limite  $\delta, \varepsilon \rightarrow 0$  et nous décrivons complètement le comportement des grandes déviations associées, suivant les relations entre  $\delta$  et  $\varepsilon$ .

La théorie des structures de régularité récemment développée permet d'analyser le comportement des solutions à intensité de bruit  $\varepsilon$  fixée dans la limite  $\delta \rightarrow 0$ . Un fait crucial est que, afin d'obtenir des limites non-triviales dans cette limite, il est nécessaire d'introduire des contretermes divergents. La théorie des structures de régularité permet d'analyser rigoureusement de telles procédures de renormalisation pour un nombre d'équations intéressantes.

Notre résultat principal est un principe de grandes déviations pour ces équations renormalisées. Il est alors intéressant de noter que les constantes de renormalisation divergentes disparaissent au niveau de la fonction de taux. Une conséquence de ce résultat est une condition optimale sur le comportement relatif de  $\delta$  et  $\varepsilon$  qui garantit l'existence d'un principe de grandes déviations également pour l'équation non-renormalisée dans certains régimes.

## 1. INTRODUCTION

The purpose of this article is to provide large deviation results for a class of non-linear stochastic PDEs which are driven by space-time white noise, in 2 and 3 spatial dimensions. We are going to study equations of the type

$$\partial_t u = \Delta u + Cu - u^3 + \sqrt{\varepsilon} \xi_\delta, \quad (1.1)$$

where  $\xi_\delta$  is some random driving noise (we will be more specific very soon) and  $C \in \mathbf{R}$ . In order to avoid complications coming from the effect of boundary

conditions, we will always assume that the spatial variable takes values in the  $d$ -dimensional torus  $\mathbf{T}^d$  and that the initial datum  $u_0$  is fixed.

Equations of this type are popular as phenomenological descriptions in various situations, e.g. for phase separation (see e.g. [HH77]). From a physical point of view, the interesting regime is that where the noise is weak (i.e. the typical strength of the noise, say when tested against smooth test functions, is of order  $\sqrt{\varepsilon} \ll 1$ ) and almost white (i.e. correlations of  $\xi$  decay on a lengthscale  $\delta \ll 1$ ). In this case, we denote the noise by  $\sqrt{\varepsilon}\xi_\delta$ .

In one spatial dimension the solution theory for equation (1.1) is well-understood, even in the case of vanishing correlation length  $\delta = 0$ , i.e. in the case where the noise term  $\xi$  is a space-time white noise. In this situation the limit of vanishing noise strength was described in detail by [FJL82] on the level of large deviation estimates.

Recently, there have been several works dedicated to studying (1.1) driven by  $\sqrt{\varepsilon}\xi_\delta$  for  $\varepsilon \ll \delta \ll 1$  in arbitrary spatial dimension. For fixed  $\delta > 0$ , the law of (1.1) satisfies a large deviation principle as  $\varepsilon \rightarrow 0$  with rate  $\varepsilon$  and some rate function  $\mathcal{I}_\delta$ . Formally, as  $\delta \rightarrow 0$ , the sequence of rate functions  $\mathcal{I}_\delta$  converges to a limiting “rate function”  $\mathcal{I}$  given by

$$\mathcal{I}(u) = \frac{1}{2} \int_0^T \int_{\mathbf{T}^d} (\partial_t u - \Delta u - Cu + u^3)^2 dx dt .$$

(With the understanding that  $\mathcal{I}(u)$  is infinite if the distribution  $\partial_t u - \Delta u - Cu + u^3$  is not represented by a square integrable function or if  $u$  does not satisfy the initial condition.) Minimisers of the rate function  $\mathcal{I}$  subject to certain initial and terminal conditions were investigated in [ERVE04, KORVE07]. In [CF11] Cerrai and Freidlin showed that the convergence  $\mathcal{I}_\delta \rightarrow \mathcal{I}$  holds in the sense of  $\Gamma$ -convergence (with respect to a suitable topology) for arbitrary dimension  $d$ . They used this result to conclude  $\Gamma$ -convergence of the associated *quasi potentials*.

These results naturally suggest that  $\mathcal{I}$  should be the rate function for a problem which does not involve  $\delta$  anymore. It would be very natural to interpret  $\mathcal{I}$  as the rate function for the solutions to (1.1) with noise  $\sqrt{\varepsilon}\xi$  where  $\xi$  denotes space-time white noise. (In one spatial dimensions  $d = 1$  it is shown in [FJL82] that this is indeed true). The problem that immediately presents itself is that while the result of [CF11] holds in any dimension, the equation (1.1) (with fixed  $C$ ) driven by space-time white noise is ill-posed in any dimension  $d \geq 2$ . As a matter of fact, if we denote by  $u_\delta^{(\varepsilon)}$  the solution to (1.1) driven by  $\sqrt{\varepsilon}\xi_\delta$  and by  $u_{\text{det}}$  the deterministic solution to the equation with  $\xi = 0$ , then one has

$$\lim_{\delta \rightarrow 0} \lim_{\varepsilon \rightarrow 0} u_\delta^{(\varepsilon)} = u_{\text{det}} .$$

On the other hand, it was shown in [HRW12] that, already in dimension 2, one has  $\lim_{\delta \rightarrow 0} u_\delta^{(\varepsilon)} = 0$  in a space of distributions for any fixed value of  $\varepsilon$ !

There is, however, a way to obtain a non-trivial stochastic limit in a scheme with  $\delta \rightarrow 0$ , but the equation (1.1) has to be modified: Consider the solution  $\hat{u}_\delta^{(\varepsilon)}$  to

$$\partial_t u = \Delta u + (C + 3\varepsilon C_\delta^{(1)} - 9\varepsilon^2 C_\delta^{(2)})u - u^3 + \sqrt{\varepsilon}\xi_\delta , \quad (1.2)$$

where the spatial variable takes values in the  $d$ -dimensional torus with  $d \in \{2, 3\}$ . In two spatial dimensions it was shown in [DPD03] that for every fixed  $\varepsilon$  the solutions to (1.2) have a non-trivial limit as  $\delta \rightarrow 0$ , provided that the  $\delta$ -dependent constants

are suitably chosen. Recently, in [Hai14c], the first named author developed a theory of regularity structures allowing to construct a non-trivial limit also in three spatial dimensions. We denote these limits by  $\hat{u}^{(\varepsilon)}$ .

*Remark 1.1.* In dimension  $d = 2$ , one can actually take  $C_\delta^{(2)} = 0$  and  $C_\delta^{(1)} = \frac{1}{4\pi} |\log \delta|$ . In dimension  $d = 3$  on the other hand, one has  $C_\delta^{(1)} \propto \delta^{-1}$  with a proportionality constant depending on the details of the regularisation of  $\xi_\delta$ , see (2.24). Furthermore, it is  $C_\delta^{(2)}$  which should be taken proportional to  $|\log \delta|$  in this case.

The purpose of this article is to investigate the large deviations of the solutions obtained from the scheme (1.2) as  $\varepsilon \rightarrow 0$ . Our main result, Theorem 4.4, states that for every  $T > 0$  and every function  $\varepsilon \mapsto \delta(\varepsilon) \geq 0$  with  $\lim_{\varepsilon \rightarrow 0} \delta(\varepsilon) = 0$ , the laws of the sequence  $\hat{u}_{\delta(\varepsilon)}^{(\varepsilon)}$  satisfy a large deviation principle in the space  $\mathcal{C}([0, T], \mathcal{C}^{1-d/2-\kappa})$  (for arbitrarily small  $\kappa > 0$ ) with rate  $\varepsilon$  and rate function  $\mathcal{I}$  (see below in Section 2 for a discussion of spaces with negative regularity). This result includes the case  $\delta = 0$ : The processes  $\hat{u}^{(\varepsilon)}$  do not depend on a correlation length  $\delta$  and satisfy a large deviation principle with rate function  $\mathcal{I}$ . A remarkable feature of this result is that the *diverging renormalisation constants*  $C_\delta^{(1)}$  and  $C_\delta^{(2)}$ , present in (1.2), disappear on the level of the large deviations, independently of the relationship between  $\delta$  and  $\varepsilon$ .

As an immediate consequence of our method, in Theorem 4.7, we make more precise the condition on  $\delta$  and  $\varepsilon$  under which the solutions  $u_\delta^{(\varepsilon)}$  of the equation (1.1) without renormalisation constants satisfy a large deviation principle. Assume that the function  $\delta(\varepsilon)$  satisfies

$$\lim_{\varepsilon \rightarrow 0} \varepsilon \delta(\varepsilon)^{-1} = \lambda^2 \in [0, \infty],$$

for  $d = 3$ , or  $\lim_{\varepsilon \rightarrow 0} \varepsilon \log(\delta(\varepsilon)^{-1}) = \lambda^2$  for  $d = 2$ . If  $\lambda = 0$  the solutions  $u_\delta^{(\varepsilon)}$  also satisfy a large deviation principle with rate function  $\mathcal{I}$ . If  $\lambda \in (0, \infty)$ , we still obtain a large deviation principle, but this time with a *modified* rate function that depends on  $\lambda$ . (And on the regularisation of  $\xi_\delta$  in dimension 3.)

The threshold  $\varepsilon \ll (\log(\delta^{-1}))^{-1}$  for  $d = 2$  and  $\varepsilon \ll \delta$  for  $d = 3$  is not surprising. Indeed, it is quite straightforward to check without reference to the theory of regularity structures that if this condition is satisfied, then as  $\varepsilon \rightarrow 0$  the solutions  $u_{\delta(\varepsilon)}^{(\varepsilon)}$  converge in probability to  $u_{\text{det}}$ , with respect to the topology of uniform convergence (see [HRW12, Section 4] for the argument if  $d = 2$ ). Indeed, the same statement holds for arbitrary spatial dimension  $d \geq 4$  provided  $\varepsilon \ll \delta^{d-2}$ . It is however important to note that our method to prove the large deviation principle relies strongly on the understanding of the renormalised equation *even for the schemes without renormalisation*. In particular, it is not clear whether a large deviations principle holds in higher dimensions, even in the regime  $\varepsilon \ll \delta^{d-2}$ .

We would also like to emphasise that the methods developed in this article are not restricted to the specific equation (1.1). The same arguments would yield similar results for all equations that can be treated with the methods developed in [Hai14c]. In particular, in  $d = 2$  the nonlinear term  $u^3$  in (1.1) could be replaced by an arbitrary polynomial of odd degree with negative leading-order coefficient, and one has a large deviations principle for the KPZ equation driven by small noise. In

the same way we could obtain a large deviation principle for the two dimensional Navier-Stokes equations driven by space-time white noise (see [DPD02] for the solution theory at  $\varepsilon > 0$  fixed, see [BLZ14] for an analysis of the corresponding rate function).

Finally, we point out that partial results in two spatial dimension were already obtained by Jona-Lasinio and Mitter in [JLM90]. In [Aid12, Aid09] Aida studied a related model in one spatial dimension, but with scaling properties akin to the two-dimensional case. We also want to mention that the theory of regularity structures grew out of an attempt to generalise the theory of *rough paths* to higher dimensions. Large deviation results for rough paths were obtained by [LQZ02, MSS06, FV07].

**1.1. Structure of the article.** In Section 2 we present a very short summary of some notions of the theory of regularity structures. In Section 3 we discuss random variables in a fixed Wiener chaos taking values in a separable Banach space and their large deviations. Section 4 contains the statements and proofs of our main results.

**Acknowledgements.** We are grateful to the anonymous referee for carefully reading the original manuscript and making various suggestions that improved the exposition. MH was supported by the Leverhulme trust through a leadership award and the Royal Society through a Wolfson research award. MH would also like to thank the Institute for Advanced Study for its warm hospitality and the ‘The Fund for Math’ for funding his stay there, over the course of which this work was completed. HW was supported by an EPSRC First Grant.

## 2. REGULARITY STRUCTURES

In order to prove the type of convergence result mentioned in the introduction, we make use of the theory of regularity structures developed in [Hai14c]. A complete self-contained exposition of the theory is of course beyond the scope of this article, so we will content ourselves with a short summary of the theory’s main concepts and results, when specialised to the specific example of the stochastic Allen-Cahn equation (1.1). For a concise exposition of the general theory, see also the lecture notes [Hai14a].

The main ingredient of the theory is that of a *regularity structure*. In our case, this consists of a graded vector space  $\mathcal{T} = \bigoplus_{\alpha \in A} \mathcal{T}_\alpha$  where  $A$  denotes a set of real-valued indices (called homogeneities) that is locally finite and bounded from below. Each of the spaces  $\mathcal{T}_\alpha$  will be finite-dimensional and come with a distinguished canonical basis. The space  $\mathcal{T}$  also comes endowed with a group  $\mathcal{G}$  of continuous linear transformations of  $\mathcal{T}$  with the property that, for every  $\Gamma \in \mathcal{G}$ , every  $\alpha \in A$ , and every  $\tau \in \mathcal{T}_\alpha$  one has

$$\Gamma\tau - \tau \in \mathcal{T}_{<\alpha} \stackrel{\text{def}}{=} \bigoplus_{\beta < \alpha} \mathcal{T}_\beta. \quad (2.1)$$

The canonical example to keep in mind is the space  $\bar{\mathcal{T}} = \bigoplus_{n \in \mathbf{N}} \bar{\mathcal{T}}_n$  of abstract polynomials in finitely many indeterminates, with  $A = \mathbf{N}$  and  $\bar{\mathcal{T}}_n$  denoting the space of monomials that are homogeneous of degree  $n$ . In this case, a natural group of transformations  $\mathcal{G}$  acting on  $\bar{\mathcal{T}}$  is given by the group of translations, which does indeed satisfy (2.1).

**2.1. Specific regularity structure.** The regularity structure that is relevant for the analysis of (1.1) is built in the following way. First, we start with the regularity structure  $\bar{\mathcal{T}}$  given by all polynomials in  $d + 1$  indeterminates, let us call them  $X_0, \dots, X_d$ , which denote the time and space directions respectively. We do however endow these with the parabolic space-time scaling instead of the more usual Euclidean scaling so that each factor of the “time” variable  $X_0$  increases the homogeneity by 2. More precisely, one has  $\mathbf{1} \in \bar{\mathcal{T}}_0$ ,  $X_0 \in \bar{\mathcal{T}}_2$ ,  $X_i \in \bar{\mathcal{T}}_1$  for  $i \in \{1, \dots, d\}$ , etc.

We then introduce two additional symbols,  $\Xi$  and  $\mathcal{I}$ , which will be interpreted as an abstract representation of the driving noise  $\xi$  and of the operation of convolution with the heat kernel respectively. Fixing some (sufficiently small in the sequel) exponent  $\kappa > 0$ , we then *postulate* that  $\Xi$  has homogeneity  $|\Xi| = -\frac{d+2}{2} - \kappa$  and, if  $\tau$  is some formal expression with homogeneity  $|\tau| = \alpha$ , then  $\mathcal{I}(\tau)$  is a new formal expression with homogeneity  $|\mathcal{I}(\tau)| = \alpha + 2$ . We also *postulate* that  $\mathcal{I}(X^k) = 0$  for every multi-index  $k$ , which will make sense in view of (2.6) below. (Here, for  $k = (k_0, \dots, k_d)$ , we have used the shorthand  $X^k = X_0^{k_0} \dots X_d^{k_d}$ .) Furthermore, if  $\tau, \bar{\tau}$  are formal expressions with respective homogeneities  $\alpha, \bar{\alpha}$ , then  $\tau\bar{\tau} = \bar{\tau}\tau$  is postulated to be a new formal expression with homogeneity  $\alpha + \bar{\alpha}$ .

A few examples of formal expression with their respective homogeneities that can in principle be built in this way are given by

$$|\mathcal{I}(\Xi)^2| = 2 - d - 2\kappa, \quad |\mathcal{I}(\mathcal{I}(\Xi)^3)| = 5 - \frac{3d}{2} - 3\kappa, \quad |\Xi\mathcal{I}(\Xi)| = -d - 2\kappa. \quad (2.2)$$

In order to define our regularity structure  $\mathcal{T}$ , we do not keep all of these formal expressions, but only those that are actually useful for the abstract reformulation of (1.1). More precisely, we consider a collection  $\mathcal{U}$  of formal expressions which is the smallest collection containing  $\mathbf{1}$ ,  $X$ , and  $\mathcal{I}(\Xi)$ , and such that

$$\tau_1, \tau_2, \tau_3 \in \mathcal{U} \quad \Rightarrow \quad \mathcal{I}(\tau_1\tau_2\tau_3) \in \mathcal{U}.$$

Here and below we use  $X$  to denote the collection of all  $X_i$  for  $i \in \{0, 1, \dots, d\}$ . We then set

$$\mathcal{W} = \{\Xi\} \cup \{\tau_1\tau_2\tau_3 : \tau_i \in \mathcal{U}\},$$

and we define  $\mathcal{T}$  as the set of all linear combinations of elements in  $\mathcal{W}$ . (Note that since  $\mathbf{1} \in \mathcal{U}$ , one does in particular have  $\mathcal{U} \subset \mathcal{W}$ .) Naturally,  $\mathcal{T}_\alpha$  consists of those linear combinations that only involve elements in  $\mathcal{W}$  that are of homogeneity  $\alpha$ . Furthermore,  $\mathcal{W}$  is the previously announced set of canonical basis elements of  $\mathcal{T}$ . In particular,  $\mathcal{T}$  contains the first two formal expressions of (2.2), but not the last one. It follows furthermore from [Hai14c, Lem. 8.10] that, for every  $\alpha \in \mathbf{R}$ ,  $\mathcal{W}$  contains only finitely many elements of homogeneity less than  $\alpha$ , so that each  $\mathcal{T}_\alpha$  is finite-dimensional.

In order to simplify expressions later, we will use the following shorthand graphical notation for elements of  $\mathcal{W}$ . For  $\Xi$ , we simply draw a dot. The integration map is then represented by a downfacing line and the multiplication of symbols is obtained by joining them at the root. For example, we have

$$\mathcal{I}(\Xi) = \dagger, \quad \mathcal{I}(\Xi)^3 = \blacktriangledown, \quad \mathcal{I}(\Xi)\mathcal{I}(\mathcal{I}(\Xi)^3) = \Psi. \quad (2.3)$$

Symbols containing factors of  $X$  have no particular graphical representation.

**2.2. Structure group.** Let us now describe the structure group  $\mathcal{G}$  associated to the space  $\mathcal{T}$ . For this, we first introduce additional symbols  $\mathcal{J}_k(\tau)$  and denote by  $\mathcal{T}_+$ , the free commutative algebra generated by

$$\mathcal{W}_+ \stackrel{\text{def}}{=} \{X\} \cup \{\mathcal{J}_k(\tau) : \tau \in \mathcal{W}, |\tau| + 2 > |k|\}, \quad (2.4)$$

where  $k$  is an arbitrary  $(d+1)$ -dimensional multi-index and  $|k|$  denotes its ‘‘parabolic length’’, i.e.

$$|k| = 2k_0 + k_1 + \dots + k_d .$$

In other words,  $\mathcal{T}_+$  consists of all linear combinations of products of formal expressions in  $\mathcal{W}_+$ . We will view  $\mathcal{J}_k$  as a map from  $\mathcal{T}$  into  $\mathcal{T}_+$  by postulating that it acts linearly on  $\mathcal{T}$  and that  $\mathcal{J}_k(\tau) = 0$  for those formal expressions  $\tau$  for which  $|\tau| + 2 \leq |k|$ .

With this definition at hand, we construct a linear map  $\Delta: \mathcal{T} \rightarrow \mathcal{T} \otimes \mathcal{T}_+$  in a recursive way. In order to streamline notations, we shall write  $\tau^{(1)} \otimes \tau^{(2)}$  as a shorthand for  $\Delta\tau$ . (This is an abuse of notation, following Sweedler, since in general  $\Delta\tau$  is a linear combination of such terms.) We then define  $\Delta$  via the identities

$$\Delta \mathbf{1} = \mathbf{1} \otimes \mathbf{1}, \quad \Delta \Xi = \Xi \otimes \mathbf{1}, \quad \Delta X_i = X_i \otimes \mathbf{1} + \mathbf{1} \otimes X_i,$$

and then recursively by the following relations:

$$\begin{aligned} \Delta \tau \bar{\tau} &= \tau^{(1)} \bar{\tau}^{(1)} \otimes \tau^{(2)} \bar{\tau}^{(2)}, \\ \Delta \mathcal{I}(\tau) &= \mathcal{I}(\tau^{(1)}) \otimes \tau^{(2)} + \sum_{\ell, m} \frac{X^\ell}{\ell!} \otimes \frac{X^m}{m!} \mathcal{J}_{\ell+m}(\tau). \end{aligned}$$

For any linear functional  $f: \mathcal{T}_+ \rightarrow \mathbf{R}$ , we can now define in a natural way a map  $\Gamma_f: \mathcal{T} \rightarrow \mathcal{T}$  by

$$\Gamma_f \tau = (I \otimes f) \Delta \tau, \quad (2.5)$$

where  $I$  denotes the identity map. Let now  $\mathcal{G}_+$  denote the set of all such linear functionals which are multiplicative in the sense that  $f(\tau \bar{\tau}) = f(\tau) f(\bar{\tau})$  for any two elements  $\tau, \bar{\tau} \in \mathcal{T}_+$ . With this definition at hand, we set

$$\mathcal{G} = \{\Gamma_f : f \in \mathcal{G}_+\} .$$

It is not difficult to see that these operators satisfy the property (2.1), but it is a non-trivial fact that the set  $\mathcal{G}$  of these linear operators does indeed form a group under composition, see [Hai14c, Sec. 8.1].

*Remark 2.1.* As a matter of fact, we will never need to consider the full space  $\mathcal{T}$  as defined above, but it will be sufficient to consider the subspace generated by all elements of homogeneity less than some large enough number  $\zeta$ . In practice, it actually turns out to be sufficient to choose  $\zeta = 2$ .

*Remark 2.2.* The construction we just explained depends on the dimension  $d$ , so we should really use the notations  $\mathcal{T}^{(d)}$ ,  $\mathcal{T}_+^{(d)}$ , etc. In order to keep the notations simple, we refrain from doing so. Our statements will always either hold for both  $d = 2, 3$ , or if they do not, then the value of  $d$  for which they hold will be made clear.

**2.3. Models.** Now that we have fixed our algebraic regularity structure  $(\mathcal{T}, \mathcal{G})$ , we introduce a family of analytical objects associated to it that will play the role of Taylor polynomials in our theory in order to allow us to describe solutions to (1.1) locally, up to arbitrarily high order, despite the fact that they are not smooth in the conventional sense.

From now on, we also fix a value  $\zeta \geq 2$  as in Remark 2.1 and we set  $\mathcal{T} = \mathcal{T}_{<\zeta}$ . This has the advantage that  $\mathcal{T}$  is finite-dimensional so we do not need to worry about topologies. We first fix a kernel  $K: \mathbf{R}^{d+1} \rightarrow \mathbf{R}$  with the following properties:

- (1) The kernel  $K$  is compactly supported in  $\{|x|^2 + |t| \leq 1\}$ .
- (2) One has  $K(t, x) = 0$  for  $t \leq 0$  and  $K(t, -x) = K(t, x)$ .
- (3) For  $(t, x)$  with  $|x|^2 + t < 1/2$  and  $t > 0$ , one has

$$K(t, x) = \frac{1}{|4\pi t|^{d/2}} e^{-\frac{|x|^2}{4t}},$$

and  $K$  is smooth on  $\{|x|^2 + |t| \geq 1/4\}$ .

- (4) For every polynomial  $P: \mathbf{R}^{d+1} \rightarrow \mathbf{R}$  of parabolic degree less than  $\zeta$ , one has

$$\int_{\mathbf{R}^{d+1}} K(t, x) P(t, x) dx dt = 0; \quad (2.6)$$

in other words,  $K$  has essentially all the properties of the heat kernel, except that it is furthermore compactly supported and satisfies (2.6). The constants  $1/2$  and  $1/4$  appearing in the third point are of course completely arbitrary as long as they are strictly between 0 and 1. The existence of a kernel  $K$  satisfying these properties is very easy to show.

We now denote by  $\mathcal{S}'$  the space of distributions on  $\mathbf{R}^{d+1}$  and by  $\mathcal{L}(\mathcal{T}, \mathcal{S}')$  the space of (necessarily continuous) linear maps from  $\mathcal{T}$  to  $\mathcal{S}'$ . Furthermore, given a continuous test function  $\varphi: \mathbf{R}^{d+1} \rightarrow \mathbf{R}$  and a point  $z = (t, x) \in \mathbf{R}^{d+1}$ , we set

$$\varphi_z^\lambda(\bar{z}) = \lambda^{-(d+2)} \varphi(\lambda^{-2}(\bar{t} - t), \lambda^{-1}(\bar{x} - x)), \quad (2.7)$$

where we also used the shorthand  $\bar{z} = (\bar{t}, \bar{x})$ . Finally, we write  $\mathcal{B}$  for the set of functions  $\varphi: \mathbf{R}^{d+1} \rightarrow \mathbf{R}$  that are smooth, compactly supported in the ball of radius one, and with their values and derivatives up to order 3 bounded by 1.

Given a kernel  $K$  as above, we then introduce a set  $\mathcal{M}$  of *admissible models* which will play the role of Taylor polynomials for our solution theory. An admissible model consists of a pair  $(\Pi, F)$  of functions

$$\begin{aligned} \Pi: \mathbf{R}^{d+1} &\rightarrow \mathcal{L}(\mathcal{T}, \mathcal{S}') & F: \mathbf{R}^{d+1} &\rightarrow \mathcal{G} \\ z &\mapsto \Pi_z & z &\mapsto F_z \end{aligned}$$

with the following properties. First, writing  $\gamma_{z\bar{z}} \in \mathcal{G}_+$  for the element such that  $F_z^{-1} \circ F_{\bar{z}} = \Gamma_{\gamma_{z\bar{z}}}$ , they satisfy the analytical bounds

$$|(\Pi_z \tau)(\varphi_z^\lambda)| \lesssim \lambda^{|\tau|}, \quad |\gamma_{z\bar{z}}(\bar{\tau})| \lesssim |z - \bar{z}|^{|\bar{\tau}|}, \quad (2.8)$$

uniformly over  $\varphi \in \mathcal{B}$ ,  $\lambda \in (0, 1]$ ,  $\tau \in \mathcal{W}$ , and  $\bar{\tau} \in \mathcal{W}_+$ . Here, with the same shorthand as before, we set

$$|z - \bar{z}| = |x - \bar{x}| + \sqrt{|t - \bar{t}|}.$$

The proportionality constants implicit in the notation  $\lesssim$  of (2.8) are assumed to be bounded uniformly for  $z$  and  $\bar{z}$  taking values in any compact set. We also assume

that one has the algebraic identity

$$\Pi_z F_z^{-1} = \Pi_{\bar{z}} F_{\bar{z}}^{-1}, \quad (2.9)$$

valid for every  $z, \bar{z}$  in  $\mathbf{R}^{d+1}$ . Finally, and this is why our models are called *admissible*, we assume that  $(\Pi_z \mathbf{1})(\bar{z}) = 1$ , that for every multi-index  $k$  and every  $\tau \in \mathcal{W}$  with  $X^k \tau \in \mathcal{T}$ ,

$$(\Pi_z X^k \tau)(\cdot) = (\cdot - z)^k \Pi_z \tau, \quad f_z(X^k) = (-z)^k, \quad (2.10a)$$

and that, for every  $\tau \in \mathcal{W}$  with  $\mathcal{I}\tau \in \mathcal{T}$  (recall that we did truncate  $\mathcal{T}$  so this is not true for *all*  $\tau$ ), one has the identities

$$f_z(\mathcal{J}_k \tau) = - \int_{\mathbf{R}^{d+1}} D^k K(\bar{z} - z) (\Pi_z \tau)(d\bar{z}), \quad |k| < |\tau| + 2, \quad (2.10b)$$

$$(\Pi_z \mathcal{I}\tau)(\bar{z}) = \int_{\mathbf{R}^{d+1}} K(\bar{z} - \bar{\bar{z}}) (\Pi_z \tau)(d\bar{\bar{z}}) + \sum_k \frac{(\bar{z} - z)^k}{k!} f_z(\mathcal{J}_k \tau). \quad (2.10c)$$

Here, we wrote similarly to above  $f_z \in \mathcal{G}_+$  for the element such that  $F_z = \Gamma_{f_z}$ . Recall that we have set  $\mathcal{J}_k \tau = 0$  if  $|k| \geq |\tau| + 2$ , so that the sum appearing on the second line is automatically always finite. It is not clear in principle that these integrals converge, but it turns out that the analytic conditions (2.8) guarantee that this is always the case, see [Hai14c, Sec. 5].

*Remark 2.3.* Since  $f_z \in \mathcal{G}_+$ , so that it is multiplicative, (2.10a) and (2.10b) do specify  $f_z$  (and therefore  $F_z$  via (2.5)) completely, once we know  $\Pi_z$ . There is therefore quite a lot of rigidity in these definitions, which makes the mere existence of admissible models a highly non-trivial fact.

*Remark 2.4.* Building further on Remark 2.3, it actually turns out that if the map  $\Pi: \mathbf{R}^{d+1} \rightarrow \mathcal{L}(\mathcal{T}, \mathcal{S}')$  satisfies the *first* analytical bound in (2.8) and is such that, for  $F$  defined from  $\Pi$  via (2.10), one has the identities (2.10) and (2.9), then the second analytical bound in (2.8) is *automatically* satisfied. This is a consequence of [Hai14c, Thm. 5.14].

Since we will only ever consider (1.1) with periodic boundary conditions, we will always assume that the model  $(\Pi, F)$  is periodic in the following sense. We are given  $d$  vectors  $e_1, \dots, e_d \in \mathbf{R}^d$  and we denote by  $T_i: \mathbf{R}^{d+1} \rightarrow \mathbf{R}^{d+1}$  the translation maps

$$T_i: (t, x) \mapsto (t, x + e_i).$$

We also define the natural action of  $T_i$  on test functions by  $(T_i \varphi)(z) = \varphi(T_i^{-1} z)$ . We then say that the model  $(\Pi, F)$  is periodic if

$$(\Pi_{T_i z} \tau)(T_i \varphi) = (\Pi_z \tau)(\varphi), \quad F_{T_i z} = F_z, \quad (2.11)$$

for every  $z \in \mathbf{R}^{d+1}$ , every smooth test function  $\varphi$ , every  $\tau$ , and every  $i$ . This definition is consistent in the sense that if  $\Pi$  satisfies the first identity in (2.11) and  $F$  is given by (2.10b), then  $F$  automatically satisfies the second identity.

Given two admissible models  $(\Pi, F)$  and  $(\bar{\Pi}, \bar{F})$ , the bounds (2.8) yield a natural notion of a semi-distance between them by considering, for any given compact domain  $D \subset \mathbf{R}^{d+1}$ , the quantity

$$\|\Pi - \bar{\Pi}\| = \sup_{z \in D} \sup_{\substack{\varphi \in \mathcal{B} \\ \lambda \in (0,1)}} \sup_{\tau \in \mathcal{W}} \lambda^{-|\tau|} |(\Pi_z \tau - \bar{\Pi}_z \tau)(\varphi_z^\lambda)|. \quad (2.12)$$



While this seminorm does of course depend on the domain  $D$ , we will always take  $D = I \times \mathbf{R}^d$  for some large enough interval  $I$ . The precise length of  $I$  does not really matter as long as it contains  $[-2, T + 2]$ , where  $T$  is the final time for which we wish to obtain our large deviations principle. We will use the following definition.

**Definition 2.5.** *The space  $\mathcal{M}^{(d)}$  consists of all periodic admissible models for the regularity structure  $(\mathcal{T}^{(d)}, \mathcal{G})$  built above in dimension  $d \in \{2, 3\}$ . It is endowed with the metric (2.12) for the domain  $D = I \times \mathbf{R}^d$  with a sufficiently large interval  $I$ .*

*Remark 2.6.* As defined here, the space  $\mathcal{M}^{(d)}$  is not separable which might lead to technical difficulties. In practice, these can easily be resolved by defining  $\mathcal{M}^{(d)}$  as the closure of the set of *smooth* admissible models, i.e. the set of models for which  $\Pi: \mathbf{R}^{d+1} \rightarrow \mathcal{L}(\mathcal{T}_-^{(d)}, \mathcal{C}^\infty(\mathbf{R}^{d+1}))$  is a smooth function, under the distance (2.12), which is a separable space.

It is extremely important at this stage to note that  $\mathcal{M}^{(d)}$  is in general *much* smaller than the space of all (periodic in the sense of (2.11)) functions  $\Pi: \mathbf{R}^{d+1} \rightarrow \mathcal{L}(\mathcal{T}, \mathcal{S}')$  such that (2.12) is finite. This is because we still have the requirement that the identities (2.9) and (2.10) hold. It may also appear that the “norm” (2.12) neglects to control the second bound in (2.8) but, by Remark 2.4, this bound actually holds automatically on  $D' = I' \times \mathbf{R}^d$  if we know that  $\|\Pi\|$  is finite on  $D = I \times \mathbf{R}^d$  for  $I$  containing a large enough neighbourhood of  $I'$ . (Essentially  $D'$  needs to be large enough to contain the support of  $K(\cdot - z)$  for every  $z \in D$  so that one can apply (2.10b).)

Given any smooth space-time function  $\xi_\delta$  which is periodic in the sense that  $T_i \xi_\delta = \xi_\delta$ , there is a canonical way of lifting it to an admissible periodic model  $\Psi(\xi_\delta) = (\Pi^\delta, F^\delta)$  as follows. First, we set  $\Pi_z^\delta \Xi = \xi_\delta$ , independently of  $z$ , and we define it on  $X^k$  as in (2.10a). Then, we define  $\Pi_z^\delta$  recursively by (2.10c), as well as the additional identity

$$(\Pi_z^\delta \tau \bar{\tau})(\bar{z}) = (\Pi_z^\delta \tau)(\bar{z})(\Pi_z^\delta \bar{\tau})(\bar{z}). \quad (2.13)$$

Note that this is guaranteed to make sense if  $\xi_\delta$  is a function but not for more general distributions! It was shown in [Hai14c, Prop. 8.27] that if we furthermore define  $F^\delta$  via (2.10b), then this does indeed define an admissible model for every continuous function  $\xi_\delta$ . It is however very important to keep in mind that not every admissible model is obtained in this way, or even as a limit of such models. We will use this fact in Section 2.6 below.

Since an admissible model  $(\Pi, F)$  is interpreted as an extension of the usual Taylor polynomials, it is quite natural to define spaces  $\mathcal{D}^{\gamma, \eta}$  which mimic a weighted version of the Hölder spaces  $\mathcal{C}^\gamma$  in the following way.

**Definition 2.7.** *A function  $U: \mathbf{R}^{d+1} \rightarrow \mathcal{T}_{<\gamma}$  belongs to  $\mathcal{D}^{\gamma, \eta}$  if, for every compact domain  $D$ , one has*

$$\|U\|_{\gamma, \eta} \stackrel{\text{def}}{=} \sup_{z \in D} \sup_{\alpha < \gamma} \frac{\|U(z)\|_\alpha}{|t|^{(\eta - \alpha) \wedge 0}} + \sup_{\substack{z \neq \bar{z} \in D \\ |z - \bar{z}| \leq 1}} \sup_{\alpha < \gamma} \frac{\|U(z) - \Gamma_{z\bar{z}} U(\bar{z})\|_\alpha}{(|t| \wedge |\bar{t}|)^{\eta - \gamma} |z - \bar{z}|^{\gamma - \alpha}} < \infty. \quad (2.14)$$

Here, we wrote  $\|\tau\|_\alpha$  for the norm of the component of  $\tau$  in  $\mathcal{T}_\alpha$  and we used as before the notation  $\Gamma_{z\bar{z}} = F_z^{-1} F_{\bar{z}}$ . We also used  $t$  and  $\bar{t}$  as shorthands for the time components of the space-time points  $z$  and  $\bar{z}$ .

*Remark 2.8.* The powers of  $t$  appearing in this definition allow elements of  $\mathcal{D}^{\gamma,\eta}$  to exhibit a singularity on the line  $\{(t, x) : t = 0\}$ . This is essential in order to be able to deal with solutions to (1.1) with distributional initial conditions.

*Remark 2.9.* In order to streamline notations, we suppressed the dependence on the domain  $D$  in this norm. This is because, similarly to our definition of admissible models, we will only ever use this on some fixed space-time domain.

The space  $\mathcal{D}^{\gamma,\eta}$  depends in a crucial way on the underlying model  $(\Pi, F)$ . Therefore, it is not obvious *a priori* how to compare elements belonging to  $\mathcal{D}^{\gamma,\eta}$ , but based on two different models. This is however crucial when investigating the continuity of solutions to (1.1) as a function of the underlying model. A natural distance between elements  $U \in \mathcal{D}^{\gamma,\eta}$  and  $\bar{U} \in \bar{\mathcal{D}}^{\gamma,\eta}$  (denoting by  $\bar{\mathcal{D}}^{\gamma,\eta}$  the space built over the model  $(\bar{\Pi}, \bar{F})$ ), is given by (2.14), with  $U(z)$  replaced by  $U(z) - \bar{U}(z)$  in the first term and  $U(z) - \Gamma_{zz}U(\bar{z})$  replaced by

$$U(z) - \bar{U}(z) - \Gamma_{zz}U(\bar{z}) + \bar{\Gamma}_{zz}\bar{U}(\bar{z})$$

in the second term. Note that this distance is *not* a function of  $U - \bar{U}$ !

The idea now is to reformulate (1.1) as a fixed point problem in  $\mathcal{D}^{\gamma,\eta}$  (based on the canonical model  $\Psi(\xi_\delta)$  built above) for suitable values of the exponents  $\gamma$  and  $\eta$ . As a matter of fact, we will view it as a fixed point problem in the subspace  $\mathcal{D}_U^{\gamma,\eta} \subset \mathcal{D}^{\gamma,\eta}$  consisting of those functions taking values in the linear span of  $\mathcal{U}$ . If we consider an admissible model consisting of smooth functions, then one can define an operator  $\mathcal{R}$  acting on  $\mathcal{D}^{\gamma,\eta}$  and taking values in the space of continuous functions by

$$(\mathcal{R}U)(z) = (\Pi_z U(z))(z). \quad (2.15)$$

A remarkable fact, and this is the content of [Hai14c, Thm 3.10], is that as soon as  $\gamma > 0$  the map

$$(\Pi, U) \mapsto \mathcal{R}U \in \mathcal{S}' \quad (2.16)$$

given by (2.15) is jointly (locally) Lipschitz continuous with respect to the metric defined in (2.14) and (2.12), so that it makes sense even in situations where the definition (2.15) is nonsensical! This of course relies very heavily on the fact that we *only* consider admissible models in (2.16) and not arbitrary functions  $\Pi: \mathbf{R}^{d+1} \rightarrow \mathcal{L}(\mathcal{T}, \mathcal{S}')$ . The map  $\mathcal{R}$  is called the *reconstruction operator* since it reconstructs the (global) distribution  $\mathcal{R}U$  from the (local) data  $U$  and  $\Pi$ .

**2.4. Continuity properties.** As we will see below, it turns out that the solution to (1.1) can be obtained, via a suitable fixed point argument, as a *jointly continuous* map of the initial condition and a suitable random model  $(\Pi, F)$  built from the underlying space-time white noise. As a consequence of the contraction principle, it will therefore be sufficient to obtain a suitable large deviations principle for the random model. For this, we want to reduce ourselves to the “minimal” amount of information necessary to reconstruct the whole model. Denote by  $\mathcal{W}_-^{(d)} \subset \mathcal{W}$  the list of symbols of negative order in dimension  $d$ , which do not contain any factor  $X_i$ . In dimension  $d = 2$  these are given by

$$\mathcal{W}_-^{(2)} = \{\Xi, \uparrow, \vee, \Psi\},$$

while in dimension  $d = 3$ , one has

$$\mathcal{W}_-^{(3)} = \{\Xi, \uparrow, \vee, \Psi, \Psi, \Psi, \Psi\}. \quad (2.17)$$

(Here and in the sequel, we assume that  $\kappa$  is small enough. Otherwise, these lists could expand.) We also denote by  $\mathcal{T}_-^{(d)} \subset \mathcal{T}$  the linear span of  $\mathcal{W}_-^{(d)}$ . We now define the space  $\mathcal{M}_-^{(d)}$  (with  $d = 2, 3$ ) of “minimal admissible models” in the following way.

It is natural to take for  $\mathcal{M}_-^{(d)}$  the closure of the set of smooth maps  $\Pi: \mathbf{R}^{d+1} \rightarrow \mathcal{L}(\mathcal{T}_-^{(d)}, \mathcal{C}^\infty(\mathbf{R}^{d+1}))$  that are periodic in the sense given previously, under the norm given by (2.18) below. We furthermore impose that the constraint (2.9) holds with  $F_z$  given by (2.10b). Besides providing a bound of the type (2.8), we also want to ensure that  $\Pi_z \uparrow$  is not just a space-time distribution of order  $|\uparrow|$ , but actually a continuous function of time with values in a space of distributions of order  $|\uparrow|$ . In order to formulate this, we denote by  $\mathcal{B}_0$  a set of test functions just like  $\mathcal{B}$ , but depending only on the spatial variable. Their translated and rescaled versions  $\varphi_x^\lambda$  are defined analogously to (2.7). We then assume that  $\Pi$  is such that

$$\begin{aligned} \sup_{\lambda \in (0,1]} \sup_{\varphi \in \mathcal{B}} \sup_{\tau \in \mathcal{W}_-^{(d)}} \sup_{z \in D} \lambda^{-|\tau|} |\langle \Pi_z \tau, \varphi_z^\lambda \rangle| &< \infty, \\ \sup_{\lambda \in (0,1]} \sup_{\varphi \in \mathcal{B}} \sup_{s \in \mathbf{R}} \sup_{z \in D} \lambda^{-|\Xi|} |\langle \mathbf{1}_{t \geq s} \Pi_z \Xi, \varphi_z^\lambda \rangle| &< \infty, \\ \sup_{\lambda \in (0,1]} \sup_{\varphi \in \mathcal{B}_0} \sup_{z \in D} \lambda^{-|\uparrow|} |\langle (\Pi_z \uparrow)(t, \cdot), \varphi_x^\lambda \rangle| &< \infty, \end{aligned} \quad (2.18)$$

where  $\langle \cdot, \cdot \rangle$  denotes the usual  $L^2$  inner product and  $D = I \times \mathbf{T}^2$  for some fixed but large enough interval  $I$ . Furthermore, we impose that the identities (2.9) and (2.10) hold. In order for this definition to make sense, we need to make sure that we have enough data to define the linear maps  $F_z$  on  $\mathcal{T}_-^{(d)}$ . In dimension  $d = 2$ , it is straightforward to verify that  $\Delta \tau = \tau \otimes \mathbf{1}$  for every  $\tau \in \mathcal{W}_-^{(2)}$ , so that the action of  $\mathcal{G}$  on  $\mathcal{T}_-^{(2)}$  is trivial. The condition (2.9) then reduces to stating that  $\Pi_z$  does actually not depend on  $z$  at all. In particular, the space  $\mathcal{M}_-^{(2)}$  is a separable Banach space.

In dimension  $d = 3$  and for  $\kappa$  sufficiently small, we note that one has the identities

$$\begin{aligned} \Delta \mathfrak{V} &= \mathfrak{V} \otimes \mathbf{1} + \mathfrak{v} \otimes \mathcal{J}(\mathfrak{v}), & \Delta \mathfrak{V} &= \mathfrak{V} \otimes \mathbf{1} + \uparrow \otimes \mathcal{J}(\mathfrak{v}), \\ \Delta \mathfrak{V} &= \mathfrak{V} \otimes \mathbf{1} + \mathfrak{v} \otimes \mathcal{J}(\mathfrak{v}), \end{aligned}$$

so that the action of  $F_z$  on  $\mathcal{T}_-^{(3)}$  only requires (2.10b) with  $\tau \in \{\mathfrak{v}, \mathfrak{v}\}$ . This data is available from  $\Pi_z \mathfrak{v}$  and  $\Pi_z \mathfrak{v}$  (which actually do not depend on  $z$ ), so that the first condition in (2.18) is indeed meaningful. Note that, unlike  $\mathcal{M}_-^{(2)}$ ,  $\mathcal{M}_-^{(3)}$  is *not* a linear space. It is however a separable metric space and can be viewed as a closed subset of the separable Banach space obtained by relaxing (2.9) to the requirement that  $\Pi_z \tau = \Pi_{\bar{z}} \tau$  for  $\tau \in \mathcal{W}_-^{(2)}$ .

The reason why the spaces  $\mathcal{M}_-^{(d)}$  are important is the following result.

**Theorem 2.10.** *For every  $\Pi \in \mathcal{M}_-^{(d)}$  there exists a unique admissible model  $(\Pi, F) \in \mathcal{M}_-^{(d)}$  such that  $\Pi_z \tau = \Pi_{\bar{z}} \tau$  for all  $\tau \in \mathcal{W}_-^{(d)}$  and  $z \in \mathbf{R}^{d+1}$ . Furthermore, the map  $\Pi \mapsto (\Pi, F)$  is locally Lipschitz continuous.*

*Proof.* This is an immediate consequence of Proposition 3.31 and Theorem 5.14 of [Hai14c].  $\square$

**2.5. Abstract fixed point problem.** We now reformulate (1.1) as a fixed point problem in  $\mathcal{D}_U^{\gamma, \eta}$  for suitable values of  $\gamma$  and  $\eta$ . Note first that by Duhamel’s

formula, (1.1) is equivalent for smooth  $\xi$  to the integral equation

$$u = P \star ((\xi + Cu - u^3)\mathbf{1}_{t>0}) + Pu_0. \quad (2.19)$$

Here,  $P$  denotes the heat kernel,  $\star$  denotes space-time convolution, and  $Pu_0$  denotes the solution to the heat equation with initial condition  $u_0$ . A local solution is a pair  $(u, \tau)$  with  $\tau > 0$  and such that (2.19) holds on  $[0, \tau] \times \mathbf{R}^d$ . In order to interpret this equation as an identity in  $\mathcal{D}^{\gamma, \eta}$ , we need to find an analogue for the operation of convolution against  $P$ .

It turns out that, provided that  $\gamma < \bar{\gamma} + 2$ , that  $\eta < |\uparrow| = 1 - \frac{d}{2} - \kappa$  and that  $\bar{\eta} > -2$ , it is possible to construct a linear operator  $\mathcal{P}: \mathcal{D}^{\bar{\gamma}, \bar{\eta}} \rightarrow \mathcal{D}_U^{\gamma, \eta}$  with the following properties:

- (1) One has the identity  $\mathcal{R}\mathcal{P}U = P \star \mathcal{R}U$ .
- (2) One has  $\mathcal{P}U = \mathcal{I}U + \tilde{\mathcal{P}}U$ , where  $\tilde{\mathcal{P}}U$  only takes values in  $\bar{\mathcal{T}}$ , the linear span of  $\{X^k\}$ .
- (3) There exists  $\theta > 0$  such that

$$\|\mathcal{P}\mathbf{1}_{t>0}U\|_{\gamma, \eta} \leq T^\theta \|U\|_{\bar{\gamma}, \bar{\eta}},$$

where the norms are taken over the domain  $[0, T] \times \mathbf{R}$ .

Furthermore, even though  $\mathbf{1}_{t>0}\Xi \notin \mathcal{D}^{\bar{\gamma}, \bar{\eta}}$  with  $\bar{\gamma}$  and  $\bar{\eta}$  as above, it turns out that if the second bound in (2.18) is satisfied (in particular if the model comes from an element of  $\mathcal{M}_-^{(d)}$  as in Theorem 2.10), then there exists an element  $\mathcal{P}\mathbf{1}_{t>0}\Xi \in \mathcal{D}_U^{\gamma, \eta}$  satisfying all of the properties (1)–(3). For the precise construction of  $\mathcal{P}$  and a proof of these properties, see Section 5, Proposition 6.16, and Theorem 7.1 in [Hai14c].

Finally, given a function  $u$  of (parabolic) class  $\mathcal{C}^\gamma$ , we write  $\mathcal{T}_\gamma u$  for its Taylor expansion of (parabolic) order  $\gamma$ , namely

$$(\mathcal{T}_\gamma u)(z) = \sum_{|k| < \gamma} \frac{X^k}{k!} (D^k u)(z) \in \bar{\mathcal{T}} \subset \mathcal{T}.$$

It was then shown in [Hai14c, Lemma 7.5] that one has  $\mathcal{T}_\gamma Pu_0 \in \mathcal{D}_U^{\gamma, \eta}$ , provided that  $u_0 \in \mathcal{C}^\alpha(\mathbf{T}^d)$  for some  $\alpha > \eta$ . Note that this is true for arbitrary  $\eta \in \mathbf{R}$ , in particular one can have  $\eta < 0$ .

With all of these notations at hand, we can lift (2.19) in a very natural way to a fixed point problem in  $\mathcal{D}_U^{\gamma, \eta}$ , by looking for solutions  $U$  to

$$U = \mathcal{P}\mathbf{1}_{t>0}(\Xi + CU - U^3) + \mathcal{T}_\gamma Pu_0. \quad (2.20)$$

According to [Hai14c, Prop. 6.12] for  $U \in \mathcal{D}_U^{\gamma, \eta}$  we can actually define  $U^3$  as a modelled distribution. Similarly to (2.19), we call a pair  $(U, \tau)$  with  $\tau > 0$  a local solution if the identity (2.20) holds on  $[0, \tau] \times \mathbf{R}^d$ .

*Remark 2.11.* For (2.20) to make sense, we need a suitable definition for the element  $\mathcal{P}\mathbf{1}_{t>0}\Xi$ . As already mentioned, this is the case if the underlying model comes from an element of  $\mathcal{M}_-^{(d)}$ .

We will say that local solutions are unique if, whenever  $(U, \tau)$  and  $(\bar{U}, \bar{\tau})$  are local solutions, one has  $U(t, x) = \bar{U}(t, x)$  for  $t \in (0, \tau \wedge \bar{\tau}]$ . Note also that if  $U \in \mathcal{D}_U^{\gamma, \eta}$  with  $\gamma > 0$  and the underlying model comes from an element of  $\mathcal{M}_-^{(d)}$ , then  $\mathcal{R}U$  is a continuous function of time with values in the space  $\mathcal{C}^\beta = \mathcal{C}^\beta(\mathbf{T}^d)$  of periodic distributions in space of regularity  $\beta = |\uparrow| = 1 - \frac{d}{2} - \kappa$  (see for example [Hai14c, Def. 3.7] for a definition of  $\mathcal{C}^\beta$  for  $\beta < 0$ ). On a bounded domain, it agrees

with the Besov space  $\mathcal{B}_{\infty, \infty}^{\beta}$ ). Given a final time  $T > 0$ , we say that a local solution  $(U, \tau)$  to (2.20) is *maximal* if either  $\tau = T$  or  $\lim_{t \rightarrow \tau} \|(\mathcal{R}U)(t, \cdot)\|_{\beta} = \infty$ . Given a final time  $T$ , it will be convenient to view solutions as elements of the space  $\mathcal{X} = \mathcal{C}([0, T], \mathcal{C}^{\eta}) \cup \{\infty\}$  for some  $\eta < \beta$ , which is a metric space by postulating that the ‘‘cemetery state’’  $\infty$  is at distance 1 of all other elements.

One then has the following result, where we implicitly assume that  $\kappa$  is sufficiently small ( $\kappa < 1/10$  is sufficient for example).

**Theorem 2.12.** *Let  $\gamma \in (1 + 4\kappa, \zeta)$ , and let  $\eta \in (-\frac{2}{3}, 1 - \frac{d}{2} - \kappa)$ . Then, for every admissible model  $(\Pi, F)$  arising from an element of  $\mathcal{M}_{-}^{(d)}$ , (2.20) admits a unique maximal solution in  $\mathcal{D}_{\mathcal{U}}^{\gamma, \eta}$ .*

*Furthermore, if there exists a smooth function  $\xi_{\delta}$  such that  $(\Pi, F) = \Psi(\xi_{\delta})$ , then this solution is global and the function  $u = \mathcal{R}U$  solves (2.19). Finally, define the solution map  $\mathcal{S}_A: \mathbf{R} \times \mathcal{C}^{\eta} \times \mathcal{M}_{-}^{(d)} \rightarrow \mathcal{X}$  which maps  $(C, u_0, \Pi)$  onto  $\mathcal{R}U$  if the maximal solution  $U$  to (2.20) satisfies  $\tau > T$  and to  $\infty$  otherwise. Then, for every  $Z = (C, u_0, \Pi)$  such that  $\mathcal{S}_A(Z) \neq \infty$ , there exists a neighbourhood  $\mathcal{O}$  of  $Z$  such that  $\mathcal{S}_A$  is Lipschitz continuous from  $\mathcal{O}$  to  $\mathcal{X}$ .*

*Proof.* This is a synthesis of the results of [Hai14c, Sec. 7] and [Hai14c, Sec. 9], see also [Hai14b] for an informal overview of the proof. The continuous dependence on  $C$  is not stated explicitly there, but is implicit in the proofs.

The only fact which remains to be shown is that the map  $(\Pi, F) \mapsto \mathcal{R}U$  is (locally) continuous with values in  $\mathcal{X}$ . Indeed, the reconstruction operator  $\mathcal{R}$  only maps the space  $\mathcal{D}_{\mathcal{U}}^{\gamma, \eta}$  into  $\mathcal{C}^{\eta}([0, T] \times \mathbf{T}^d)$  in general [Hai14c, Prop. 6.9], which is strictly larger than  $\mathcal{X}$ . However, one does get continuity of  $\mathcal{S}_A$  into  $\mathcal{X}$ . The reason why this is true is that by (2.19) the solution  $\mathcal{R}U$  is of the form

$$\mathcal{R}U = P \star (\xi + C\mathcal{R}U - \mathcal{R}(U^3))\mathbf{1}_{t>0} + Pu_0,$$

with  $\mathcal{R}$  the reconstruction operator. Since  $U$  takes values in  $\mathcal{D}_{\mathcal{U}}^{\gamma, \eta}$ , it follows from [Hai14c, Prop. 6.12] that  $U^3$  takes values in  $\mathcal{D}^{\gamma', 3\eta}$  for some  $\gamma' > 0$ . By [Hai14c, Prop. 6.9], one therefore obtains that  $\mathcal{R}(U^3) \in \mathcal{C}^{3(1-\frac{d}{2}-\kappa)}([0, T] \times \mathbf{T}^d)$ , which in particular shows by usual (parabolic) Schauder estimates that  $P \star \mathcal{R}(U^3)$  takes values in  $\mathcal{C}([0, T] \times \mathbf{T}^d) \subset \mathcal{X}$ . Combining this with the fact that both  $Pu_0$  and  $P \star \xi$  belong to  $\mathcal{X}$  (the latter as a consequence of the second and third bound in (2.18)) concludes the proof.  $\square$

**2.6. Renormalised model.** It turns out that, if we take for  $\xi_{\delta}$  some mollification of space-time white noise then the sequence  $\Psi(\xi_{\delta})$ , with the canonical lift  $\Psi$  defined via (2.13), does *not* converge in  $\mathcal{M}^{(d)}$  (and a fortiori not in  $\mathcal{M}_{-}^{(d)}$ ). Instead, we fix two renormalisation constants  $C_{\delta}^{(1)}$  and  $C_{\delta}^{(2)}$  and we define a ‘‘renormalised lift’’  $\hat{\Psi}_{\delta}$  in the following way. Regarding  $\hat{\Pi}_z^{\delta} \dagger$  (which actually doesn’t depend on  $z$ ), we define it from  $\hat{\Pi}_z^{\delta} \Xi = \xi_{\delta}$  via (2.10c).

In dimension  $d = 2$ , we then replace (2.13) by the identities

$$\hat{\Pi}_z^{\delta} \heartsuit = (\hat{\Pi}_z^{\delta} \dagger)^2 - C_{\delta}^{(1)}, \quad \hat{\Pi}_z^{\delta} \spadesuit = (\hat{\Pi}_z^{\delta} \dagger)^3 - 3C_{\delta}^{(1)} \hat{\Pi}_z^{\delta} \dagger. \quad (2.21)$$

We then use Theorem 2.10 to extend this to all of  $\mathcal{T}^{(2)}$ .

In dimension  $d = 3$ , we also define  $\hat{\Pi}_z^{\delta}$  on  $\heartsuit$  and  $\spadesuit$  as in (2.21), but we furthermore set

$$\hat{\Pi}_z^{\delta} \clubsuit = (\hat{\Pi}_z^{\delta} \heartsuit)(\hat{\Pi}_z^{\delta} \heartsuit) - C_{\delta}^{(2)}, \quad \hat{\Pi}_z^{\delta} \blacklozenge = (\hat{\Pi}_z^{\delta} \spadesuit)(\hat{\Pi}_z^{\delta} \heartsuit) - 3C_{\delta}^{(2)} \hat{\Pi}_z^{\delta} \dagger,$$

$$\hat{\Pi}_z^\delta \Psi = (\hat{\Pi}_z^\delta \Psi) (\hat{\Pi}_z^\delta \uparrow), \quad (2.22)$$

where  $\hat{\Pi}_z^\delta \Upsilon$  and  $\hat{\Pi}_z^\delta \Psi$  are defined from  $\hat{\Pi}_z^\delta \mathbf{V}$  and  $\hat{\Pi}_z^\delta \Psi$  through (2.10c). Again, we then use Theorem 2.10 to extend this to all of  $\mathcal{T}^{(3)}$ . Here, the non-trivial fact that  $\hat{\Pi}^\delta$  defined in this way indeed belongs to  $\mathcal{M}_-^{(3)}$  was shown in Theorem 8.44 and Section 9.2 of [Hai14c].

As usual, we define  $\hat{F}_z^\delta$  through (2.10b) and we write  $(\hat{\Pi}^\delta, \hat{F}^\delta) = \hat{\Psi}_\delta(\xi_\delta)$ . With these definitions, the following theorem is the main result of [Hai14c, Sec. 10].

**Theorem 2.13.** *Let  $\varrho: \mathbf{R}^{d+1} \rightarrow \mathbf{R}$  be a compactly supported smooth function with  $\int \varrho = 1$  and set*

$$\varrho_\delta(t, x) = \delta^{-d-2} \varrho(\delta^{-2}t, \delta^{-1}x), \quad \xi_\delta = \varrho_\delta \star \xi,$$

where  $\xi$  is periodic space-time white noise.

There exist choices of constants  $C_\delta^{(1)}$  and  $C_\delta^{(2)}$  such that the sequence  $\hat{\Psi}_\delta(\xi_\delta)$  converges in probability in  $\mathcal{M}_-^{(d)}$  to a limiting model  $\hat{\Pi}$ . Furthermore, the function  $u = \mathbf{R}U$ , where  $U$  is the (local) solution to (2.20) with model  $\hat{\Pi}^\delta$  is the classical solution to

$$\partial_t u = \Delta u + (3C_\delta^{(1)} - 9C_\delta^{(2)})u - u^3 + \xi_\delta.$$

*Remark 2.14.* In dimension 2, we do not need  $C_\delta^{(2)}$  (i.e. we set it equal to 0), and  $C_\delta^{(1)}$  can be chosen as

$$C_\delta^{(1)} = \frac{1}{4\pi} \log \delta^{-1} + c_\varrho, \quad (2.23)$$

for some constant  $c_\varrho$  of order 1 which depends on the mollifier  $\varrho$ . The precise expression for  $c_\varrho$  does not matter at this stage.

In dimension 3 on the other hand, one should take

$$C_\delta^{(1)} = \delta^{-1} \int_{\mathbf{R}^4} (P \star \varrho)(t, x)^2 dt dx, \quad (2.24)$$

where  $\star$  denotes space-time convolution as before and  $P$  denotes the heat kernel (without any truncation). This time, it is the constant  $C_\delta^{(2)}$  that has to be chosen to diverge logarithmically:

$$C_\delta^{(2)} = c \log \delta^{-1} + \bar{c}_\varrho, \quad (2.25)$$

with  $c$  a universal constant independent of  $\varrho$  and  $\bar{c}_\varrho$  some constant that depends on the mollifier. The precise values for  $c$  and  $\bar{c}_\varrho$  are irrelevant for our current purpose.

### 3. LARGE DEVIATIONS FOR WIENER CHAOS

In this section we state and prove a large deviation statement for random variables taking values in a fixed inhomogeneous Wiener chaos over an arbitrary Gaussian probability space. These results are then applied in Section 4 to obtain a large deviation result for the renormalised models constructed from a white noise with small amplitude.

Large deviation results for random variables in a fixed Wiener chaos have been studied by several authors, see e.g. [Bor78, Bor84b, Led90, MWNPA92]. For the reader's convenience, we give a complete exposition in our context. The core of our argument is a *generalised contraction principle* in the spirit of [DS89, Lemma 2.1.4]. We start by recalling the following definition from [DS89].

**Definition 3.1.** Let  $(S, d)$  be a separable metric space. We say that a family  $\{\mu_\varepsilon, \varepsilon > 0\}$  of probability measures on  $S$  satisfies a large deviation principle with rate  $\varepsilon$  and rate function  $\mathcal{I}$  if

- (1)  $\mathcal{I} : S \rightarrow [0, +\infty]$  is lower semicontinuous, has compact sublevel sets and is not identical to  $+\infty$ .
- (2) For every closed set  $\mathcal{C} \subseteq S$  we have

$$\limsup_{\varepsilon \rightarrow 0} \varepsilon \log \mu_\varepsilon(\mathcal{C}) \leq - \inf_{s \in \mathcal{C}} \mathcal{I}(s) . \quad (3.1)$$

- (3) For every open subset  $\mathcal{O} \subseteq S$  we have

$$\liminf_{\varepsilon \rightarrow 0} \varepsilon \log \mu_\varepsilon(\mathcal{O}) \geq - \inf_{s \in \mathcal{O}} \mathcal{I}(s) . \quad (3.2)$$

We say that a family of  $S$ -valued random variables satisfies a large deviation principle if their distributions do.

*Remark 3.2.* All of the large deviation principles in this paper have rate  $\varepsilon$  and we will not repeat the rate every time.

We also recall the following version of the contraction principle:

**Lemma 3.3.** Let  $\{\mu_\varepsilon, \varepsilon > 0\}$  be a family of probability measures on a separable metric space  $(S, d)$  and let  $\mathcal{I} : S \rightarrow [0, \infty]$  satisfy the first condition of Definition 3.1. Let  $(S', d')$  be another separable metric space, and let  $\{\Psi_\varepsilon, \varepsilon \geq 0\}$  be a family of mappings from  $S$  to  $S'$  which are continuous on a neighbourhood of  $\{s \in S : \mathcal{I}(s) < \infty\}$ . We assume that

- (1) The probability measures  $\{\mu_\varepsilon, \varepsilon > 0\}$  satisfy a large deviation principle on  $S$  with rate function  $\mathcal{I}$ .
- (2) For every  $c \in \mathbf{R}$ , there exists a neighbourhood  $\mathcal{O}_c$  of  $\{s \in S : \mathcal{I}(s) \leq c\}$  such that the mappings  $\Psi_\varepsilon$  converge uniformly on  $\mathcal{O}_c$  to  $\Psi_0$ .

Then the image measures  $\mu_\varepsilon^{-1} \circ \Psi_\varepsilon^{-1}$  satisfy a large deviation principle on  $S'$  with rate function

$$\mathcal{I}'(s') = \inf\{\mathcal{I}(s) : s \in S, \Psi_0(s) = s'\} ,$$

with the convention that the inf equals  $+\infty$  if the set is empty.

*Proof.* This is a direct consequence of [DS89, Exercice 2.1.20 (ii)] (see also [DS89, Lemma 2.1.4]). Indeed, in this exercise it is shown that the conclusion of the theorem holds if the  $\Psi_\varepsilon$  are continuous on all of  $S$  and if for all  $\alpha > 0$  we have

$$\limsup_{\varepsilon \rightarrow 0} \varepsilon \log \mu_\varepsilon\left(\xi : d'(\Psi_\varepsilon(\xi), \Psi_0(\xi)) > \alpha\right) = -\infty .$$

Furthermore, given any  $c \in \mathbf{R}$ , the uniform convergence of  $\Psi_\varepsilon$  on  $\mathcal{O}_c$  implies that for every  $\alpha > 0$  one can find  $\varepsilon_0$  such that for  $\varepsilon < \varepsilon_0$  the condition  $d'(\Psi_\varepsilon(s), \Psi_0(s)) > \alpha$  implies  $s \notin \mathcal{O}_c$ . The LDP for  $\mu_\varepsilon$  then implies that  $\limsup_{\varepsilon \rightarrow 0} \varepsilon \log \mu_\varepsilon(\{s : d'(\Psi_\varepsilon(s), \Psi_0(s)) > \alpha\}) \leq -c$  for arbitrary  $c$ , so that this limit has to be  $-\infty$ .

Furthermore, the condition that the  $\Psi_\varepsilon$  should be continuous everywhere can be replaced by continuity in an open neighbourhood of  $\{s \in S : \mathcal{I}(s) < \infty\}$  – indeed, just apply the above reasoning to the measures  $\mu_\varepsilon$  conditioned on this open neighbourhood of  $\{s \in S : \mathcal{I}(s) < \infty\}$  on which the functions  $\Psi_\varepsilon$  are continuous. These conditioned measures have the same large deviation behaviour as the original measures.  $\square$

Now we proceed to the discussion of large deviations for Banach-valued Wiener chaos. Let  $(B, \mathcal{H}, \mu)$  be an abstract Wiener space, i.e.  $\mu$  is a centred Gaussian measure on a separable Banach space  $B$  with Cameron-Martin space  $\mathcal{H} \subset B$ . Let  $(e_i)_{i=1}^\infty$  be an orthonormal basis of  $\mathcal{H}$  consisting of elements of the dual space  $B^*$ , where  $B^* \subset \mathcal{H}$  is defined by dualising the inclusion  $\mathcal{H} \subset B$  and identifying  $\mathcal{H}^*$  with  $\mathcal{H}$ . (Such a basis can always be found.) Then for every  $e_i$  the linear mapping  $\xi \mapsto \langle \xi, e_i \rangle$  is continuous on  $B$  and defines a centred Gaussian random variable with  $\int \langle \xi, e_i \rangle \langle \xi, e_j \rangle \mu(d\xi) = \delta_{ij}$ .

Recall that for  $\xi \in \mathbf{R}$  the Hermite polynomials  $H_k(\xi)$  are defined by

$$\exp\left(\lambda\xi - \frac{1}{2}\lambda^2\right) = \sum_{k=0}^{\infty} \frac{\lambda^k}{\sqrt{k!}} H_k(\xi),$$

i.e.  $H_0(\xi) = 1$ ,  $H_1(\xi) = \xi$ ,  $H_2(\xi) = \frac{1}{\sqrt{2}}(\xi^2 - 1)$ ,  $H_3(\xi) = \frac{1}{\sqrt{6}}(\xi^3 - 3\xi)$  etc.

*Remark 3.4.* The convention used here is that if  $\xi$  is a centred normal random variable with variance 1, then the random variables  $H_k(\xi)$  are orthonormal in  $L^2(\Omega)$ .

We call multi-index any sequence  $\alpha = (\alpha_1, \alpha_2, \dots) \in \mathbf{N}_0^{\mathbf{N}}$  with only finitely many non-zero entries. We use the convention  $\alpha! = \prod_{i=1}^{\infty} (\alpha_i!)$  and  $|\alpha| = \sum_{i=1}^{\infty} \alpha_i$ .

Then for any multi-index  $\alpha$  we define the  $\mathbf{R}$ -valued function

$$H_\alpha(\xi) = \prod_{i=1}^{\infty} H_{\alpha_i}(\langle \xi, e_i \rangle) \quad \text{for } \xi \in B.$$

For  $k \in \mathbf{N}$  the  $k$ -th homogeneous  $\mathbf{R}$ -valued Wiener chaos  $\mathcal{H}^{(k)}(\mu) = \mathcal{H}^{(k)}(\mu, \mathbf{R})$  is defined as the closure in  $L^2(\mu, \mathbf{R})$  of the linear space generated by the functions  $H_\alpha(\xi)$  with  $|\alpha| = k$ .

Let  $E$  be a real separable Banach space and denote by  $L^p(\mu, E)$  ( $p \in [1, \infty)$ ) the space of all  $E$ -valued measurable functions  $\Psi$  on  $(B, \mu)$  with  $\int \|\Psi(\xi)\|_E^p \mu(d\xi) < \infty$ . For any  $k \in \mathbf{N}$  the  $k$ -th  $E$ -valued Wiener chaos  $\mathcal{H}^{(k)}(\mu, E)$  is defined as the closure in  $L^2(\mu, E)$  of the linear space generated by random variables of the form  $H_\alpha(\xi) y$  with  $|\alpha| = k$  and  $y \in E$ . Equivalently, one can set

$$\mathcal{H}^{(k)}(\mu, E) \stackrel{\text{def}}{=} \left\{ \Psi \in L^2(\mu, E) : \int \Psi(\xi) H_\alpha(\xi) \mu(d\xi) = 0 \text{ for all } |\alpha| \neq k \right\}. \quad (3.3)$$

Indeed, it is obvious that each random variable  $\Psi(\xi) = \sum_{|\alpha|=k} y_\alpha H_\alpha(\xi)$  with only finitely many summands satisfies the condition specified in (3.3) and that this condition is stable under convergence in  $L^2(\mu, E)$ . On the other hand, if  $\Psi \in L^2(\mu, E)$  satisfies the condition of (3.3), then the conditional expectation of  $\Psi$  with respect to the  $\sigma$ -algebra  $\mathcal{F}_N \stackrel{\text{def}}{=} \sigma(\langle \xi, e_1 \rangle, \dots, \langle \xi, e_N \rangle)$  is given by

$$\Psi_N(\xi) \stackrel{\text{def}}{=} \sum_{\substack{|\alpha|=k \\ \alpha_i=0, i>N}} \left( \int \Psi(\xi') H_\alpha(\xi') \mu(d\xi') \right) H_\alpha(\xi), \quad (3.4)$$

and the convergence theorem for Banach-valued martingales [Nev75, Prop. V.2.6] shows that this series converges to  $\Psi$  in  $L^2(\mu, E)$  as  $N$  goes to infinity. Actually, the strong integrability properties of random variables in a fixed Wiener chaos (see



e.g. [Bor84a, Thm. 4.1]) imply the much stronger statement that there exists a constant  $C$  and a sequence  $\beta_N$  with  $\lim \beta_N \rightarrow \infty$  such that

$$\int \exp \left( \beta_N \|\Psi(\xi) - \Psi_N(\xi)\|_{\frac{2}{E}} \right) \mu(d\xi) \leq C . \quad (3.5)$$

It is important to note here that even though we are using what looks like an  $L^2$  theory,  $E$  is *not* assumed to be a Hilbert space. As a matter of fact, this point is absolutely crucial to our argument.

Fix now a

$$\Psi(\xi) = \sum_{|\alpha|=k} y_\alpha H_\alpha(\xi) \quad (3.6)$$

in  $\mathcal{H}^{(k)}(\mu, E)$ . Denote by  $\Psi_N$  the same sum restricted to those indices  $\alpha$  with  $\alpha_i = 0$  for  $i > N$ . For any  $h \in \mathcal{H}$  we can define the *homogeneous part* of  $\Psi$  by

$$\Psi_{\text{hom}}(h) = \int \Psi(\xi + h) \mu(d\xi) . \quad (3.7)$$

The Cameron-Martin theorem implies that for every  $h \in \mathcal{H}$  the random variable  $\Psi(\xi + h)$  is well defined  $\mu$ -almost surely and that the expectation in (3.7) converges. For the truncated sum  $\Psi_N$  it is easy to check that

$$(\Psi_N)_{\text{hom}}(h) \stackrel{\text{def}}{=} \int \Psi_N(\xi + h) \mu(d\xi) = \sum_{\substack{|\alpha|=k \\ \alpha_i=0, i>N}} y_\alpha h^\alpha ,$$

where  $h^\alpha \stackrel{\text{def}}{=} \prod_{i=1}^{\infty} \langle h, e_i \rangle^{\alpha_i}$ . On the other hand we have by the Cameron-Martin theorem and Cauchy-Schwarz inequality

$$\begin{aligned} & \left\| (\Psi_N)_{\text{hom}}(h) - \Psi_{\text{hom}}(h) \right\|_E \\ &= \left\| \int (\Psi_N(\xi + h) - \Psi(\xi + h)) \mu(d\xi) \right\|_E \\ &= \left\| \int \exp \left( \langle h, \xi \rangle - \frac{1}{2} |h|_{\mathcal{H}}^2 \right) (\Psi_N(\xi) - \Psi(\xi)) \mu(d\xi) \right\|_E \\ &\leq \exp \left( \frac{1}{2} |h|_{\mathcal{H}}^2 \right) \left\| \Psi_N(\xi) - \Psi(\xi) \right\|_{L^2(\mu, E)} , \end{aligned} \quad (3.8)$$

which implies uniform convergence of the sum

$$\Psi_{\text{hom}}(h) = \sum_{|\alpha|=k} y_\alpha h^\alpha$$

on bounded subsets of  $\mathcal{H}$ .

From now on, we will consider random variables  $\Psi = \bigoplus_{\tau \in \mathcal{W}} \Psi_\tau$ , where  $\mathcal{W}$  is some finite index set and each component  $\Psi_\tau$  lies in an  $E_\tau$ -valued inhomogeneous Wiener chaos of order  $K_\tau$ . This means that for each  $\tau$  we can write

$$\Psi_\tau = \sum_{k=0}^{K_\tau} \Psi_{\tau,k} ,$$

where  $\Psi_{\tau,k} \in \mathcal{H}^{(k)}(\mu, E_\tau)$  and  $E_\tau$  is some separable Banach space. Let  $(\Psi_\delta, \delta \in (0, 1))$  be a family of such random variables. We are going to say that  $\Psi_\delta$  converges

to  $\Psi$  if for each  $\tau \in \mathcal{W}$  and  $k \leq K_\tau$

$$\lim_{\delta \rightarrow 0} \int \|\Psi_{\delta;\tau,k}(\xi) - \Psi_{\tau,k}(\xi)\|_{E_\tau}^2 \mu(d\xi) = 0. \quad (3.9)$$

We also define the homogenous part of  $\Psi$  as  $\Psi_{\text{hom}} = \bigoplus_{\tau \in \mathcal{W}} (\Psi_{\tau, K_\tau})_{\text{hom}}$ . Note that in this definition there is no contribution corresponding to the  $\Psi_{\tau,k}$  for  $k < K_\tau$  to  $\Psi_{\text{hom}}$ . The reason for this is that this respects the scaling used in (3.10).

The main result of this section is a large deviation statement for the random variables  $\Psi$  when changing the *noise intensity*. The natural rescaling of the different coordinates in our context is given by

$$\Psi^{(\varepsilon)}(\xi) \stackrel{\text{def}}{=} \bigoplus_{\tau \in \mathcal{W}} \varepsilon^{\frac{K_\tau}{2}} \Psi_\tau(\xi). \quad (3.10)$$

The main result of this section states that under the measure  $\mu$ , the random variables  $\Psi^{(\varepsilon)}$  satisfy a large deviation principle that is stable under convergence to  $\Psi$ . For any  $\mathbf{s} = \bigoplus_{\tau \in \mathcal{W}} s_\tau \in \mathbf{E} \stackrel{\text{def}}{=} \bigoplus_{\tau \in \mathcal{W}} E_\tau$ , the rate function is given by

$$\mathcal{I}(\mathbf{s}) = \inf \left\{ \frac{1}{2} |h|_{\mathcal{H}}^2 : h \in \mathcal{H} \text{ with } \Psi_{\text{hom}}(h) = \mathbf{s} \right\}, \quad (3.11)$$

where the infimum is interpreted as  $+\infty$  if the set is empty.

**Theorem 3.5** (Large deviations in Wiener chaos). *Assume that  $\Psi_\delta$  converges to  $\Psi$  as  $\delta$  tends to 0 in the sense of (3.9). Choose  $\delta = \delta(\varepsilon) \geq 0$  such that  $\lim_{\varepsilon \downarrow 0} \delta(\varepsilon) = 0$ . Then the random variables  $\Psi_{\delta(\varepsilon)}^{(\varepsilon)}(\xi)$  satisfy a large deviation principle on  $\mathbf{E}$  with rate  $\varepsilon$  and rate function  $\mathcal{I}$  given by (3.11).*

*Remark 3.6.* The reason that the lower order contributions  $\Psi_{\tau,k}$  for  $k < K_\tau$  vanish on the level of the large deviation principle is the scaling (3.10). In order to have a contribution from  $\Psi_{\tau,k}$  one would have to consider  $\varepsilon^{\frac{k}{2}} \Psi_{\tau,k}$  instead of  $\varepsilon^{\frac{K_\tau}{2}} \Psi_{\tau,k}$ .

As stated above, our argument follows roughly the strategy employed in [DS89] to prove the classical Freidlin-Wentzell estimates (see [DS89, Sec. 1.4 and Lemma 2.1.4]).

Suppose that for every  $\delta \geq 0$ , and every  $\tau \in \mathcal{W}$  and  $k \leq K_\tau$  we have

$$\Psi_{\delta;\tau,k}(\xi) = \sum_{|\alpha|=k} y_{\alpha;\delta;\tau,k} H_\alpha(\xi). \quad (3.12)$$

Denote as before by  $\Psi_{N;\delta}(\xi)$  the conditional expectation of  $\Psi_\delta(\xi)$  with respect to  $\mathcal{F}_N$ , i.e. the series expansion (3.12) of each  $\Psi_{\delta;\tau,k}(\xi)$  is restricted to indices  $\alpha$  with  $\alpha_i = 0$  for  $i > N$ .

**Lemma 3.7.** *For any fixed  $N \in \mathbf{N}$  the random variables  $\Psi_{N,\delta(\varepsilon)}^{(\varepsilon)}(\xi)$  satisfy a large deviation principle with rate function*

$$\mathcal{I}_N(\mathbf{s}) = \inf \left\{ \frac{1}{2} |h|_{\mathcal{H}}^2 : h \in \mathcal{H} \text{ with } (\Psi_N)_{\text{hom}}(h) = \mathbf{s} \right\}. \quad (3.13)$$

*Proof.* As stated above, the random functions  $\Psi_{N;\delta;\tau,k}$  are continuous functions  $B \rightarrow E_\tau$  for every fixed  $N$ .

Note that for every  $\tau \in \mathcal{W}$  and  $k \leq K_\tau$  and for every multi-index  $\alpha$  with  $|\alpha| = k$  we have by Cauchy-Schwarz inequality

$$\begin{aligned} \|y_{\alpha;\delta;\tau,k} - y_{\alpha;\tau,k}\|_{E_\tau}^2 &= \left\| \int (\Psi_{\delta;\tau,k}(\xi) - \Psi_{\tau,k}(\xi)) H_\alpha(\xi) \mu(d\xi) \right\|_{E_\tau}^2 \\ &\leq \int \|\Psi_{\delta;\tau,k}(\xi) - \Psi_{\tau,k}(\xi)\|_{E_\tau}^2 \mu(d\xi). \end{aligned} \quad (3.14)$$

This converges to zero as  $\delta \rightarrow 0$  by assumption.

In order to apply the contraction principle, Lemma 3.3, it is useful to rewrite the rescaled random variables as

$$\varepsilon^{\frac{K_\tau}{2}} \Psi_{N;\delta(\varepsilon);\tau,k}(\xi) = \Phi_{N;\tau,k}^{(\varepsilon)}(\varepsilon^{\frac{1}{2}} \xi),$$

where

$$\Phi_{N;\tau,k}^{(\varepsilon)}(\xi) \stackrel{\text{def}}{=} \varepsilon^{\frac{K_\tau-k}{2}} \sum_{\substack{|\alpha|=k \\ \alpha_i=0, i>N}} y_{\alpha;\delta(\varepsilon);\tau,k} \Phi_\alpha^{(\varepsilon)}(\xi),$$

and

$$\Phi_\alpha^{(\varepsilon)}(\xi) = \prod_{i=1}^{\infty} \varepsilon^{\frac{\alpha_i}{2}} H_{\alpha_i}(\varepsilon^{-\frac{1}{2}} \langle \xi, e_i \rangle).$$

Now it is easy to check, using (3.14) that for  $k < K_\tau$  the functions  $\Phi_{N;\tau,k}^{(\varepsilon)}$  converge to 0 and the functions  $\Phi_{N;\tau,K_\tau}^{(\varepsilon)}$  converge to  $(\Psi_{N;\tau,K_\tau})_{\text{hom}}$ , in both cases uniformly on bounded subsets of  $B$ . This implies in particular that the slightly weaker condition (2) in Lemma 3.3 holds.

Furthermore, it is well known (see e.g. [Led96, Chapter 4]) that under  $\mu$ , the random vectors  $\sqrt{\varepsilon}\xi$  satisfy a large deviation principle on  $B$  with rate  $\varepsilon$  and rate function

$$\mathcal{I}_\xi(h) = \begin{cases} \frac{1}{2} |h|_{\mathcal{H}}^2 & \text{for } h \in \mathcal{H}, \\ +\infty & \text{else.} \end{cases}$$

Hence, the Lemma follows from the contraction principle, Lemma 3.3.  $\square$

**Lemma 3.8.** *Let  $\mathcal{I}$  be given by (3.11) and let  $\mathcal{I}_N$  be given by (3.13). Then  $\mathcal{I} : \mathbf{E} \rightarrow [0, \infty]$  is lower semicontinuous with compact sublevel sets. Furthermore for any closed set  $\mathcal{C} \subseteq \mathbf{E}$  we have*

$$\lim_{\lambda \rightarrow 0} \liminf_{N \rightarrow \infty} \inf_{\mathbf{s} \in \mathcal{C}_\lambda} \mathcal{I}_N(\mathbf{s}) = \inf_{\mathbf{s} \in \mathcal{C}} \mathcal{I}(\mathbf{s}). \quad (3.15)$$

Here  $\mathcal{C}_\lambda \stackrel{\text{def}}{=} \{\mathbf{s} \in \mathbf{E} : \text{dist}_{\mathbf{E}}(\mathbf{s}, \mathcal{C}) \leq \lambda\}$ .

*Proof of Lemma 3.8.* As a consequence of (3.8), the assumptions of [DS89, Lemma 2.1.4] are satisfied, from which the first statement immediately follows. The estimate (3.15) then follows from the first part of the proof of [DS89, Lemma 2.1.4].  $\square$

**Lemma 3.9** (Exponential equivalence). *We have for every  $\lambda > 0$*

$$\limsup_{N \rightarrow \infty} \limsup_{\varepsilon \rightarrow 0} \varepsilon \log \mu(\xi : \|\Psi_{\delta(\varepsilon)}^{(\varepsilon)}(\xi) - \Psi_{N;\delta(\varepsilon)}^{(\varepsilon)}(\xi)\|_{\mathbf{E}} \geq \lambda) = -\infty.$$

*Proof.* We write

$$\int \|\Psi_{\delta(\varepsilon);\tau}(\xi) - \Psi_{N;\delta(\varepsilon);\tau}(\xi)\|_{E_\tau}^2 \mu(d\xi) \leq 3 \int \|\Psi_{\delta(\varepsilon);\tau}(\xi) - \Psi_\tau(\xi)\|_{E_\tau}^2 \mu(d\xi)$$

$$\begin{aligned}
& + 3 \int \|\Psi_\tau(\xi) - \Psi_{N;\tau}(\xi)\|_{E_\tau}^2 \mu(d\xi) \\
& + 3 \int \|\Psi_{N;\tau}(\xi) - \Psi_{N;\delta(\varepsilon);\tau}(\xi)\|_{E_\tau}^2 \mu(d\xi) .
\end{aligned}$$

The assumption (3.9) on convergence of the  $\Psi_\delta$  to  $\Psi$  implies that the first term on the right hand side goes to zero as  $\varepsilon \rightarrow 0$ . The conditional expectation is a contraction on  $L^2(\mu, E_\tau)$ , and hence the third term goes to zero as well. In particular, for fixed  $N$  we can find an  $\varepsilon_N$  such that for  $\varepsilon < \varepsilon_N$  the right hand side can be bounded by  $4 \int \|\Psi_\tau(\xi) - \Psi_{N;\tau}(\xi)\|_{E_\tau}^2 \mu(d\xi)$ . As a consequence of the integrability properties of random variables belonging to a fixed Wiener chaos [Bor84a, Theorem 4.1], we conclude from this that there exists a sequence  $\beta_N$  increasing to  $\infty$  as  $N \rightarrow \infty$  and a constant  $0 < C < \infty$  such that for every  $N$  and for every  $\varepsilon < \varepsilon_N$  we have

$$\int \exp\left(\beta_N \|\Psi_{\delta(\varepsilon);\tau}(\xi) - \Psi_{N;\delta(\varepsilon);\tau}(\xi)\|_{E_\tau}^{\frac{2}{K_\tau}}\right) \mu(d\xi) \leq C .$$

Using Chebyshev's inequality we get for every  $N$  and every  $0 < \varepsilon < \varepsilon_N$  that

$$\begin{aligned}
& \mu\left(\xi : \varepsilon^{\frac{K_\tau}{2}} \|\Psi_{\delta(\varepsilon);\tau}(\xi) - \Psi_{N;\delta(\varepsilon);\tau}(\xi)\|_{E_\tau} \geq \lambda\right) \\
& = \mu\left(\xi : \exp\left(\beta_N \|\Psi_{\delta(\varepsilon);\tau}(\xi) - \Psi_{N;\delta(\varepsilon);\tau}(\xi)\|_{E_\tau}^{\frac{2}{K_\tau}}\right) \geq \exp\left(\beta_N \lambda^{\frac{2}{K_\tau}} \varepsilon^{-1}\right)\right) \\
& \leq C \exp\left(-\beta_N \lambda^{\frac{2}{K_\tau}} \varepsilon^{-1}\right) .
\end{aligned}$$

So the result follows.  $\square$

With Lemma 3.7, Lemma 3.8, and Lemma 3.9 in hand Theorem 3.5 follows easily.

*Proof of Theorem 3.5.* It was already shown in Lemma 3.8 that the rate function  $\mathcal{I}$  satisfies the first condition in Definition 3.1, so it remains to establish the upper bound (3.1) and the lower bound (3.2).

Let  $\mathcal{C} \subset \mathbf{E}$  be closed. Then we have for every  $\lambda > 0$  and  $N \in \mathbf{N}$  that

$$\begin{aligned}
\mu(\xi : \Psi_{\delta(\varepsilon)}^{(\varepsilon)}(\xi) \in \mathcal{C}) & \leq \mu(\xi : \Psi_{N;\delta(\varepsilon)}^{(\varepsilon)}(\xi) \in \mathcal{C}_\lambda) \\
& \quad + \mu(\xi : \|\Psi_{\delta(\varepsilon)}^{(\varepsilon)}(\xi) - \Psi_{N;\delta(\varepsilon)}^{(\varepsilon)}(\xi)\|_{\mathbf{E}} \geq \lambda) ,
\end{aligned}$$

where as above  $\mathcal{C}_\lambda \stackrel{\text{def}}{=} \{\mathbf{x} : \text{dist}_{\mathbf{E}}(\mathbf{x}, \mathcal{C}) \leq \lambda\}$ . According to Lemma 3.9 for any fixed  $\lambda > 0$  we can choose  $N$  large enough (depending on  $\mathcal{C}$ ) to ensure

$$\limsup_{\varepsilon \rightarrow 0} \varepsilon \log \mu(\xi : \|\Psi_{\delta(\varepsilon)}^{(\varepsilon)}(\xi) - \Psi_{N;\delta(\varepsilon)}^{(\varepsilon)}(\xi)\|_{\mathbf{E}} \geq \lambda) \leq - \inf_{\mathbf{s} \in \mathcal{C}} \mathcal{I}(\mathbf{s}) .$$

On the other hand  $\mathcal{C}_\lambda$  is closed, so that Lemma 3.7 implies that for any fixed  $N$

$$\limsup_{\varepsilon \rightarrow 0} \varepsilon \log \mu(\xi : \Psi_{N;\delta(\varepsilon)}^{(\varepsilon)}(\xi) \in \mathcal{C}_\lambda) \leq - \inf_{\mathbf{s} \in \mathcal{C}_\lambda} \mathcal{I}_N(\mathbf{s}) .$$

Letting first  $N$  go to  $\infty$  and then  $\lambda$  to 0 and applying Lemma 3.8 finishes the proof of the upper bound (3.1).

In a similar way, let  $\mathcal{O} \subset \mathbf{E}$  be open and  $\mathbf{s} \in \mathcal{O}$  and  $\lambda > 0$  be such that  $\{\mathbf{x} \in \mathbf{E} : \|\mathbf{x} - \mathbf{s}\|_{\mathbf{E}} \leq 2\lambda\} \subseteq \mathcal{O}$ . Then we have

$$\mu(\xi : \Psi_{\delta(\varepsilon)}^{(\varepsilon)}(\xi) \in \mathcal{O}) \geq \mu(\xi : \|\Psi_{N;\delta(\varepsilon)}^{(\varepsilon)}(\xi) - \mathbf{s}\|_{\mathbf{E}} < \lambda)$$

$$- \mu(\xi: \|\Psi_{\delta(\varepsilon)}^{(\varepsilon)}(\xi) - \Psi_{N;\delta(\varepsilon)}^{(\varepsilon)}(\xi)\|_{\mathbf{E}} \geq \lambda) .$$

As above, Lemma 3.9 implies that for fixed  $\lambda > 0$  and  $N$  large enough (depending on  $\mathbf{s}$ ) we have

$$\limsup_{\varepsilon \rightarrow 0} \varepsilon \log \mu(\xi: \|\Psi_{\delta(\varepsilon)}^{(\varepsilon)}(\xi) - \Psi_{N;\delta(\varepsilon)}^{(\varepsilon)}(\xi)\|_{\mathbf{E}} \geq \lambda) \leq -\mathcal{I}(\mathbf{s}) - 1 ,$$

and we get from Lemma 3.7 that for any  $N$

$$\liminf_{\varepsilon \rightarrow 0} \varepsilon \log \mu(\xi: \|\Psi_{N;\delta(\varepsilon)}^{(\varepsilon)}(\xi) - \mathbf{s}\|_{\mathbf{E}} < \lambda) \geq - \inf_{\{\|\mathbf{x} - \mathbf{s}\| \leq \frac{1}{2}\lambda\}} \mathcal{I}_N(\mathbf{x}) .$$

Letting again  $N$  go to  $\infty$  first and then  $\lambda$  to zero, we obtain the required lower bound from Lemma 3.8.  $\square$

#### 4. LARGE DEVIATIONS FOR STOCHASTIC PDES

In this section we apply our abstract large deviation result, Theorem 3.5, to the specific setting of the models arising in the solution theory for the stochastic Allen-Cahn equation. We start by recalling that  $\mathcal{W}_-^{(3)} = \{\Xi, \uparrow, \vee, \Psi, \Psi, \Psi\}$ . We set  $\mathbf{E} \stackrel{\text{def}}{=} \bigoplus_{\tau \in \mathcal{W}_-^{(3)}} E_\tau$ , where for  $\tau \in \mathcal{W}_-^{(3)} \setminus \{\Xi, \uparrow\}$  the spaces  $E_\tau$  are given by the closure of the set of smooth functions  $(z, \bar{z}) \mapsto (\Pi_z \tau)(\bar{z})$  in the topology given by the norms

$$\|\Pi \tau\| = \sup_{\varphi, \lambda, z} \lambda^{-|\tau|} |(\Pi_z \tau)(\varphi_z^\lambda)| ,$$

where the supremum runs over  $\lambda \in (0, 1]$ ,  $z \in [-1, T + 1] \times \mathbf{T}^d$ , and  $\varphi$  runs over a suitable set of test functions as in (2.12). For  $\Xi$  and  $\uparrow$  the topology is given by the norms listed in (2.18). Since these norms are weaker than the supremum norm over a compact set, it follows from the Stone-Weierstrass theorem that smooth functions can be approximated by polynomials with rational coefficients, so that the spaces  $E_\tau$  are separable.

Note that  $\mathcal{M}_-^{(d)}$  can be viewed as a (nonlinear and rather complicated) closed subset of  $\mathbf{E}$ . In the case  $d = 2$  one can simplify the situation considerably by restricting oneself to  $\tau \in \{\Xi, \uparrow, \vee, \Psi\}$  and by dropping the  $z$ -dependence of the functions  $\Pi \tau$ . For the remainder of this section, we will only consider the case  $d = 3$ , the case  $d = 2$  follows *mutatis mutandis*.

Let  $\mathcal{H}$  be  $L^2([-2, T + 2] \times \mathbf{T}^3)$ . The Gaussian measure  $\mu$  with Cameron-Martin space  $\mathcal{H}$  can be realised on  $B$ , the closure of the space of smooth functions in the Besov space  $\mathcal{C}^{-\frac{5}{2}-\kappa}$  for any  $\kappa > 0$ , and the triple  $(B, \mathcal{H}, \mu)$  is an abstract Wiener space. Let  $\varrho$  be a mollifying kernel, i.e.  $\varrho: \mathbf{R} \times \mathbf{R}^3 \rightarrow \mathbf{R}$  is a smooth function with  $\int \varrho(z) dz = 1$  and compact support contained in the unit ball of  $\mathbf{R} \times \mathbf{R}^3$ . Then for  $\xi \in B$  and for  $\delta \in (0, 1)$  set

$$\xi_\delta(z) = \langle \xi, \varrho_z^\delta \rangle , \quad (4.1)$$

where  $\varrho_z^\delta$  is defined as above in (2.7). Note that under  $\mu$  the random distribution  $\xi$  is a realisation of space-time white noise and  $\xi_\delta$  is a smooth approximation to  $\xi$ .

*Remark 4.1.* One could also have considered different approximations, like for example regularisations in space only or in time only. These could also be handled in the same way, but we restrict ourselves to space-time regularisations in order to be able to reuse the convergence results obtained in [Hai14c].

Let  $\Psi(\xi_\delta) = (\Pi^{\xi_\delta}, F^{\xi_\delta})$  be the canonical model constructed from  $\Pi_y^{\xi_\delta} \Xi = \xi_\delta$  as described above in (2.10c) and (2.13) (we have added a superscript  $\xi_\delta$  to make more explicit the dependence on the specific realisation of the noise). As discussed in Section 2.3 and Section 2.4 ( $\Pi^{\xi_\delta}, F^{\xi_\delta}$ ) lives in  $\mathcal{M}^{(d)}$  and it is uniquely characterised by the *minimal model* i.e. the  $\Pi^{\xi_\delta} \tau$  for  $\tau \in \mathcal{W}_-^{(3)}$ .

We show now that the mapping  $\Psi_\delta$  that maps the white noise  $\xi$  to the minimal model fits exactly into the framework developed in Section 3. For any  $\delta > 0$  set  $\Psi_\delta = \bigoplus_{\tau \in \mathcal{W}_-^{(3)}} \Psi_{\delta;\tau}$  where  $\Psi_{\delta;\tau}(\xi) \stackrel{\text{def}}{=} \Pi^{\xi_\delta} \tau$  as above. By construction, for a fixed realisation of the noise  $\xi$ , the distributions  $\varphi \mapsto \Pi_z^{\xi_\delta} \tau(\varphi)$  can be identified with a smooth function in the two variables  $z$  and  $\bar{z}$ . In particular, it also takes values in  $E_\tau$ .

It follows from the analysis in [Hai14c, Sec. 10] that for each  $\tau \in \mathcal{W}_-^{(3)}$  and for  $\delta > 0$  the mapping  $\xi \mapsto \Psi_{\delta;\tau}(\xi)$  belongs to the  $E_\tau$ -valued inhomogenous Wiener chaos of order  $K_\tau$ , where  $K_\tau$  is the number of occurrences of the symbol  $\Xi$  in  $\tau$ . More precisely, we have

$$K_\Xi = K_\dagger = 1, \quad K_\vee = 2, \quad K_\Psi = 3, \quad K_{\Psi\ddagger} = K_{\Psi\ddot{\Psi}} = 4, \quad K_{\Psi\ddot{\Psi}\ddot{\Psi}} = 5.$$

We briefly outline this analysis. Let  $\mathcal{H}^{\otimes_s k}$  denote the  $k$ -fold symmetric tensor power of  $\mathcal{H}$ . We identify  $\mathcal{H}^{\otimes_s k}$  with the space of symmetric square integrable functions  $\mathcal{W}$  in  $k$  arguments  $z_i \in [-2, T+2] \times \mathbf{T}^3$ . Then it is well known (see e.g. [Nua06, Sec 1.1]) that for every  $k \geq 1$  there is an isometry (up to a factor  $\sqrt{k!}$ )

$$I_k: \mathcal{H}^{\otimes_s k} \rightarrow \mathcal{H}^{(k)}(\mu, \mathbf{R})$$

given by iterated stochastic integrals: For all  $\xi$  in a set of  $\mu$  measure one we have

$$I_k(\mathcal{W})(\xi) = \int \cdots \int \mathcal{W}(z_1, \dots, z_k) \xi(dz_1) \cdots \xi(dz_k).$$

We can thus identify every random variable in  $\mathcal{H}^{(k)}(\mu, \mathbf{R})$  with a symmetric kernel in  $k$  arguments. We can also apply  $I_k$  to arbitrary kernels in  $L^2([-2, T+2] \times \mathbf{T}^3)^{\otimes k}$  by precomposing with the symmetrisation map

$$\mathcal{W}^{(k)}(z_1, z_2, \dots, z_k) \mapsto \frac{1}{k!} \sum_{\sigma} \mathcal{W}^{(k)}(z_{\sigma(1)}, \dots, z_{\sigma(k)}), \quad (4.2)$$

where the sum runs over all elements  $\sigma$  in the group of permutations of  $\{1, \dots, k\}$ . The symmetrisation (4.2) is a contraction on  $L^2([-2, T+2] \times \mathbf{T}^3)^{\otimes k}$ . So we obtain

$$\int (I_k(\mathcal{W}^{(k)})(\xi))^2 \mu(d\xi) \leq k! \int \mathcal{W}^{(k)}(z_1, z_2, \dots, z_k)^2 dz_1 \cdots dz_k. \quad (4.3)$$

It is shown recursively in [Hai14c, Sec. 10], that for any  $\tau \in \mathcal{W}_-^{(3)}$  and for  $\delta > 0$ ,  $\Pi_y^{\xi_\delta}(\tau)$  can be characterised by kernels  $\mathcal{W}_{\delta;\tau,k}(\bar{y}, y, \cdot) \in \mathcal{H}^{\otimes k}$  for  $k = K_\tau, K_\tau - 2, K_\tau - 4, \dots$ . The precise form of the kernels can be read off directly from the graphs  $\dagger, \vee, \Psi$  etc. in (2.3) above (in fact, this is the reason for this notation).

We illustrate this construction for  $\tau = \Psi\ddot{\Psi}$ , which is the most involved term. All other symbols can be dealt with in a very similar (but easier) way. The kernel  $\mathcal{W}_{\delta,\Psi,5}(\bar{y}, y, \cdot)$  for the contribution in the highest Wiener chaos  $K_{\Psi\ddot{\Psi}} = 5$  is given by

$$\mathcal{W}_{\delta,\Psi,5}(\bar{y}, y; z_1, \dots, z_5) \stackrel{\text{def}}{=} \int dx K_\delta(z_1 - x) K_\delta(z_2 - x) K_\delta(z_3 - x) \quad (4.4)$$

$$(K(x - \bar{y}) - K(x - y)) K_\delta(z_4 - y) K_\delta(z_5 - y) .$$

Here  $K$  denotes the kernel introduced in Section 2.3 and  $K_\delta \stackrel{\text{def}}{=} K \star \varrho_\delta$ . The connection between (4.4) and the graphical representation  $\Psi$  is the following: Each vertex of  $\Psi$  corresponds to a variable in  $[-2, T + 2] \times \mathbf{T}^3$ - the leafs are  $z_1, \dots, z_5$ , the root corresponds to  $\bar{y}$  and  $y$  and the central vertex corresponds to  $x$ . Then as explained in (2.3) every down facing line in  $\Psi$  corresponds to one occurrence of a kernel  $K$  or  $K_\delta$  in (4.4): If the line is connected to a leaf, the kernel is  $K_\delta$  and the kernel is  $K$  otherwise. The variable  $x$  corresponding to a vertex that are neither a leaf nor the root is integrated out. The fact that we have to subtract one occurrence of the kernel in  $(K(x - \bar{y}) - K(x - y))$  corresponds to the second term in (2.10c) which has to be added to guarantee for the right behaviour near the diagonal  $\bar{y} = y$  (see [Hai14c, Sec. 6 and Sec. 10] for details).

If  $h \in \mathcal{H}$  is an  $L^2$ -function then it is easy to check from the recursive construction in (2.10c) and (2.13) that for every test function  $\varphi$  the iterated integral

$$\int \langle \mathcal{W}_{\delta, \Psi, 5}(\cdot, y, z_1, \dots, z_5), \varphi \rangle h(z_1) \dots h(z_5) dz_1 \dots dz_5 \quad (4.5)$$

yields the value  $\Pi_y^{h_\delta} \Psi(\varphi)$  for the canonical model constructed from  $\Pi_y^{h_\delta} \Xi = h_\delta \stackrel{\text{def}}{=} h \star \varrho_\delta$ . Here we have set

$$\langle \mathcal{W}_{\delta, \Psi, 5}(\cdot, y, z_1, \dots, z_5), \varphi \rangle \stackrel{\text{def}}{=} \int \mathcal{W}_{\delta, \Psi, 5}(\bar{y}, y, z_1, \dots, z_5) \varphi(\bar{y}) d\bar{y} . \quad (4.6)$$

In the case of the regularised white noise  $\xi_\delta$  the relationship between the kernels and the canonical model is slightly more complicated. The iterated stochastic integral

$$I_5(\langle \mathcal{W}_{\delta, \Psi, 5}(\cdot, y, \dots), \varphi \rangle) = \int \langle \mathcal{W}_{\delta, \Psi, 5}(\cdot, y, z_1, \dots, z_5), \varphi \rangle \xi(dz_1) \dots \xi(dz_5) \quad (4.7)$$

does not give the value  $\Pi_y^{\xi_\delta} \Psi(\varphi)$  for the canonical model contracted from  $\xi_\delta$ . Indeed, iterated integrals obey an *Itô-type* chain rule that produces some extra terms. These terms can be represented by additional kernels  $\mathcal{W}_{\delta, \Psi, 3}(\bar{y}, y, z_1, z_2, z_3)$  and  $\mathcal{W}_{\delta, \Psi, 1}(\bar{y}, y, z_1)$ . These kernels are obtained through suitable *contractions* of the graph  $\Psi$ . Their precise form is not relevant for our discussion and we refer to [Hai14c, Sec. 10.5] and [Nua06] for more detailed explanations. Once these kernels are constructed, the random variable  $\Pi_y^{\xi_\delta} \Psi$  is given as

$$\Pi_y^{\xi_\delta} \Psi(\varphi) \stackrel{\text{def}}{=} \sum_{k \in \{5, 3, 1\}} I_k(\langle \mathcal{W}_{\delta, \Psi, k}(\cdot, y, \dots), \varphi \rangle) . \quad (4.8)$$

As above in (4.6) the pointed bracket on the right hand side denotes integration of the kernels  $\mathcal{W}_{\delta, \Psi, k}$  against  $\varphi$  in the first variable  $\bar{y}$ , whereas  $I_k$  represents the iterated stochastic integration with respect to the remaining variables  $z_i$ .

The decomposition (4.8) is useful for several reasons. On one hand, it gives a natural decomposition of  $\Pi_y^{\xi_\delta} \Psi$  into its components in homogenous Wiener chaoses. Furthermore, (4.8) in conjunction with the estimate (4.3) reduces the problem of estimating the second moment of the expression  $\Pi_y^{\xi_\delta} \Psi(\varphi)$  to deriving bounds on the  $L^2$ -norm of the kernels  $\mathcal{W}_{\delta, \Psi, k}$ . Using the equivalence of all moments in a fixed Wiener chaos, these bounds in turn can then be used as input into a *Kolmogorov-like theorem* (see [Hai14c, Thm 10.7]) to obtain bounds on  $\int \|\Pi_y^{\xi_\delta} \tau(\xi)\|_{E_\tau}^p \mu(d\xi)$  for arbitrary  $1 < p < \infty$ .

On the other hand, we can immediately read off from (4.5) and (4.8) the homogeneous part of  $\Psi_{\delta;\Psi}$ . According to the discussion in Section 3, the terms corresponding to the kernels  $\mathcal{W}_{\Psi,k}$  for  $k = 1, 3$  do not influence the value of  $(\Psi_{\delta;\Psi})_{\text{hom}}$ . For  $k = K_{\Psi} = 5$  the construction of the iterated integrals (see [Nua06, Section 1.1]) yields an explicit expression of the quantity  $\Psi_{\delta;\Psi,5}$  evaluated at a shifted noise  $\xi + h$  for  $h \in \mathcal{H}$ . For any  $h \in \mathcal{H}$  and every test function  $\varphi$  we get for  $\mu$ -almost every  $\xi$

$$\Psi_{\delta;\Psi,5}(\xi + h)[y, \varphi] = \int \langle \mathcal{W}_{\delta;\Psi,5}(\cdot, y, z_1, \dots, z_5), \varphi \rangle \quad (4.9)$$

$$(\xi(dz_1) + h(z_1)dz_1) \cdots (\xi(dz_5) + h(z_5)dz_5) .$$

Here we use square brackets to denote the evaluation of  $\Psi_{\delta;\Psi,5}(\xi + h) \in E_{\Psi}$  at  $y$  and at a test function  $\varphi$ . According to (3.7)  $(\Psi_{\delta;\Psi})_{\text{hom}}(h)$  evaluated at  $(y, \varphi)$  is given by the expectation of this expression. But all of the iterated integrals that involve at least one power of  $\xi$  have vanishing expectation. The only term with non-zero expectation with respect to  $\mu$  is given by

$$\int \langle \mathcal{W}_{\delta;\Psi,5}(\cdot, y, z_1, \dots, z_5), \varphi \rangle h(z_1) \cdots h(z_5) dz_1 \cdots dz_5 .$$

According to the discussion around (4.5) this is precisely  $\Pi_y^{h_{\delta};\Psi}(\varphi)$  for the canonical model constructed from  $h_{\delta}$ .

The reasoning for all the other  $\tau \in \mathcal{W}$  is similar – we refer to [Hai14c, Sec. 10] for the precise expressions of the kernels  $\mathcal{W}_{\delta;\tau,k}$ . In each case the  $\Psi_{\delta;\tau,k}$  live in  $\mathcal{H}^{(k)}(\mu, E_{\tau})$  and in each case the homogenous part  $(\Psi_{\delta;\tau})_{\text{hom}}(h)$  is given by the canonical model constructed from  $h_{\delta}$ .

Finally, the recursive definition (2.10c) and (2.13) is homogenous in the input  $\xi$ . Replacing the building block  $\xi$  by  $\sqrt{\varepsilon}\xi$  produces  $\Pi^{\sqrt{\varepsilon}\xi_{\delta}}\tau = \varepsilon^{\frac{K_{\tau}}{2}}\Pi^{\xi_{\delta}}\tau$  almost surely. This is precisely the rescaling assumed in the abstract setting in (3.10). Hence, we are in the situation described in Section 3 and we can apply Theorem 3.5 directly to the functions  $\Psi_{\delta}^{(\varepsilon)}$  for fixed  $\delta$  to obtain the following large deviation result.

**Theorem 4.2.** *For any  $\varepsilon, \delta > 0$  let  $(\Pi^{\sqrt{\varepsilon}\xi_{\delta}}, F^{\sqrt{\varepsilon}\xi_{\delta}}) \in \mathcal{M}^{(3)}$  be the canonical model defined recursively via (2.10c) and (2.13) from  $\Pi_y^{\sqrt{\varepsilon}\xi_{\delta}}\Xi = \sqrt{\varepsilon}\xi_{\delta}$ . In the same way, for  $h \in \mathcal{H}$  let  $(\Pi^{h_{\delta}}, F^{h_{\delta}})$  be the canonical model constructed from  $h_{\delta} = h \star \varrho_{\delta}$ .*

*Then for any fixed value of  $\delta > 0$  the distributions of  $\{(\Pi^{\sqrt{\varepsilon}\xi_{\delta}}, F^{\sqrt{\varepsilon}\xi_{\delta}}), \varepsilon > 0\}$  under  $\mu$  satisfy a large deviation principle on  $\mathcal{M}^{(3)}$  with rate  $\varepsilon$  and rate function*

$$\mathcal{I}_{\text{Model } \delta}((\Pi, F)) \stackrel{\text{def}}{=} \inf \left\{ \frac{1}{2} |h|_{\mathcal{H}}^2 : (\Pi^{h_{\delta}}, F^{h_{\delta}}) = (\Pi, F) \right\} .$$

*Proof.* Theorem 3.5 applied to  $\{\Psi_{\delta}^{(\varepsilon)}; \varepsilon > 0\}$  for fixed value of  $\delta$  yields a large deviation principle for the random variable  $\bigoplus_{\tau \in \mathcal{W}} \Pi^{\sqrt{\varepsilon}\xi_{\delta}}\tau \in \mathbf{E}$ . For  $\mathbf{s} = \bigoplus_{\tau \in \mathcal{W}} s_{\tau}$  the rate function is given by

$$\widetilde{\mathcal{I}}_{\delta}(\mathbf{s}) = \inf \left\{ \frac{1}{2} |h|_{\mathcal{H}}^2 : \Pi^{h_{\delta}}\tau = s_{\tau} \text{ for } \tau \in \mathcal{W}_-^{(3)} \right\} .$$

By construction this random variable takes values in  $\mathcal{M}_-^{(3)}$  almost surely. As  $\mathcal{M}_-^{(3)}$  is a closed subset of  $\mathbf{E}$  it follows automatically that the large deviation principle also holds on  $\mathcal{M}_-^{(3)}$  with respect to the relative topology.



According to Theorem 2.10 the unique extension to an admissible model is continuous from  $\mathcal{M}_-^{(d)}$  to  $\mathcal{M}^{(d)}$ . Hence the contraction principle, Lemma 3.3, implies a large deviation principle for the full model. The rate function is given by

$$\begin{aligned} \mathcal{I}_{\text{Model } \delta}((\Pi, F)) &= \inf \{ \widetilde{\mathcal{I}}_\delta(\mathbf{s}) : \mathbf{s} \in \mathcal{M}_-^{(3)} \text{ extension of } \mathbf{s} \text{ is } (\Pi, F) \} \\ &= \inf \left\{ \frac{1}{2} |h|_{\mathcal{H}}^2 : \Pi^{h_\delta} \tau = s_\tau \text{ for } \tau \in \mathcal{W}_-^{(3)} \text{ and} \right. \\ &\quad \left. \text{extension of } \mathbf{s} = \bigoplus s_\tau \text{ is } (\Pi, F) \right\} \\ &= \inf \left\{ \frac{1}{2} |h|_{\mathcal{H}}^2 : (\Pi^{h_\delta}, F^{h_\delta}) = (\Pi, F) \right\}, \end{aligned}$$

which is precisely our claim.  $\square$

Combining this result with the contraction principle, Lemma 3.3 and Theorem 2.12 we recover the Freidlin-Wentzell type bounds for the stochastic PDE (1.1) driven by a noise term  $\sqrt{\varepsilon} \xi_\delta$  for fixed correlation length  $\delta$  as the noise strength  $\varepsilon$  goes to zero. Of course, such bounds can also be derived without any reference to the theory of regularity structures.

Our approach produces meaningful new results only when considering the limit  $\delta \rightarrow 0$ . Indeed, if we could show that the  $\Psi_\delta$  converges to a limit  $\Psi$  in the sense of (3.9) we could apply Theorem 3.5 to obtain a large deviation principle for any diagonal sequence  $\delta, \varepsilon \rightarrow 0$ . Unfortunately, as already explained above in Section (2.6) this is not true. However, if we replace the models  $\Pi^{\xi_\delta}$  by the *renormalised* models  $\hat{\Pi}^{\xi_\delta}$  as defined above in (2.22) we obtain uniform bounds and can pass to the limit  $\delta \rightarrow 0$ .

More precisely, for any  $\varepsilon > 0$  and  $\delta > 0$ , let  $(\hat{\Pi}^{\sqrt{\varepsilon} \xi_\delta}, \hat{F}^{\sqrt{\varepsilon} \xi_\delta}) \in \mathcal{M}^{(3)}$  be the model defined recursively via (2.10c), (2.21), and (2.22) from  $\hat{\Pi}_y^{(\varepsilon)} \Xi = \sqrt{\varepsilon} \xi_\delta$ , and then extended to all of  $\mathcal{M}^{(3)}$  by Theorem 2.10. For  $\delta = 0$ , we write  $(\hat{\Pi}^{\sqrt{\varepsilon} \xi}, \hat{F}^{\sqrt{\varepsilon} \xi}) = (\hat{\Pi}^{\sqrt{\varepsilon} \xi_0}, \hat{F}^{\sqrt{\varepsilon} \xi_0}) \in \mathcal{M}^{(3)}$  for the limit as  $\delta \rightarrow 0$  of these models, which exists by Theorem 2.13. As above, for  $h \in \mathcal{H}$ , we write  $(\Pi^h, F^h)$  for the canonical model constructed from  $h$ . We then have the following result.

**Theorem 4.3.** *Let  $\varepsilon \mapsto \delta(\varepsilon) \geq 0$  be any function satisfying  $\lim_{\varepsilon \rightarrow 0} \delta(\varepsilon) = 0$ . Then the distributions of  $\{(\hat{\Pi}^{\sqrt{\varepsilon} \xi_{\delta(\varepsilon)}}, \hat{F}^{\sqrt{\varepsilon} \xi_{\delta(\varepsilon)}}), \varepsilon > 0\}$  under  $\mu$  satisfy a large deviation principle on  $\mathcal{M}^{(3)}$  with rate  $\varepsilon$  and rate function*

$$\mathcal{I}_{\text{Model}}((\Pi, F)) \stackrel{\text{def}}{=} \inf \left\{ \frac{1}{2} |h|_{\mathcal{H}}^2 : (\Pi^h, F^h) = (\Pi, F) \right\}.$$

*Proof.* For  $\delta > 0$ , define the  $\mathbf{E}$ -valued functions  $\hat{\Psi}_\delta(\xi)$  as  $\hat{\Psi}_{\delta; \tau}(\xi) = \hat{\Pi}^{\xi_\delta} \tau$ . It is shown in [Hai14c, Thms 10.7 and 10.22] that for every  $\tau \in \mathcal{W}_-^{(3)}$  these  $\hat{\Psi}_{\delta; \tau}(\xi)$  converge to the limit  $\hat{\Psi}_\tau(\xi)$  in the sense that

$$\lim_{\delta \rightarrow 0} \int \|\hat{\Psi}_{\delta; \tau}(\xi) - \hat{\Psi}_\tau(\xi)\|_{E_\tau}^2 \mu(d\xi) = 0.$$

Furthermore, each  $\hat{\Psi}_\tau(\xi)$  is given by iterated stochastic integrals against kernels  $\hat{\mathcal{W}}_{\tau; k}$  as, for example, above in (4.7). For each  $\tau \in \mathcal{W}_-^{(3)}$  the contribution in the highest Wiener chaos  $\mathcal{H}^{(K_\tau)}(\mu, E_\tau)$  is obtained by simply removing the convolution

with  $\varrho_\delta$  in the definitions of the the integral kernels  $K_\delta$ . For example, we have

$$\begin{aligned} \hat{\mathcal{W}}_{\mathfrak{Y},5}(\bar{y}, y, z_1, \dots, z_5) &\stackrel{\text{def}}{=} \int dx K(z_1 - x) K(z_2 - x) K(z_3 - x) \\ &\quad (K(x - \bar{y}) - K(x - y)) K(z_4 - y) K(z_5 - y). \end{aligned} \quad (4.10)$$

This is because the renormalisation procedure described in (2.21) and (2.22) only affects the contributions in  $\mathcal{H}^{(k)}(\mu, E_\tau)$  for  $k < K_\tau$ . This implies that for each  $\tau \in \mathcal{W}_-^{(3)}$  the homogeneous part  $(\hat{\Psi}_\tau)_{\text{hom}}(h)$  is given by  $\Pi^h \tau$ , the contribution at level  $\tau$  for the canonical model constructed from  $\Pi_z \Xi = h$ . Hence, for any function  $\delta(\varepsilon) \geq 0$  with  $\lim_{\varepsilon \rightarrow 0} \delta(\varepsilon) = 0$ , Theorem 3.5 yields a large deviation principle on  $\mathbf{E}$  for the sequence of random variables  $\hat{\Psi}_{\delta(\varepsilon)}^{(\varepsilon)}$ . From this we obtain a large deviation principle for the full renormalised models  $(\hat{\Pi}^{\sqrt{\varepsilon}\xi_{\delta(\varepsilon)}}, \hat{F}^{\sqrt{\varepsilon}\xi_{\delta(\varepsilon)}})$  using Theorem 2.10 as well as the contraction principle, Lemma 3.3, in the same way as in the proof of Theorem 4.2  $\square$

Applying Theorem 2.12 and the contraction principle, Lemma 3.3, once more, we immediately obtain our main result. In order to formulate it, we use the notation

$$\mathcal{C}_{u_0}([0, T], \mathcal{C}^\eta) \stackrel{\text{def}}{=} \{u \in \mathcal{C}([0, T], \mathcal{C}^\eta) : u(0, \cdot) = u_0\},$$

and  $\mathcal{X}_{u_0} \stackrel{\text{def}}{=} \mathcal{C}_{u_0}([0, T], \mathcal{C}^\eta) \cup \{\infty\}$ . As above we endow  $\mathcal{X}_{u_0}$  with a metric by postulating that  $\{\infty\}$  is at distance 1 from every point in  $\mathcal{C}_{u_0}([0, T], \mathcal{C}^\eta)$ .

**Theorem 4.4.** *Let  $\eta \in (-3/2, 1 - \frac{d}{2} - \kappa)$  as before and, for any  $\varepsilon, \delta > 0$  and for any initial datum  $u_0 \in \mathcal{C}^\eta$ , let  $\hat{u}_\delta^{(\varepsilon)}$  be the unique solution of the SPDE*

$$\partial_t u = \Delta u + (C + 3\varepsilon C_\delta^{(1)} - 9\varepsilon^2 C_\delta^{(2)})u - u^3 + \sqrt{\varepsilon}\xi_\delta \quad (4.11)$$

with initial datum  $u_0$ . Here  $C \in \mathbf{R}$  is arbitrary and the constants  $C_\delta^{(i)}$  are defined in (2.24) and (2.25). For  $\varepsilon > 0$ , let  $\hat{u}_0^{(\varepsilon)}$  be the limit as  $\delta \rightarrow 0$  of these solutions, constructed in theorems 2.12 and 2.13. (As before we set it equal to  $\infty$  if the maximal existence time is less or equal to  $T$ .)

Let  $\varepsilon \mapsto \delta(\varepsilon) \geq 0$  be a function with

$$\lim_{\varepsilon \rightarrow 0} \delta(\varepsilon) = 0.$$

Then the sequence  $\hat{u}_{\delta(\varepsilon)}^{(\varepsilon)}$  satisfies a large deviation principle in  $\mathcal{X}_{u_0}$  with rate function  $\mathcal{I}(\infty) = +\infty$  and

$$\mathcal{I}(u) = \frac{1}{2} \int_0^T \int_{\mathbf{T}^d} (\partial_t u - \Delta u - C u + u^3)^2 dx dt. \quad (4.12)$$

This is with the understanding that  $\mathcal{I}(u) = +\infty$  if  $u \notin L^3([0, T] \times \mathbf{T}^3)$  or if the distribution appearing in the right hand side of (4.12) is not a square integrable function.

*Remark 4.5.* The fact that the prefactor for  $C_\delta^{(1,2)}$  in (4.11) is given by  $\varepsilon$  and  $\varepsilon^2$  respectively follows from (3.10) since  $K_\vee = 2$  and  $K_{\mathfrak{Y}} = 4$ . This is also consistent with the definitions (10.35) and (10.41) in [Hai14c].

*Proof.* According to Theorem 2.13 the solution of (4.11) can be obtained by first applying the solution map  $\mathcal{S}_A$  to the renormalised model  $(\hat{\Pi}^{\sqrt{\varepsilon}\xi_{\delta(\varepsilon)}}, \hat{F}^{\sqrt{\varepsilon}\xi_{\delta(\varepsilon)}})$ .

According to Theorem 4.3 the renormalised models  $(\hat{\Pi}^{\sqrt{\varepsilon}\xi_{\delta(\varepsilon)}}, \hat{F}^{\sqrt{\varepsilon}\xi_{\delta(\varepsilon)}})$  satisfy a large deviation principle on  $\mathcal{M}^{(3)}$ . Denote by  $\mathcal{M}_{\mathcal{H}}^{(3)} \subset \mathcal{M}^{(3)}$  the set of models obtained by the canonical lift of some element  $h \in \mathcal{H}$ . Since the solutions to the Allen-Cahn equation driven by a noise term in  $\mathcal{H}$  are global, the map  $\mathcal{S}_A$  does not take the value  $\infty$  on  $\mathcal{M}_{\mathcal{H}}^{(3)}$ . Furthermore, as a consequence of Theorem 2.12, the solution operator  $\mathcal{S}_A$  (for fixed  $u_0$  and  $C$ ) is continuous from an open neighbourhood of  $\mathcal{M}_{\mathcal{H}}^{(3)}$  in  $\mathcal{M}^{(3)}$  into  $\mathcal{C}_{u_0}([0, T], \mathcal{C}^\eta)$ . Since  $\mathcal{M}_{\mathcal{H}}^{(3)} = \{\mathcal{I}_{\text{Model}} < +\infty\}$ , the contraction principle, Lemma 3.3, immediately implies that the  $\hat{u}_{\delta(\varepsilon)}^{(\varepsilon)}$  satisfy a large deviation principle on  $\mathcal{X}_{u_0} = \mathcal{C}_{u_0}([0, T], \mathcal{C}^\eta) \cup \{\infty\}$ .

The rate function is given by

$$\begin{aligned} \mathcal{I}(u) &= \inf\{\mathcal{I}_{\text{Model}}(\Pi, F) : \mathcal{S}_A(\Pi, F) = u\} \\ &= \inf\left\{\frac{1}{2}|h|_{\mathcal{H}}^2 : \mathcal{S}_A(\Pi^h, F^h) = u\right\}. \end{aligned}$$

According to Theorem 2.13, for  $h \in \mathcal{H}$  the function  $\mathcal{S}_A(\Pi^h, F^h)$  is simply the classical solution to

$$\partial_t u = \Delta u + Cu - u^3 + h.$$

Standard parabolic regularity theory implies that for  $h \in L^2([0, T] \times \mathbf{T}^3)$  (and for such an irregular choice of initial condition)  $u$  will at least attain values in  $L^3([0, T] \times \mathbf{T}^3)$ . Conversely, if  $u$  can be obtained in this way from  $h$  we can simply recover  $h$  by setting

$$h = \partial_t u - (\Delta u + Cu - u^3).$$

This concludes the argument.  $\square$

*Remark 4.6.* Actually, we do not expect the renormalised solutions to explode either for  $d = 2$  or  $d = 3$ . Our argument does not imply this, but we can conclude that the probability of finite time explosion decays faster than  $e^{-c\varepsilon^{-1}}$  for any  $c \in (0, \infty)$ .

In a very similar way, Theorem 4.3 implies the following large deviation principle for solutions of the SPDE without renormalisation:

**Theorem 4.7.** *Let  $\eta \in (-3/2, 1 - \frac{d}{2} - \kappa)$  as before. For any  $\varepsilon, \delta > 0$  and for any initial datum  $u^0 \in \mathcal{C}^\eta$  let  $u_\delta^{(\varepsilon)}$  be the unique (global) solution of the SPDE*

$$\partial_t u = \Delta u + Cu - u^3 + \sqrt{\varepsilon}\xi_\delta, \quad (4.13)$$

with initial datum  $u_0$ .

Let  $\varepsilon \mapsto \delta(\varepsilon) > 0$  be a function with  $\lim_{\varepsilon \rightarrow 0} \delta(\varepsilon) = 0$  and

$$\lim_{\varepsilon \rightarrow 0} \varepsilon \delta^{-1} = \lambda^2 \in [0, \infty). \quad (4.14)$$

Then the  $u_{\delta(\varepsilon)}^{(\varepsilon)}$  satisfy a large deviation principle in  $\mathcal{C}([0, T], \mathcal{C}^\eta)$  with rate function

$$\mathcal{I}(u) = \frac{1}{2} \int_0^T \int_{\mathbf{T}^d} (\partial_t u - \Delta u + C_\lambda u + u^3)^2 dx dt.$$

where  $C_\lambda = C - 3\lambda^2 \int_{\mathbf{R}^4} (P \star \varrho)(t, x)^2 dt dx$ .

*Remark 4.8.* Recall that  $P$  denotes the heat kernel on the 3-dimensional torus without any truncation. (See also Remark 2.14.)

*Proof.* We start by rewriting (4.13) as

$$\partial_t u = \Delta u + \tilde{C}(\varepsilon)u - (u^3 - (3\varepsilon C_\delta^{(1)} - 9\varepsilon^2 C_\delta^{(2)})u) + \sqrt{\varepsilon}\xi_\delta$$

where  $\tilde{C}(\varepsilon) = C - 3\varepsilon C_\delta^{(1)} + 9\varepsilon^2 C_\delta^{(2)}$ . According to Theorem 2.13 the solution to this equation can be obtained by applying the solution operator  $\mathcal{S}_A$  corresponding to the choice  $C = \tilde{C}(\varepsilon)$  to the renormalised model  $(\hat{\Pi}^{\sqrt{\varepsilon}\xi_\delta(\varepsilon)}, \hat{F}^{\sqrt{\varepsilon}\xi_\delta(\varepsilon)})$ .

According to the assumption (4.14) and to the definition (2.24) and (2.25) of the constants  $C_\delta^{(1)}$  and  $C_\delta^{(2)}$  the  $\tilde{C}(\varepsilon)$  converge to  $C_\lambda$ . Now the argument proceeds like the proof of Theorem 4.4, noting only that the solution map is locally uniformly continuous in the choice of  $C$  when applying the contraction principle, Lemma 3.3.  $\square$

*Remark 4.9.* The analogous result for  $d = 2$  can be obtained in a similar way. There, according to (2.23) we get for the value of the modified constant  $C_\lambda = C - \lambda^2 \frac{3}{4\pi}$  under the assumption that  $\lim_{\varepsilon \rightarrow 0} \varepsilon \log \delta(\varepsilon)^{-1} = \lambda^2 \in [0, \infty)$ . In particular, the value of  $C_\lambda$  does not depend on the choice of mollifying kernel  $\varrho$ .

*Remark 4.10.* For  $d = 3$  we could also consider schemes of the form

$$\partial_t u = \Delta u + \tilde{C}(\varepsilon)u - (u^3 - 3\varepsilon C_\delta^{(1)}u) + \sqrt{\varepsilon}\xi_\delta,$$

i.e. schemes that ignore the logarithmic sub divergence. In this case we recover a large deviation result with rate function depending on  $\lim_{\varepsilon \rightarrow 0} \varepsilon^2 \log \delta(\varepsilon)^{-1} = \lambda^2 \in [0, \infty)$ .

*Remark 4.11.* Even though our statement does not make reference to regularity structures, the proof uses heavily the fact that there is a non-trivial renormalised limit as  $\delta \rightarrow 0$  for fixed  $\varepsilon > 0$ . In particular, we do not know if a similar statement (but with stronger assumptions on how  $\delta(\varepsilon) \rightarrow 0$ ) holds true in four spatial dimensions where no renormalised solutions are available.

## REFERENCES

- [Aid09] S. AIDA. Semi-classical limit of the lowest eigenvalue of a Schrödinger operator on a Wiener space. II.  $P(\varphi)_2$ -model on a finite volume. *J. Funct. Anal.* **256**, no. 10, (2009), 3342–3367. doi:10.1016/j.jfa.2008.10.001.
- [Aid12] S. AIDA. Tunneling for spatially cut-off  $P(\varphi)_2$ -Hamiltonians. *J. Funct. Anal.* **263**, no. 9, (2012), 2689–2753. arXiv:1104.0486. doi:10.1016/j.jfa.2012.08.017.
- [BLZ14] F. BOUCHET, J. LAURIE, and O. ZABORONSKI. Langevin dynamics, large deviations and instantons for the quasi-geostrophic model and two-dimensional euler equations. *ArXiv e-prints* (2014). arXiv:1403.0216.
- [Bor78] C. BORELL. Tail probabilities in Gauss space. In *Vector Space Measures and Applications I*, 73–82. Springer, 1978.
- [Bor84a] C. BORELL. On polynomial chaos and integrability. *Probab. Math. Statist* **3**, no. 2, (1984), 191–203.
- [Bor84b] C. BORELL. On the Taylor series of a Wiener polynomial. In *Seminar Notes on multiple stochastic integration, polynomial chaos and their integration. Case Western Reserve University, Cleveland*. 1984.
- [CF11] S. CERRAI and M. FREIDLIN. Approximation of quasi-potentials and exit problems for multidimensional RDE's with noise. *Trans. Amer. Math. Soc.* **363**, no. 7, (2011), 3853–3892. doi:10.1090/S0002-9947-2011-05352-3.
- [DPD02] G. DA PRATO and A. DEBUSSCHE. Two-dimensional Navier-Stokes equations driven by a space-time white noise. *J. Funct. Anal.* **196**, no. 1, (2002), 180–210. doi:10.1006/jfan.2002.3919.

- [DPD03] G. DA PRATO and A. DEBUSSCHE. Strong solutions to the stochastic quantization equations. *Ann. Probab.* **31**, no. 4, (2003), 1900–1916. doi:10.1214/aop/1068646370.
- [DS89] J.-D. DEUSCHEL and D. W. STROOCK. *Large deviations*, vol. 137 of *Pure and Applied Mathematics*. Academic Press Inc., Boston, MA, 1989.
- [ERVE04] W. E. W. REN, and E. VANDEN-EIJNDEN. Minimum action method for the study of rare events. *Comm. Pure Appl. Math.* **57**, no. 5, (2004), 637–656. doi:10.1002/cpa.20005.
- [FJL82] W. G. FARIS and G. JONA-LASINIO. Large fluctuations for a nonlinear heat equation with noise. *J. Phys. A* **15**, no. 10, (1982), 3025–3055. doi:10.1088/0305-4470/15/10/011.
- [FV07] P. FRIZ and N. VICTOIR. Large deviation principle for enhanced Gaussian processes. *Ann. Inst. H. Poincaré Probab. Statist.* **43**, no. 6, (2007), 775–785. arXiv:math/0512213. doi:10.1016/j.anihpb.2006.11.002.
- [Hai14a] M. HAIRER. Introduction to Regularity Structures. *ArXiv e-prints* (2014). arXiv:1401.3014. To appear in *Braz. J. Prob. Stat.*
- [Hai14b] M. HAIRER. Singular stochastic PDEs. *ArXiv e-prints* (2014). arXiv:1403.6353. To appear in *Proc. ICM*.
- [Hai14c] M. HAIRER. A theory of regularity structures. *Invent. Math.* **198**, no. 2, (2014), 269–504. arXiv:1303.5113. doi:10.1007/s00222-014-0505-4.
- [HH77] P. C. HOHENBERG and B. I. HALPERIN. Theory of dynamic critical phenomena. *Reviews of Modern Physics* **49**, no. 3, (1977), 435. doi:10.1103/RevModPhys.49.435.
- [HRW12] M. HAIRER, M. D. RYSER, and H. WEBER. Triviality of the 2D stochastic Allen-Cahn equation. *Electron. J. Probab.* **17**, (2012), no. 39, 14. arXiv:1201.3089. doi:10.1214/EJP.v17-1731.
- [JLM90] G. JONA-LASINIO and P. K. MITTER. Large deviation estimates in the stochastic quantization of  $\varphi_2^4$ . *Comm. Math. Phys.* **130**, no. 1, (1990), 111–121. doi:10.1007/BF02099877.
- [KORVE07] R. KOHN, F. OTTO, M. G. REZNIKOFF, and E. VANDEN-EIJNDEN. Action minimization and sharp-interface limits for the stochastic Allen-Cahn equation. *Comm. Pure Appl. Math.* **60**, no. 3, (2007), 393–438. doi:10.1007/s00526-005-0370-5.
- [Led90] M. LEDOUX. A note on large deviations for Wiener chaos. In *Séminaire de Probabilités, XXIV, 1988/89*, vol. 1426 of *Lecture Notes in Math.*, 1–14. Springer, Berlin, 1990.
- [Led96] M. LEDOUX. Isoperimetry and Gaussian analysis. In *Lectures on probability theory and statistics*, 165–294. Springer, 1996.
- [LQZ02] M. LEDOUX, Z. QIAN, and T. ZHANG. Large deviations and support theorem for diffusion processes via rough paths. *Stochastic Process. Appl.* **102**, no. 2, (2002), 265–283. doi:10.1016/S0304-4149(02)00176-X.
- [MSS06] A. MILLET and M. SANZ-SOLÉ. Large deviations for rough paths of the fractional Brownian motion. *Ann. Inst. H. Poincaré Probab. Statist.* **42**, no. 2, (2006), 245–271. doi:10.1016/j.anihpb.2005.04.003.
- [MWNPA92] E. MAYER-WOLF, D. NUALART, and V. PÉREZ-ABREU. Large deviations for multiple Wiener-Itô integral processes. In *Séminaire de Probabilités XXVI*, 11–31. Springer, 1992.
- [Nev75] J. NEVEU. *Discrete-parameter martingales*. North-Holland Publishing Co., Amsterdam-Oxford; American Elsevier Publishing Co., Inc., New York, revised ed., 1975. Translated from the French by T. P. Speed, North-Holland Mathematical Library, Vol. 10.
- [Nua06] D. NUALART. *The Malliavin calculus and related topics*. Springer, 2006.

MARTIN HAIRER, UNIVERSITY OF WARWICK  
*E-mail address:* M.Hairer@Warwick.ac.uk

HENDRIK WEBER, UNIVERSITY OF WARWICK  
*E-mail address:* hendrik.weber@warwick.ac.uk