

# A SCALING LIMIT FROM THE WAVE MAP TO THE HEAT FLOW INTO $\mathbb{S}^2$

NING JIANG, YI-LONG LUO, SHAOJUN TANG, AND ARGHIR ZARNESCU

ABSTRACT. In this paper we study a limit connecting a scaled wave map with the heat flow into the unit sphere  $\mathbb{S}^2$ . We show quantitatively how that the two equations are connected by means of an initial layer correction. This limit is motivated as a first step into understanding the limit of zero inertia for the hyperbolic-parabolic Ericksen-Leslie's liquid crystal model.

## 1. INTRODUCTION

**1.1. Wave map and heat flow.** We consider a hyperbolic system for functions  $\mathbf{d} : \mathbb{R}^+ \times \mathbb{R}^3 \rightarrow \mathbb{S}^2$ :

$$\partial_t \mathbf{d} = -\square \mathbf{d} + (|\nabla \mathbf{d}|^2 - |\partial_t \mathbf{d}|^2) \mathbf{d}, \quad (1.1)$$

subject to initial data: for any  $x \in \mathbb{R}^3$ ,

$$\mathbf{d}|_{t=0} = \mathbf{d}^0(x) \in \mathbb{S}^2, \quad \partial_t \mathbf{d}|_{t=0} = \tilde{\mathbf{d}}^0(x) \in \mathbb{R}^3, \quad \mathbf{d}^0(x) \cdot \tilde{\mathbf{d}}^0(x) = 0, \quad (1.2)$$

where  $\square = \partial_{tt} - \Delta$  is the standard wave operator, and the compatibility condition  $\mathbf{d}^0 \cdot \tilde{\mathbf{d}}^0 = 0$  on the initial data is due to the fact that  $|\mathbf{d}^0| \equiv 1$ .

The system (1.1) is a wave map from  $\mathbb{R}^3$  to the unit sphere  $\mathbb{S}^2$ , with a damping term  $\partial_t \mathbf{d}$ . One way of interpreting this system is as follows: setting the righthand side of (1.1) equal to 0, we obtain  $\square \mathbf{d} = (|\nabla \mathbf{d}|^2 - |\partial_t \mathbf{d}|^2) \mathbf{d}$ . This is the well-known *wave map*, which can be characterized variationally as a critical point of the functional

$$\mathcal{A}(\mathbf{d}) = \frac{1}{2} \iint (|\nabla \mathbf{d}|^2 - |\partial_t \mathbf{d}|^2) dx dt, \quad (1.3)$$

among maps  $\mathbf{d}$  satisfying the target constraint,  $\mathbf{d} : \mathbb{R}^+ \times \mathbb{R}^3 \rightarrow \mathbb{S}^2$ . Thus the full system (1.1) can be viewed as a “gradient flow” of the functional (1.3).

Another gradient flow can be obtained by formally dropping some terms out of the previous system, and obtaining the *heat flow*

$$\partial_t \mathbf{d} = \Delta \mathbf{d} + |\nabla \mathbf{d}|^2 \mathbf{d}. \quad (1.4)$$

Similarly as before setting the right-hand side equal to zero we obtain the equations for the harmonic map from  $\mathbb{R}^3$  to the unit sphere  $\mathbb{S}^2$  namely

$$\Delta \mathbf{d} + |\nabla \mathbf{d}|^2 \mathbf{d} = 0, \quad (1.5)$$

which is a critical point of the energy functional

$$E(\mathbf{d}) = \frac{1}{2} \int |\nabla \mathbf{d}|^2 dx. \quad (1.6)$$

There exist deep relations between the two systems, (1.1) and (1.4) and one way to see this is by considering the following parabolic scaling:

$$\mathbf{d}^\varepsilon(t, x) := \mathbf{d}\left(\frac{t}{\varepsilon}, \frac{x}{\sqrt{\varepsilon}}\right), \quad (1.7)$$

Then,  $\mathbf{d}^\varepsilon$  satisfies the following scaled wave map:

$$\partial_t \mathbf{d}^\varepsilon = -(\varepsilon \partial_{tt} - \Delta) \mathbf{d}^\varepsilon + (|\nabla \mathbf{d}^\varepsilon|^2 - \varepsilon |\partial_t \mathbf{d}^\varepsilon|^2) \mathbf{d}^\varepsilon, \quad (1.8)$$

on  $\mathbb{R}^+ \times \mathbb{R}^3$ . For this scaled system we take the initial values independent of  $\varepsilon$ , namely:

$$\mathbf{d}^\varepsilon|_{t=0} = \mathbf{d}^{in}(x) \in \mathbb{S}^2, \quad \partial_t \mathbf{d}^\varepsilon|_{t=0} = \tilde{\mathbf{d}}^{in}(x) \in \mathbb{R}^3. \quad (1.9)$$

The finite-time behaviour of the limit  $\varepsilon \rightarrow 0$  for the system (1.8) with initial data (1.9) is the focus of this paper. It is easy to see that letting  $\varepsilon = 0$  in (1.8) will formally give the heat flow (1.4). However, a refined analysis and the introduction of an *initial layer* is needed in order to overcome the singular character of this limit and understand the relationship between the system (1.8) and its formal limit, as it will be seen in the Theorem 1.1 below.

**Notations and conventions:** Throughout this paper, we use the following standard notations:

$$|\mathbf{d}|_{L^p}^p = \int_{\mathbb{R}^3} |\mathbf{d}|^p dx, \quad |\mathbf{d}|_{H^k} = \sum_{\gamma \leq k} |\nabla^\gamma \mathbf{d}|_{L^2}, \quad |\mathbf{d}|_{\dot{H}^k} = \sum_{1 \leq \gamma \leq k} |\nabla^\gamma \mathbf{d}|_{L^2}.$$

Additionally, for the Hilbert space  $L^2 \equiv L^2(dx, \mathbb{R}^3)$ , we use the following notation to denote the standard inner product:

$$\langle f, g \rangle = \int_{\mathbb{R}^3} fg dx.$$

Furthermore, if there is a generic constant  $C > 0$  such that the inequality  $f(t) \leq Cg(t)$  holds for all  $t \geq 0$ , we denote this inequality by

$$f \lesssim g.$$

**1.2. Ericksen-Leslie's hyperbolic liquid crystal model.** Our motivation for considering the previously mentioned limit comes from the hydrodynamic theory of nematic liquid crystals.

The most widely accepted equations of nematics were proposed by Ericksen [5, 6, 7] and Leslie [10, 11] in the 1960's (see for more details Section 5.1 of [12]). The general hyperbolic-parabolic Ericksen-Leslie system consists of an equation for velocity  $u$  of the centers of mass of the rod-like molecules, coupled with an equation for the direction  $d$  of these molecules. More specifically we have the following equations (in non-dimensional form):

$$\begin{cases} \partial_t \mathbf{u} + \mathbf{u} \cdot \nabla \mathbf{u} - \frac{1}{2} \mu_4 \Delta \mathbf{u} + \nabla p = -\operatorname{div}(\nabla \mathbf{d} \odot \nabla \mathbf{d}) + \operatorname{div} \tilde{\sigma}, \\ \operatorname{div} \mathbf{u} = 0, \\ \varepsilon D_{\mathbf{u}}^2 \mathbf{d} = \Delta \mathbf{d} + \gamma \mathbf{d} + \lambda_1 (D_{\mathbf{u}} \mathbf{d} - \mathbf{B} \mathbf{d}) + \lambda_2 \mathbf{A} \mathbf{d} \end{cases} \quad (1.10)$$

on  $\mathbb{R}^+ \times \mathbb{R}^3$  with constraint  $|\mathbf{d}| = 1$ , where  $\mathbf{A} = \frac{1}{2}(\nabla \mathbf{u} + \nabla \mathbf{u}^\top)$  and  $\mathbf{B} = \frac{1}{2}(\nabla \mathbf{u} - \nabla \mathbf{u}^\top)$ ,  $D_{\mathbf{u}} f = \partial_t f + \mathbf{u} \cdot \nabla f$  is the material derivative of  $f$  respect to the vector  $\mathbf{u}$ ,  $D_{\mathbf{u}}^2 \mathbf{d} = \partial_t D_{\mathbf{u}} \mathbf{d} + \mathbf{u} \cdot \nabla D_{\mathbf{u}} \mathbf{d}$ . The Lagrangian  $\gamma$  that enforces the unit-length constraint  $|\mathbf{d}| = 1$  is given by

$$\gamma \equiv \gamma(\mathbf{u}, \mathbf{d}, D_{\mathbf{u}} \mathbf{d}) = -\varepsilon |D_{\mathbf{u}} \mathbf{d}|^2 + |\nabla \mathbf{d}|^2 - \lambda_2 \mathbf{d}^\top \mathbf{A} \mathbf{d}.$$

The stress tensor appearing in the equation for  $u$  is given by:

$$\begin{aligned} \tilde{\sigma}_{ij} \equiv (\tilde{\sigma}(\mathbf{u}, \mathbf{d}, D_{\mathbf{u}} \mathbf{d}))_{ij} &= \mu_1 \mathbf{d}_k \mathbf{d}_p \mathbf{A}_{kp} \mathbf{d}_i \mathbf{d}_j + \mu_2 \mathbf{d}_j ((D_{\mathbf{u}} \mathbf{d})_i + \mathbf{B}_{ki} \mathbf{d}_k) \\ &\quad + \mu_3 \mathbf{d}_i ((D_{\mathbf{u}} \mathbf{d})_j + \mathbf{B}_{kj} \mathbf{d}_k) + \mu_5 \mathbf{d}_j \mathbf{d}_k \mathbf{A}_{ki} + \mu_6 \mathbf{d}_i \mathbf{d}_k \mathbf{A}_{kj}. \end{aligned}$$

The constant  $\varepsilon > 0$  measures the inertial effects. The constants  $\mu_i$  ( $1 \leq i \leq 6$ ) are known as Leslie coefficients and one has  $\mu_4 > 0$ . Furthermore, we have:

$$\lambda_1 = \mu_2 - \mu_3, \quad \lambda_2 = \mu_5 - \mu_6, \quad \mu_2 + \mu_3 = \mu_6 - \mu_5,$$

where the last relation is called *Parodi* relation. For the more background and derivation of (1.10), see [10] and [9].

For any fixed  $\varepsilon > 0$ , in [9] the first two named authors of the current paper proved the local well-posedness of the system (1.10) under assumptions on the Leslie coefficients which

ensure the dissipativity of the basic energy law, and global well-posedness with small initial data under further damping effect, i.e.  $\lambda_1 < 0$ .

As noted in the ‘‘Conclusion’’ section of [9], the inertial constant  $\varepsilon > 0$  is, physically, in most common non-dimensionalisations and materials, very small. Formally, letting  $\varepsilon = 0$  will give the parabolic Ericksen-Leslie system which is basically a coupling of Navier-Stokes equations and an extension of the heat flow to the unit sphere. However it is a very challenging task to obtaining estimates uniform in  $\varepsilon$  for the full system (1.10), in order to understand the limit  $\varepsilon \rightarrow 0$ . In the current paper, the problem we consider what appears as a simple instance of this general problem, namely the case where the bulk velocity  $\mathbf{u} = \mathbf{0}$  and the coefficient  $\lambda_1 = -1$  in (1.10). For this case, the system (1.10) is reduced to the scaled wave map (1.8), i.e. the wave map (1.8) with a damping can be regarded as an Ericksen-Leslie’s liquid crystal flow unaffected by the fluid velocity.

**1.3. Initial layer and the main result.** As mentioned in the previous two subsections the formal limit of the equation (1.8), obtained by setting  $\varepsilon = 0$  is provided by the heat flow for functions with values into  $\mathbb{S}^2$ :

$$\partial_t \mathbf{d}_0 = \Delta \mathbf{d}_0 + |\nabla \mathbf{d}_0|^2 \mathbf{d}_0, \mathbf{d}_0 \in \mathbb{S}^2, \quad (1.11)$$

The limit we consider is a singular limit, as the character of the equations changes, from a hyperbolic-type system for  $\varepsilon > 0$  to a parabolic system for  $\varepsilon = 0$ . An immediate manifestation of the difference between the two types of equations is related to the initial conditions, which for the limit equation take the form:

$$\mathbf{d}_0|_{t=0} = \mathbf{d}^{in}(x) \in \mathbb{S}^2. \quad (1.12)$$

Thus, we note that the wave map is a system of hyperbolic equations with two initial conditions, while the heat flow is a parabolic system with only one initial condition. Usually the solution of the heat flow does not satisfy the second initial condition in (1.9). This disparity between the initial conditions of the wave map (1.8) and of the heat flow (1.11) indicates that in one should expect an ‘‘initial layer’’ in time, appearing in the limiting process  $\varepsilon \rightarrow 0$ . A formal derivation (postponed for later, in Section 2) indicates that this should be of the form:

$$\begin{aligned} \mathbf{d}_0^I\left(\frac{t}{\varepsilon}, x\right) &= -\varepsilon(\tilde{\mathbf{d}}^{in}(x) - \partial_t \mathbf{d}_0(0, x)) \exp\left(-\frac{t}{\varepsilon}\right) \\ &= -\varepsilon(\tilde{\mathbf{d}}^{in}(x) - \Delta \mathbf{d}^{in}(x) - |\nabla \mathbf{d}^{in}(x)|^2 \mathbf{d}^{in}(x)) \exp\left(-\frac{t}{\varepsilon}\right) \\ &\equiv -\varepsilon \mathbf{D}(x) \exp\left(-\frac{t}{\varepsilon}\right), \end{aligned} \quad (1.13)$$

where  $\mathbf{D}(x)$  is defined as

$$\mathbf{D}(x) \equiv \tilde{\mathbf{d}}^{in}(x) - \Delta \mathbf{d}^{in}(x) - |\nabla \mathbf{d}^{in}(x)|^2 \mathbf{d}^{in}(x).$$

Our study of the limit from the wave map (1.8) to the heat flow (1.5) is inspired by the classical approach of Caffisch on the compressible Euler limit of the Boltzmann equation [1]. This approach is based on the Hilbert expansion in which the leading term is given by solutions of the limit equation. The Caffisch’s approach assumes that a solution of the limiting equation (which in our case is the heat flow (1.5)) is known beforehand. Then the solution to the original equation (which in our case is the wave map (1.8)) can be constructed around the limiting equation with perturbations as expansions in powers of  $\varepsilon$ . Based on the arguments above and the formal analysis in Section 2, in the expansions, besides the heat flow, the leading term should also include an initial layer. More specifically, we take the following ansatz of the solution  $\mathbf{d}^\varepsilon$  to the system (1.8):

$$\mathbf{d}^\varepsilon(t, x) = \mathbf{d}_0(t, x) + \mathbf{d}_0^I\left(\frac{t}{\varepsilon}, x\right) + \sqrt{\varepsilon} \mathbf{d}_R^\varepsilon(t, x), \quad (1.14)$$

where  $\mathbf{d}_0(t, x)$  obeys the heat flow (1.11) and the initial layer  $\mathbf{d}_0^I(\frac{t}{\varepsilon}, x)$  is defined in (1.13). Plugging (1.14) into the system (1.8), the remainder term  $\mathbf{d}_R^\varepsilon(t, x)$  must satisfy the system

$$\partial_{tt}\mathbf{d}_R^\varepsilon + \frac{1}{\varepsilon}\partial_t\mathbf{d}_R^\varepsilon - \frac{1}{\varepsilon}\Delta\mathbf{d}_R^\varepsilon = \mathcal{S}(\mathbf{d}_R^\varepsilon) + \mathcal{R}(\mathbf{d}_R^\varepsilon) \quad (1.15)$$

with the initial conditions

$$\mathbf{d}_R^\varepsilon(0, x) = \sqrt{\varepsilon}\mathbf{D}(x), \quad \partial_t\mathbf{d}_R^\varepsilon(0, x) = 0, \quad (1.16)$$

where the singular term  $\mathcal{S}(\mathbf{d}_R^\varepsilon)$  is

$$\begin{aligned} \mathcal{S}(\mathbf{d}_R^\varepsilon) &= -\frac{1}{\sqrt{\varepsilon}}\partial_{tt}\mathbf{d}_0 - \frac{1}{\sqrt{\varepsilon}}\Delta\mathbf{D}(x)\exp(-\frac{t}{\varepsilon}) + \frac{1}{\varepsilon}|\nabla\mathbf{d}_0|^2\mathbf{d}_R^\varepsilon + \frac{1}{\sqrt{\varepsilon}}|\nabla\mathbf{d}_R^\varepsilon|^2\mathbf{d}_0 \\ &\quad - \frac{1}{\sqrt{\varepsilon}}|\partial_t\mathbf{d}_0 + \mathbf{D}(x)\exp(-\frac{t}{\varepsilon})|^2\mathbf{d}_0 - \frac{1}{\sqrt{\varepsilon}}|\nabla\mathbf{d}_0|^2\mathbf{D}(x)\exp(-\frac{t}{\varepsilon}) \\ &\quad + \frac{2}{\sqrt{\varepsilon}}(\nabla\mathbf{d}_0 \cdot \nabla\mathbf{d}_R^\varepsilon)\mathbf{d}_R^\varepsilon - \frac{2}{\sqrt{\varepsilon}}(\nabla\mathbf{d}_0 \cdot \nabla\mathbf{D}(x))\exp(-\frac{t}{\varepsilon})\mathbf{d}_0 + \frac{2}{\varepsilon}(\nabla\mathbf{d}_0 \cdot \nabla\mathbf{d}_R^\varepsilon)\mathbf{d}_0, \end{aligned}$$

and the regular term  $\mathcal{R}(\mathbf{d}_R^\varepsilon)$  is

$$\begin{aligned} \mathcal{R}(\mathbf{d}_R^\varepsilon) &= -|\partial_t\mathbf{d}_0 + \mathbf{D}(x)\exp(-\frac{t}{\varepsilon}) + \sqrt{\varepsilon}\partial_t\mathbf{d}_R^\varepsilon|^2[-\sqrt{\varepsilon}\mathbf{D}(x)\exp(-\frac{t}{\varepsilon}) + \mathbf{d}_R^\varepsilon] \\ &\quad - [2(\partial_t\mathbf{d}_0 + \mathbf{D}(x)\exp(-\frac{t}{\varepsilon})) \cdot \partial_t\mathbf{d}_R^\varepsilon + \sqrt{\varepsilon}|\partial_t\mathbf{d}_R^\varepsilon|^2]\mathbf{d}_0 - 2(\nabla\mathbf{D}(x) \cdot \nabla\mathbf{d}_R^\varepsilon)\exp(-\frac{t}{\varepsilon})\mathbf{d}_0 \\ &\quad + \sqrt{\varepsilon}|\nabla\mathbf{D}(x)|^2\exp(-\frac{2t}{\varepsilon})\mathbf{d}_0 + |\nabla\mathbf{d}_R^\varepsilon|^2\mathbf{d}_R^\varepsilon - 2(\nabla\mathbf{d}_0 \cdot \nabla\mathbf{D}(x))\exp(-\frac{t}{\varepsilon})\mathbf{d}_R^\varepsilon \\ &\quad - 2(\nabla\mathbf{d}_0 \cdot \nabla\mathbf{d}_R^\varepsilon)\mathbf{D}(x)\exp(-\frac{t}{\varepsilon}) - 2\sqrt{\varepsilon}(\nabla\mathbf{D}(x) \cdot \nabla\mathbf{d}_R^\varepsilon)\exp(-\frac{t}{\varepsilon})\mathbf{d}_R^\varepsilon \\ &\quad - \sqrt{\varepsilon}|\nabla\mathbf{d}_R^\varepsilon|^2\mathbf{D}(x)\exp(-\frac{t}{\varepsilon}) + \varepsilon|\nabla\mathbf{D}(x)|^2\exp(-\frac{2t}{\varepsilon})\mathbf{d}_R^\varepsilon - \varepsilon^{\frac{3}{2}}|\nabla\mathbf{D}(x)|^2\exp(-\frac{3t}{\varepsilon})\mathbf{D}(x) \\ &\quad + 2\sqrt{\varepsilon}(\nabla\mathbf{d}_0 \cdot \nabla\mathbf{D}(x))\exp(-\frac{2t}{\varepsilon})\mathbf{D}(x) + 2\varepsilon(\nabla\mathbf{D}(x) \cdot \nabla\mathbf{d}_R^\varepsilon)\exp(-\frac{2t}{\varepsilon})\mathbf{D}(x). \end{aligned}$$

According to Eells-Sampson's classical result in [4], for the heat flow (1.11) on the unit sphere  $\mathbb{S}^2$ , one can have the following results of local well-posedness:

**Proposition 1.1.** *For any given  $\mathbf{d}^{in} \in \mathbb{S}^2$  satisfying  $\mathbf{d}^{in} \in \dot{H}^k(\mathbb{R}^3)$  for any integer  $k > 2$ , there exists a time  $T = T(|\mathbf{d}^{in}|_{\dot{H}^k}) > 0$  such that (1.11) admits a unique classical solution  $\mathbf{d}_0 \in L^\infty(0, T; \dot{H}^k) \cap L^2(0, T; \dot{H}^{k+1})$ . Moreover, there is a constant  $C^* = C^*(|\mathbf{d}^{in}|_{\dot{H}^k}, T) > 0$  such that the solution  $\mathbf{d}_0$  satisfies*

$$|\mathbf{d}_0|_{L^\infty(0, T; \dot{H}^k)}^2 + |\nabla\mathbf{d}_0|_{L^2(0, T; H^k)}^2 \leq C^*.$$

The proof can be found in Chapter 5 in the book [13].

Now we state the main result of this paper:

**Theorem 1.1.** *We consider vector fields  $\mathbf{d}^{in} : \mathbb{R}^3 \rightarrow \mathbb{S}^2$  and  $\tilde{\mathbf{d}}^{in} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ , satisfying the compatibility condition  $\mathbf{d}^{in} \cdot \tilde{\mathbf{d}}^{in} \equiv 0$ . Assume that  $\nabla\mathbf{d}^{in} \in H^6$ ,  $\tilde{\mathbf{d}}^{in} \in H^5$ , and let  $T > 0$  be the time interval of existence of the solution of the heat flow (1.11) with initial condition  $\mathbf{d}^{in}$ , determined in Proposition 1.1.*

*Then, there exists an  $\varepsilon_0 \equiv \varepsilon_0(|\nabla\mathbf{d}^{in}|_{H^6}, |\tilde{\mathbf{d}}^{in}|_{H^5}, T) \in (0, \frac{1}{2})$  such that for all  $\varepsilon \in (0, \varepsilon_0)$  we have that on the interval  $[0, T]$  the wave map equation (1.8) with the initial conditions (1.9) admits a unique solution with the form (1.14), i.e.,*

$$\mathbf{d}^\varepsilon(t, x) = \mathbf{d}_0(t, x) + \mathbf{d}_0^I(\frac{t}{\varepsilon}, x) + \sqrt{\varepsilon}\mathbf{d}_R^\varepsilon(t, x),$$

*where  $\mathbf{d}_0$  is the solution of the heat flow (1.11) with initial condition  $\mathbf{d}^{in}$  and  $\mathbf{d}_0^I(\frac{t}{\varepsilon}, x)$  is the initial layer (1.13). Moreover, there exists a positive constant  $C_0 = C_0(\mathbf{d}^{in}, \tilde{\mathbf{d}}^{in}, T) > 0$ , such that the remainder term  $\mathbf{d}_R^\varepsilon$  satisfies the bound*

$$|\partial_t\mathbf{d}_R^\varepsilon|_{L^\infty(0, T; H^2)}^2 + \frac{1}{\varepsilon}|\mathbf{d}_R^\varepsilon|_{L^\infty(0, T; H^3)}^2 \leq C_0 \quad (1.17)$$

*for all  $\varepsilon \in (0, \varepsilon_0)$ .*

**Remark 1.1.** *The rate of convergence we obtain is optimal. Indeed, in order to see this, it suffices to note that the limit we study contains as a particular case the linear scalar case of the singular limit of the damped wave equation to the heat equation.*

*Indeed, let us consider a solution of the scalar damped wave equation:*

$$\epsilon \partial_{tt} \theta^\epsilon + \partial_t \theta^\epsilon = \Delta \theta^\epsilon \quad (1.18)$$

for  $\theta^\epsilon : \mathbb{R} \times \mathbb{R}^3 \rightarrow \mathbb{R}$  with initial datas:

$$\theta^\epsilon(0, x) = \theta_0(x), \partial_t \theta^\epsilon(0, x) = \theta_1(x) \quad (1.19)$$

all smooth functions.

Also consider the solution of the heat equation:

$$\partial_t \theta^0 = \Delta \theta^0 \quad (1.20)$$

for  $\theta^0 : \mathbb{R} \times \mathbb{R}^3 \rightarrow \mathbb{R}$  with initial data:

$$\theta^0(0, x) = \theta_0(x), \quad (1.21)$$

all smooth functions.

Denoting  $n^0(t, x) := (\cos \theta^0(t, x), \sin \theta^0(t, x), 0)$  and  $n^\epsilon(t, x) := (\cos \theta^\epsilon(t, x), \sin \theta^\epsilon(t, x), 0)$  we have that  $n^0$  is a solution of the heat-flow:

$$\partial_t n^0 = \Delta n^0 + n^0 |\nabla n^0|^2 \quad (1.22)$$

with initial data

$$n^0(0, x) = n_0(x) := (\cos \theta_0, \sin \theta_0, 0) \quad (1.23)$$

while  $n^\epsilon$  is a solution of the wave-map flow:

$$\epsilon (\partial_{tt} n^\epsilon + n^\epsilon |\partial_t n^\epsilon|^2) + \partial_t n^\epsilon = \Delta n^\epsilon + n^\epsilon |\nabla n^\epsilon|^2 \quad (1.24)$$

with initial datas:

$$n^\epsilon(0, x) = n_0(x) := (\cos \theta_0(x), \sin \theta_0(x), 0), \quad (1.25)$$

$$\partial_t n^\epsilon(0, x) = n_1(x) = (-\sin \theta_0(x), \cos \theta_0(x), 0) \theta_1(x) \quad (1.26)$$

Taking  $\theta_1 = \Delta \theta_0$  the claimed optimality of the rate of converge is shown in [2].

A rigorous justification of the formal expansion (1.14) in the context of classical solutions is provided in this paper. For the original wave map (1.8) with a damping the energy bounds of  $\mathbf{d}^\epsilon$  uniform in small  $\epsilon > 0$  do not seem available. By taking the expansion (1.14) of the solutions  $\mathbf{d}^\epsilon$  to the system (1.8) with the initial conditions (1.9), one can yield a remainder system (1.15)-(1.16) of  $\mathbf{d}_R^\epsilon$ . Although the remainder system (1.15) with the initial data conditions (1.16) is still nonlinear and singular (with singular terms of the type  $\frac{1}{\epsilon} \cdot$ ), it has weaker nonlinearities than the original system (1.8). More precisely, after using the expansion (1.14), the nonlinear term  $(-\epsilon |\partial_t \mathbf{d}^\epsilon|^2 + |\nabla \mathbf{d}^\epsilon|^2) \mathbf{d}^\epsilon$  is replaced by either linear terms (the unknown  $\mathbf{d}^\epsilon$  is superseded by the known  $\mathbf{d}_0$ ) or a nonlinear term with the same form but with some higher order powers of  $\epsilon$  in front. So, by standard energy estimates, we can get uniform bounds in small  $\epsilon > 0$  of the remainder system (1.15)-(1.16).

The organization of this paper is as follows: in next section, we give the formal analysis for the asymptotic behavior of the wave map (1.8) with a damping and initial conditions (1.9) as the inertia density  $\epsilon \rightarrow 0$  by constructing the initial layer  $\mathbf{d}_0^I(\frac{t}{\epsilon}, x)$  to deal with the compatibility of the original initial conditions (1.9) and the initial condition of the limit system (1.11). In Section 3, we estimate the uniform energy bounds on small  $\epsilon > 0$  of the remainder

system (1.15)-(1.16). Finally, based on the uniform energy estimates in the previous section, Theorem 1.1 of this paper is proved in Section 4.

## 2. FORMAL ANALYSIS

In this section we present the formal analysis of the limit  $\varepsilon \rightarrow 0$  for the damped wave map (1.8) with the initial conditions (1.9). Out of the equation (1.8) we note that the formal limit, obtaining by setting  $\varepsilon = 0$ , is the heat flow system (1.11) for functions taking values into  $\mathbb{S}^2$ . We can then naturally take the ansatz

$$\mathbf{d}^\varepsilon(t, x) = \mathbf{d}_0(t, x) + \hat{\mathbf{d}}_R^\varepsilon(t, x), \quad (2.1)$$

where  $\mathbf{d}_0(t, x)$  is a solution of the heat flow system (1.11) and  $\hat{\mathbf{d}}_R^\varepsilon(t, x)$  satisfies a hyperbolic system, formally similar to (1.8) but without the geometric constraint of taking values into  $\mathbb{S}^2$ .

If the ansatz (2.1) were reasonable then  $\hat{\mathbf{d}}_R^\varepsilon(t, x) = O(\varepsilon^\alpha)$  in some sense for some  $\alpha > 0$  as  $\varepsilon > 0$  is small enough. However, by the second initial condition in (1.9) and the heat flow system (1.11), we know that

$$\begin{aligned} \partial_t \hat{\mathbf{d}}_R^\varepsilon(0, x) &= \partial_t \mathbf{d}^\varepsilon(0, x) - \partial_t \mathbf{d}_0(0, x) \\ &= \tilde{\mathbf{d}}^{in}(x) - \partial_t \mathbf{d}_0(0, x), \end{aligned}$$

which will not go to 0 as  $\varepsilon \rightarrow 0$  for arbitrarily given vectors  $\tilde{\mathbf{d}}^{in}(x)$  and  $\mathbf{d}^{in}(x)$ . As a consequence,  $\hat{\mathbf{d}}_R^\varepsilon(t, x) \neq O(\varepsilon^\alpha)$  uniformly in time for any  $\alpha > 0$ , and then the ansatz (2.1) is not satisfactory.

Therefore, in order to compensate the effect of the initial data, we need to introduce a correction term  $\mathbf{d}_0^I(\frac{t}{\varepsilon^\beta}, x)$  for some  $\beta > 0$  to be determined, called *initial layer*, such that the second initial condition in (1.9) is satisfied by  $\mathbf{d}_0(t, x) + \mathbf{d}_0^I(\frac{t}{\varepsilon^\beta}, x)$  without disturbing too much the first initial condition in (1.9), namely

$$\mathbf{d}_0^I(\frac{0}{\varepsilon^\beta}, x) = O(\varepsilon^\alpha)$$

for some  $\alpha > 0$  as  $\varepsilon \rightarrow 0$ . Thus we take the alternative ansatz

$$\mathbf{d}^\varepsilon(t, x) = \mathbf{d}_0(t, x) + \mathbf{d}_0^I(\frac{t}{\varepsilon^\beta}, x) + \sqrt{\varepsilon} \mathbf{d}_R^\varepsilon(t, x), \quad (2.2)$$

where the power  $\sqrt{\varepsilon}$  in front of the remainder term is motivated by the scaling we chose. This measures the rate of convergence, as it will be shown in the proof of Theorem 1.1.

Recalling that  $\mathbf{d}_0$  is a solution of the heat flow (1.11), we plug (2.2) into the system (1.8) and obtain:

$$\begin{aligned} &\varepsilon \partial_{tt}(\mathbf{d}_0 + \sqrt{\varepsilon} \mathbf{d}_R^\varepsilon) + \sqrt{\varepsilon} \partial_t \mathbf{d}_R^\varepsilon + (\varepsilon \partial_{tt} \mathbf{d}_0^I + \partial_t \mathbf{d}_0^I) - \Delta(\mathbf{d}_0^I + \sqrt{\varepsilon} \mathbf{d}_R^\varepsilon) \\ &= -\varepsilon |\partial_t(\mathbf{d}_0 + \mathbf{d}_0^I + \sqrt{\varepsilon} \mathbf{d}_R^\varepsilon)|^2 (\mathbf{d}_0 + \mathbf{d}_0^I + \sqrt{\varepsilon} \mathbf{d}_R^\varepsilon) \\ &\quad + |\nabla(\mathbf{d}_0 + \mathbf{d}_0^I + \sqrt{\varepsilon} \mathbf{d}_R^\varepsilon)|^2 (\mathbf{d}_0^I + \sqrt{\varepsilon} \mathbf{d}_R^\varepsilon) + [2\nabla \mathbf{d}_0 \cdot \nabla(\mathbf{d}_0^I + \sqrt{\varepsilon} \mathbf{d}_R^\varepsilon) + |\nabla(\mathbf{d}_0^I + \sqrt{\varepsilon} \mathbf{d}_R^\varepsilon)|^2] \mathbf{d}_0. \end{aligned}$$

Then we construct the initial layer in order to cancel certain time-dependent terms in the previous equations and to accommodate the discrepancy in the initial data, namely we take  $\mathbf{d}_0^I$  satisfying the  $x$ -dependent ODE and the initial-data condition:

$$\begin{cases} \varepsilon \partial_{tt} \mathbf{d}_0^I(\frac{t}{\varepsilon^\beta}, x) + \partial_t \mathbf{d}_0^I(\frac{t}{\varepsilon^\beta}, x) = 0, \\ \partial_t \mathbf{d}_0^I(\frac{0}{\varepsilon^\beta}, x) = \tilde{\mathbf{d}}^{in}(x) - \partial_t \mathbf{d}_0(0, x). \end{cases}$$

Furthermore, since  $\mathbf{d}_0^I$  is an initial layer, the following condition at infinity is required:

$$\mathbf{d}_0^I(\infty, x) = \lim_{y \rightarrow \infty} \mathbf{d}_0^I(y, x) = 0. \quad (2.3)$$

By solving the ODE system with the given boundary conditions we have

$$\mathbf{d}_0^I\left(\frac{t}{\varepsilon^\beta}, x\right) = -\varepsilon(\tilde{\mathbf{d}}^{in}(x) - \partial_t \mathbf{d}_0(0, x)) \exp\left(-\frac{t}{\varepsilon}\right). \quad (2.4)$$

We remark that the initial layer  $\mathbf{d}_0^I\left(\frac{t}{\varepsilon^\beta}, x\right)$  in (2.4) is, in fact, independent of  $\beta > 0$  and  $\mathbf{d}_0^I\left(\frac{0}{\varepsilon^\beta}, x\right) = -\varepsilon(\tilde{\mathbf{d}}^{in}(x) - \partial_t \mathbf{d}_0(0, x)) \rightarrow 0$  for any given  $\tilde{\mathbf{d}}^{in}(x)$  and  $\mathbf{d}^{in}(x)$  as  $\varepsilon \rightarrow 0$ . Consequently, the ansatz (2.2) is reasonable.

Without loss of generality, we take  $\beta = 1$  in the ansatz (2.2). Thus, by substituting (1.11) and (1.14) into the system (1.8), we derive the equation satisfied by the remainder  $\mathbf{d}_R^\varepsilon(t, x)$  as follows:

$$\begin{aligned} & \varepsilon^{\frac{3}{2}} \partial_{tt} \mathbf{d}_R^\varepsilon + \varepsilon \partial_{tt} \mathbf{d}_0 + \sqrt{\varepsilon} \partial_t \mathbf{d}_R^\varepsilon - \sqrt{\varepsilon} \Delta \mathbf{d}_R^\varepsilon + \varepsilon \Delta \mathbf{D}(x) \exp\left(-\frac{t}{\varepsilon}\right) \\ &= -\varepsilon \left| \partial_t (\mathbf{d}_0 - \varepsilon \mathbf{D}(x) \exp\left(-\frac{t}{\varepsilon}\right) + \sqrt{\varepsilon} \mathbf{d}_R^\varepsilon) \right|^2 (\mathbf{d}_0 - \varepsilon \mathbf{D}(x) \exp\left(-\frac{t}{\varepsilon}\right) + \sqrt{\varepsilon} \mathbf{d}_R^\varepsilon) \\ & \quad + \left| \nabla (\mathbf{d}_0 - \varepsilon \mathbf{D}(x) \exp\left(-\frac{t}{\varepsilon}\right) + \sqrt{\varepsilon} \mathbf{d}_R^\varepsilon) \right|^2 (-\varepsilon \mathbf{D}(x) \exp\left(-\frac{t}{\varepsilon}\right) + \sqrt{\varepsilon} \mathbf{d}_R^\varepsilon) \\ & \quad + \left[ 2 \nabla \mathbf{d}_0 \cdot \nabla (-\varepsilon \mathbf{D}(x) \exp\left(-\frac{t}{\varepsilon}\right) + \sqrt{\varepsilon} \mathbf{d}_R^\varepsilon) + \left| \nabla (-\varepsilon \mathbf{D}(x) \exp\left(-\frac{t}{\varepsilon}\right) + \sqrt{\varepsilon} \mathbf{d}_R^\varepsilon) \right|^2 \right] \mathbf{d}_0, \end{aligned}$$

which, after multiplication by  $\varepsilon^{-\frac{3}{2}}$  is the equation (1.15) we used before.

### 3. UNIFORM ENERGY ESTIMATES

In this section, we will provide, by energy methods, bounds that are uniform with respect to small inertia constant  $\varepsilon > 0$ , for the remainder system (1.15)-(1.16). By Proposition 1.1, the  $\mathbf{d}_0$ , which obeys the heat flow (1.11) into the unit sphere  $\mathbb{S}^2$  is regarded as a known quantity in the remainder system (1.15)-(1.16).

To conveniently state our results, we need to introduce the following energy functionals:

$$\begin{aligned} E_\varepsilon(t) &= |\partial_t \mathbf{d}_R^\varepsilon|_{H^2}^2 + \left(\frac{1}{\varepsilon} - 1\right) |\mathbf{d}_R^\varepsilon|_{H^2}^2 + \frac{2}{\varepsilon} |\nabla \mathbf{d}_R^\varepsilon|_{H^2}^2 + |\partial_t \mathbf{d}_R^\varepsilon + \mathbf{d}_R^\varepsilon|_{H^2}^2, \\ F_\varepsilon(t) &= \left(\frac{1}{\varepsilon} - \frac{1}{2}\right) |\partial_t \mathbf{d}_R^\varepsilon|_{H^2}^2 + \frac{1}{2\varepsilon} |\nabla \mathbf{d}_R^\varepsilon|_{H^2}^2. \end{aligned}$$

The following lemma provides the claimed uniform energy estimates :

**Lemma 3.1.** *Let  $\mathbf{d}^{in} \in \dot{H}^7(\mathbb{R}^3; \mathbb{S}^2)$  and  $[0, T]$  be the interval of existence of the the solution of the heat flow with initial data  $\mathbf{d}^{in}$ , as provided in Proposition 1.1.*

*For  $\tilde{\mathbf{d}}^{in} \in H^5$  assume that there exists a  $\varepsilon_0 \equiv \varepsilon_0(|\nabla \mathbf{d}^{in}|_{H^6}, |\tilde{\mathbf{d}}^{in}|_{H^5}, T) \in (0, \frac{1}{2})$  such that for all  $\varepsilon \in (0, \varepsilon_0)$  we that  $\mathbf{d}_R^\varepsilon \in L^\infty([0, T]; H^3)$  and  $\partial_t \mathbf{d}^\varepsilon \in L^\infty([0, T]; H^2)$  is a solution to the remainder system (1.15)-(1.16). Then there exists a positive constant  $C = C(|\nabla \mathbf{d}^{in}|_{H^6}, |\tilde{\mathbf{d}}^{in}|_{H^5}, T) > 0$  such that the inequality*

$$\frac{d}{dt} E_\varepsilon(t) + 3F_\varepsilon(t) \leq C[1 + E_\varepsilon(t)][1 + \varepsilon E_\varepsilon(t)] \quad (3.1)$$

holds for all  $\varepsilon \in (0, \frac{1}{2})$  and  $t \in [0, T]$ .

We remark that the condition  $0 < \varepsilon < \frac{1}{2}$  guarantees the relation

$$\frac{1}{2\varepsilon} < \frac{1}{\varepsilon} - 1 < \frac{1}{\varepsilon},$$

which makes the energy functionals  $E_\varepsilon(t)$  and  $F_\varepsilon(t)$  non-negative. Since our goal is to rigorously analyze the asymptotic behavior as  $\varepsilon \rightarrow 0$  for the wave map (1.8)-(1.9), the condition  $0 < \varepsilon < \frac{1}{2}$  is sufficient.

*Proof.* For the convenience of notations, we rewrite the singular terms of the remainder system (1.15) as

$$\begin{aligned}
\mathcal{S}(\mathbf{d}_R^\varepsilon) &= -\frac{1}{\sqrt{\varepsilon}} \left( \partial_{tt} \mathbf{d}_0 + \Delta \mathbf{D}(x) \exp(-\frac{t}{\varepsilon}) + |\partial_t \mathbf{d}_0 + \mathbf{D}(x) \exp(-\frac{t}{\varepsilon})|^2 \mathbf{d}_0 \right. \\
&\quad \left. + |\nabla \mathbf{d}_0|^2 \mathbf{D}(x) \exp(-\frac{t}{\varepsilon}) + 2 \nabla \mathbf{d}_0 \cdot \nabla \mathbf{D}(x) \exp(-\frac{t}{\varepsilon}) \mathbf{d}_0 \right) \Bigg\} \mathcal{S}_1 \\
&\quad + \frac{1}{\sqrt{\varepsilon}} (2(\nabla \mathbf{d}_0 \cdot \nabla \mathbf{d}_R^\varepsilon) \mathbf{d}_R^\varepsilon + |\nabla \mathbf{d}_R^\varepsilon|^2 \mathbf{d}_0) \Bigg\} \mathcal{S}_2 \\
&\quad + \frac{1}{\varepsilon} [|\nabla \mathbf{d}_0|^2 \mathbf{d}_R^\varepsilon + 2(\nabla \mathbf{d}_0 \cdot \nabla \mathbf{d}_R^\varepsilon) \mathbf{d}_0] \Bigg\} \mathcal{S}_3 \\
&\triangleq \mathcal{S}_1 + \mathcal{S}_2 + \mathcal{S}_3
\end{aligned} \tag{3.2}$$

and the regular terms as

$$\begin{aligned}
\mathcal{R}(\mathbf{d}_R^\varepsilon) &= \sqrt{\varepsilon} |\partial_t \mathbf{d}_0 + \mathbf{D}(x) \exp(-\frac{t}{\varepsilon})|^2 \mathbf{D}(x) \exp(-\frac{t}{\varepsilon}) + \sqrt{\varepsilon} |\nabla \mathbf{D}(x)|^2 \exp(-\frac{2t}{\varepsilon}) \mathbf{d}_0 \\
&\quad - \varepsilon^{\frac{3}{2}} |\nabla \mathbf{D}(x)|^2 \exp(-\frac{3t}{\varepsilon}) \mathbf{D}(x) + 2\sqrt{\varepsilon} (\nabla \mathbf{d}_0 \cdot \nabla \mathbf{D}(x)) \exp(-\frac{2t}{\varepsilon}) \mathbf{D}(x) \Bigg\} \mathcal{R}_1 \\
&\quad - 2(\partial_t \mathbf{d}_0 + \mathbf{D}(x) \exp(-\frac{t}{\varepsilon})) \cdot \partial_t \mathbf{d}_R^\varepsilon \mathbf{d}_0 - |\partial_t \mathbf{d}_0 + \mathbf{D}(x) \exp(-\frac{t}{\varepsilon})|^2 \mathbf{d}_R^\varepsilon \\
&\quad + 2\varepsilon (\partial_t \mathbf{d}_0 + \mathbf{D}(x) \exp(-\frac{t}{\varepsilon})) \cdot \partial_t \mathbf{d}_R^\varepsilon \mathbf{D}(x) \exp(-\frac{t}{\varepsilon}) + \varepsilon |\nabla \mathbf{D}|^2 \exp(-\frac{2t}{\varepsilon}) \mathbf{d}_R^\varepsilon \\
&\quad - 2(\nabla \mathbf{d}_0 \cdot \nabla \mathbf{D}(x)) \exp(-\frac{t}{\varepsilon}) \mathbf{d}_R^\varepsilon - 2(\nabla \mathbf{d}_0 \cdot \nabla \mathbf{d}_R^\varepsilon) \mathbf{D}(x) \exp(-\frac{t}{\varepsilon}) \\
&\quad - 2(\nabla \mathbf{D}(x) \cdot \nabla \mathbf{d}_R^\varepsilon) \exp(-\frac{t}{\varepsilon}) \mathbf{d}_0 + 2\varepsilon (\nabla \mathbf{D}(x) \cdot \nabla \mathbf{d}_R^\varepsilon) \exp(-\frac{2t}{\varepsilon}) \mathbf{D}(x) \Bigg\} \mathcal{R}_2 \\
&\quad + \varepsilon^{\frac{3}{2}} \mathbf{D}(x) \exp(-\frac{t}{\varepsilon}) |\partial_t \mathbf{d}_R^\varepsilon|^2 - 2\sqrt{\varepsilon} (\partial_t \mathbf{d}_0 + \mathbf{D}(x) \exp(-\frac{t}{\varepsilon})) \cdot \partial_t \mathbf{d}_R^\varepsilon \mathbf{d}_R^\varepsilon \\
&\quad - \sqrt{\varepsilon} |\partial_t \mathbf{d}_R^\varepsilon|^2 \mathbf{d}_0 - 2\sqrt{\varepsilon} (\nabla \mathbf{D}(x) \cdot \nabla \mathbf{d}_R^\varepsilon) \exp(-\frac{t}{\varepsilon}) \mathbf{d}_R^\varepsilon - \sqrt{\varepsilon} |\nabla \mathbf{d}_R^\varepsilon|^2 \mathbf{D}(x) \exp(-\frac{t}{\varepsilon}) \Bigg\} \mathcal{R}_3 \\
&\quad - \varepsilon |\partial_t \mathbf{d}_R^\varepsilon|^2 \mathbf{d}_R^\varepsilon + |\nabla \mathbf{d}_R^\varepsilon|^2 \mathbf{d}_R^\varepsilon \Bigg\} \mathcal{R}_4 \\
&\triangleq \mathcal{R}_1 + \mathcal{R}_2 + \mathcal{R}_3 + \mathcal{R}_4.
\end{aligned} \tag{3.3}$$

*Step 1.  $L^2$ -estimates.* Multiplying the remainder equation (1.15) by  $\partial_t \mathbf{d}_R^\varepsilon$ , integrating over  $\mathbb{R}^3$  and by parts, we obtain the following equation:

$$\frac{1}{2} \frac{d}{dt} (|\partial_t \mathbf{d}_R^\varepsilon|_{L^2}^2 + \frac{1}{\varepsilon} |\nabla \mathbf{d}_R^\varepsilon|_{L^2}^2) + \frac{1}{\varepsilon} |\partial_t \mathbf{d}_R^\varepsilon|_{L^2}^2 = \langle \mathcal{S}(\mathbf{d}_R^\varepsilon), \partial_t \mathbf{d}_R^\varepsilon \rangle + \langle \mathcal{R}(\mathbf{d}_R^\varepsilon), \partial_t \mathbf{d}_R^\varepsilon \rangle. \tag{3.4}$$

(I) *Estimates for the singular terms  $\langle \mathcal{S}(\mathbf{d}_R^\varepsilon), \partial_t \mathbf{d}_R^\varepsilon \rangle$ :*

For estimating  $\langle \mathcal{S}_1, \partial_t \mathbf{d}_R^\varepsilon \rangle$ , we use the Hölder inequality, the Sobolev embedding theorems, the facts that  $\exp(-\frac{t}{\varepsilon}) \leq 1$  and  $|\mathbf{d}_0| = 1$  to obtain:

$$\frac{1}{\sqrt{\varepsilon}} \langle \partial_{tt} \mathbf{d}_0 + \Delta \mathbf{D}(x) \exp(-\frac{t}{\varepsilon}), \partial_t \mathbf{d}_R^\varepsilon \rangle \lesssim (|\partial_{tt} \mathbf{d}_0|_{L_t^\infty L_x^2} + |\Delta \mathbf{D}(x)|_{L^2}) \left| \frac{\partial_t \mathbf{d}_R^\varepsilon}{\sqrt{\varepsilon}} \right|_{L^2}$$

and

$$\begin{aligned}
&\frac{1}{\sqrt{\varepsilon}} \langle |\nabla \mathbf{d}_0|^2 \mathbf{D}(x) \exp(-\frac{t}{\varepsilon}) + 2 \nabla \mathbf{d}_0 \cdot \nabla \mathbf{D}(x) \exp(-\frac{t}{\varepsilon}) \mathbf{d}_0, \partial_t \mathbf{d}_R^\varepsilon \rangle \\
&\lesssim (|\nabla \mathbf{d}_0|_{L^\infty}^2 |\mathbf{D}(x)|_{L^2} + |\nabla \mathbf{d}_0|_{L^\infty} |\nabla \mathbf{D}(x)|_{L^2}) \left| \frac{\partial_t \mathbf{d}_R^\varepsilon}{\sqrt{\varepsilon}} \right|_{L^2} \\
&\lesssim (|\nabla \mathbf{d}_0|_{L_t^\infty H_x^2}^2 + |\nabla \mathbf{d}_0|_{L_t^\infty H_x^2}) |\mathbf{D}(x)|_{H^1} \left| \frac{\partial_t \mathbf{d}_R^\varepsilon}{\sqrt{\varepsilon}} \right|_{L^2}.
\end{aligned}$$

Similarly:

$$\begin{aligned}
\frac{1}{\sqrt{\varepsilon}} \langle |\partial_t \mathbf{d}_0 + \mathbf{D}(x) \exp(-\frac{t}{\varepsilon})|^2 \mathbf{d}_0, \partial_t \mathbf{d}_R^\varepsilon \rangle &\lesssim |\partial_t \mathbf{d}_0 + \mathbf{D}(x) \exp(-\frac{t}{\varepsilon})|_{L^4}^2 \left| \frac{\partial_t \mathbf{d}_R^\varepsilon}{\sqrt{\varepsilon}} \right|_{L^2} \\
&\lesssim (|\partial_t \mathbf{d}_0|_{L_t^\infty L_x^4}^2 + |\mathbf{D}(x)|_{L^4}^2) \left| \frac{\partial_t \mathbf{d}_R^\varepsilon}{\sqrt{\varepsilon}} \right|_{L^2} \\
&\lesssim (|\partial_t \mathbf{d}_0|_{L_t^\infty H_x^1}^2 + |\mathbf{D}(x)|_{H^1}^2) \left| \frac{\partial_t \mathbf{d}_R^\varepsilon}{\sqrt{\varepsilon}} \right|_{L^2}.
\end{aligned}$$

Summarizing, we estimate  $\langle \mathcal{S}_1, \partial_t \mathbf{d}_R^\varepsilon \rangle$  as follows:

$$\begin{aligned}
\langle \mathcal{S}_1, \partial_t \mathbf{d}_R^\varepsilon \rangle &\lesssim \left[ |\partial_{tt} \mathbf{d}_0|_{L_t^\infty L_x^2} + |\partial_t \mathbf{d}_0|_{L_t^\infty H_x^1}^2 + |\mathbf{D}(x)|_{H^1}^2 \right. \\
&\quad \left. + (|\nabla \mathbf{d}_0|_{L_t^\infty H_x^2}^2 + 1) |\mathbf{D}(x)|_{H^2} \right] \left| \frac{\partial_t \mathbf{d}_R^\varepsilon}{\sqrt{\varepsilon}} \right|_{L_x^2}.
\end{aligned} \tag{3.5}$$



It is easy to derive the estimates of  $\langle \mathcal{S}_2, \partial_t \mathbf{d}_R^\varepsilon \rangle$  and  $\langle \mathcal{S}_3, \partial_t \mathbf{d}_R^\varepsilon \rangle$  as follows:

$$\begin{aligned} \langle \mathcal{S}_2, \partial_t \mathbf{d}_R^\varepsilon \rangle &= \frac{1}{\sqrt{\varepsilon}} \langle 2(\nabla \mathbf{d}_0 \cdot \nabla \mathbf{d}_R^\varepsilon) \mathbf{d}_R^\varepsilon + |\nabla \mathbf{d}_R^\varepsilon|^2 \mathbf{d}_0, \partial_t \mathbf{d}_R^\varepsilon \rangle \\ &\lesssim (|\nabla \mathbf{d}_0|_{L^\infty} |\mathbf{d}_R^\varepsilon|_{L^\infty} |\nabla \mathbf{d}_R^\varepsilon|_{L^2} + |\nabla \mathbf{d}_R^\varepsilon|_{L^4}^2) \left| \frac{\partial_t \mathbf{d}_R^\varepsilon}{\sqrt{\varepsilon}} \right|_{L^2} \\ &\lesssim (|\nabla \mathbf{d}_0|_{L_t^\infty H_x^2} + 1) |\mathbf{d}_R^\varepsilon|_{H^2}^2 \left| \frac{\partial_t \mathbf{d}_R^\varepsilon}{\sqrt{\varepsilon}} \right|_{L^2} \end{aligned} \quad (3.6)$$

and

$$\begin{aligned} \langle \mathcal{S}_3, \partial_t \mathbf{d}_R^\varepsilon \rangle &= \frac{1}{\varepsilon} \langle |\nabla \mathbf{d}_0|^2 \mathbf{d}_R^\varepsilon + 2(\nabla \mathbf{d}_0 \cdot \nabla \mathbf{d}_R^\varepsilon) \mathbf{d}_0, \partial_t \mathbf{d}_R^\varepsilon \rangle \\ &\lesssim (|\nabla \mathbf{d}_0|_{L^\infty}^2 \left| \frac{\mathbf{d}_R^\varepsilon}{\sqrt{\varepsilon}} \right|_{L^2} + |\nabla \mathbf{d}_0|_{L^\infty} \left| \frac{\nabla \mathbf{d}_R^\varepsilon}{\sqrt{\varepsilon}} \right|_{L^2}) \left| \frac{\partial_t \mathbf{d}_R^\varepsilon}{\sqrt{\varepsilon}} \right|_{L^2} \\ &\lesssim (|\nabla \mathbf{d}_0|_{L_t^\infty H_x^2} + |\nabla \mathbf{d}_0|_{L_t^\infty H_x^2}^2) ( \left| \frac{\mathbf{d}_R^\varepsilon}{\sqrt{\varepsilon}} \right|_{L^2} + \left| \frac{\nabla \mathbf{d}_R^\varepsilon}{\sqrt{\varepsilon}} \right|_{L^2} ) \left| \frac{\partial_t \mathbf{d}_R^\varepsilon}{\sqrt{\varepsilon}} \right|_{L^2}. \end{aligned} \quad (3.7)$$

Hence we have the estimate of  $\langle \mathcal{S}(\mathbf{d}_R^\varepsilon), \partial_t \mathbf{d}_R^\varepsilon \rangle$  by combining the inequalities (3.5), (3.6) and (3.7):

$$\langle \mathcal{S}(\mathbf{d}_R^\varepsilon), \partial_t \mathbf{d}_R^\varepsilon \rangle \leq C_{11} \left( 1 + |\mathbf{d}_R^\varepsilon|_{H^2}^2 + \left| \frac{\mathbf{d}_R^\varepsilon}{\sqrt{\varepsilon}} \right|_{L^2} + \left| \frac{\nabla \mathbf{d}_R^\varepsilon}{\sqrt{\varepsilon}} \right|_{L^2} \right) \left| \frac{\partial_t \mathbf{d}_R^\varepsilon}{\sqrt{\varepsilon}} \right|_{L^2}, \quad (3.8)$$

where the constant

$$C_{11} = C \left[ |\partial_{tt} \mathbf{d}_0|_{L_t^\infty L_x^2} + |\partial_t \mathbf{d}_0|_{L_t^\infty H_x^1}^2 + |\mathbf{D}(x)|_{H^1}^2 + (|\nabla \mathbf{d}_0|_{L_t^\infty H_x^2}^2 + 1)(|\mathbf{D}(x)|_{H^2} + 1) \right] > 0$$

for some computable positive constant  $C > 0$ .

(II) *Estimates for the regular terms  $\langle \mathcal{R}(\mathbf{d}_R^\varepsilon), \partial_t \mathbf{d}_R^\varepsilon \rangle$ :*

We have divided the regular terms  $\mathcal{R}(\mathbf{d}_R^\varepsilon)$  into four parts, which we will estimate separately.

The estimate of  $\langle \mathcal{R}_1, \partial_t \mathbf{d}_R^\varepsilon \rangle$ . By the Hölder inequality and the Sobolev embedding theorems we have

$$\begin{aligned} &\sqrt{\varepsilon} \langle |\partial_t \mathbf{d}_0 + \mathbf{D}(x) \exp(-\frac{t}{\varepsilon})|^2 \mathbf{D}(x) \exp(-\frac{t}{\varepsilon}), \partial_t \mathbf{d}_R^\varepsilon \rangle \\ &\lesssim \sqrt{\varepsilon} |\partial_t \mathbf{d}_0 + \mathbf{D}(x) \exp(-\frac{t}{\varepsilon})|_{L^\infty}^2 |\mathbf{D}(x)|_{L^2} |\partial_t \mathbf{d}_R^\varepsilon|_{L^2} \\ &\lesssim \sqrt{\varepsilon} (|\partial_t \mathbf{d}_0|_{L_t^\infty H_x^2}^2 + |\mathbf{D}(x)|_{H^2}^2) |\mathbf{D}(x)|_{L^2} |\partial_t \mathbf{d}_R^\varepsilon|_{L^2} \end{aligned}$$

and

$$\begin{aligned} \sqrt{\varepsilon} \langle (\nabla \mathbf{d}_0 \nabla \mathbf{D}(x)) \exp(-\frac{2t}{\varepsilon}) \mathbf{D}(x), \partial_t \mathbf{d}_R^\varepsilon \rangle &\lesssim \sqrt{\varepsilon} |\nabla \mathbf{d}_0|_{L^\infty} |\mathbf{D}(x)|_{L^\infty} |\nabla \mathbf{D}(x)|_{L^2} |\partial_t \mathbf{d}_R^\varepsilon|_{L^2} \\ &\lesssim \sqrt{\varepsilon} |\nabla \mathbf{d}_0|_{L_t^\infty H_x^2} |\mathbf{D}(x)|_{H^2}^2 |\partial_t \mathbf{d}_R^\varepsilon|_{L^2}, \end{aligned}$$

where we have used the fact that  $\exp(-\frac{t}{\varepsilon}) \leq 1$ .

The other two terms in  $\langle \mathcal{R}_1, \partial_t \mathbf{d}_R^\varepsilon \rangle$  are similarly estimated, as follows:

$$\begin{aligned} \sqrt{\varepsilon} \langle |\nabla \mathbf{D}(x)|^2 \exp(-\frac{2t}{\varepsilon}) \mathbf{d}_0, \partial_t \mathbf{d}_R^\varepsilon \rangle &\lesssim \sqrt{\varepsilon} |\nabla \mathbf{D}(x)|_{L^4}^2 |\partial_t \mathbf{d}_R^\varepsilon|_{L^2} \\ &\lesssim \sqrt{\varepsilon} |\mathbf{D}(x)|_{H^2}^2 |\partial_t \mathbf{d}_R^\varepsilon|_{L^2}, \\ \varepsilon^{\frac{3}{2}} \langle |\nabla \mathbf{D}(x)|^2 \exp(-\frac{3t}{\varepsilon}) \mathbf{D}(x), \partial_t \mathbf{d}_R^\varepsilon \rangle &\lesssim \varepsilon^{\frac{3}{2}} |\mathbf{D}(x)|_{L^\infty} |\nabla \mathbf{D}(x)|_{L^4}^2 |\partial_t \mathbf{d}_R^\varepsilon|_{L^2} \\ &\lesssim \varepsilon^{\frac{3}{2}} |\mathbf{D}(x)|_{H^2}^3 |\partial_t \mathbf{d}_R^\varepsilon|_{L^2}. \end{aligned}$$

For any small enough  $\varepsilon$  such that  $\varepsilon \in (0, \frac{1}{2}]$ , we have  $\varepsilon^{\frac{3}{2}} \leq \varepsilon^{\frac{1}{2}}$ . So from the above inequalities, we obtain the following estimate:

$$\langle \mathcal{R}_1, \partial_t \mathbf{d}_R^\varepsilon \rangle \lesssim \sqrt{\varepsilon} \left[ (|\partial_t \mathbf{d}_0|_{L_t^\infty H_x^2}^2 + |\nabla \mathbf{d}_0|_{L_t^\infty H_x^2}^2 + |\mathbf{D}(x)|_{H^2}^2) |\mathbf{D}(x)|_{H^2} + |\mathbf{D}(x)|_{H^2}^2 \right] |\partial_t \mathbf{d}_R^\varepsilon|_{L^2}. \quad (3.9)$$

The estimate of  $\langle \mathcal{R}_2, \partial_t \mathbf{d}_R^\varepsilon \rangle$ . We have the following estimates for the first three terms in  $\langle \mathcal{R}_2, \partial_t \mathbf{d}_R^\varepsilon \rangle$ :

$$\begin{aligned} \langle |\partial_t \mathbf{d}_0 + \mathbf{D}(x) \exp(-\frac{t}{\varepsilon})|^2 \mathbf{d}_R^\varepsilon, \partial_t \mathbf{d}_R^\varepsilon \rangle &\lesssim |\partial_t \mathbf{d}_0 + \mathbf{D}(x) \exp(-\frac{t}{\varepsilon})|_{L^\infty}^2 |\mathbf{d}_R^\varepsilon|_{L^2} |\partial_t \mathbf{d}_R^\varepsilon|_{L^2} \\ &\lesssim (|\partial_t \mathbf{d}_0|_{L_t^\infty H_x^2}^2 + |\mathbf{D}(x)|_{H^2}^2) |\mathbf{d}_R^\varepsilon|_{L^2} |\partial_t \mathbf{d}_R^\varepsilon|_{L^2}, \\ \langle (\partial_t \mathbf{d}_0 + \mathbf{D}(x) \exp(-\frac{t}{\varepsilon})) \cdot \partial_t \mathbf{d}_R^\varepsilon \mathbf{d}_0, \partial_t \mathbf{d}_R^\varepsilon \rangle &\lesssim (|\partial_t \mathbf{d}_0|_{L^\infty} + |\mathbf{D}(x)|_{L^\infty}) |\partial_t \mathbf{d}_R^\varepsilon|_{L^2}^2 \\ &\lesssim (|\partial_t \mathbf{d}_0|_{L_t^\infty H_x^2} + |\mathbf{D}(x)|_{H^2}) |\partial_t \mathbf{d}_R^\varepsilon|_{L^2}^2 \end{aligned}$$

and

$$\begin{aligned} &\varepsilon \langle (\partial_t \mathbf{d}_0 + \mathbf{D}(x) \exp(-\frac{t}{\varepsilon})) \cdot \mathbf{D}(x) \exp(-\frac{t}{\varepsilon}) \partial_t \mathbf{d}_R^\varepsilon, \partial_t \mathbf{d}_R^\varepsilon \rangle \\ &\lesssim \varepsilon |\partial_t \mathbf{d}_0 + \mathbf{D}(x) \exp(-\frac{t}{\varepsilon})|_{L^\infty} |\mathbf{D}(x)|_{L^\infty} |\partial_t \mathbf{d}_R^\varepsilon|_{L^2}^2 \\ &\lesssim \varepsilon (|\partial_t \mathbf{d}_0|_{L_t^\infty H_x^2} + |\mathbf{D}(x)|_{H^2}) |\mathbf{D}(x)|_{H^2} |\partial_t \mathbf{d}_R^\varepsilon|_{L^2}^2, \end{aligned}$$

where we have used the Hölder inequality, the Sobolev embedding theorems and the fact that  $|\mathbf{d}_0| = 1$ .

As for the following three terms, one can easily obtain:

$$\begin{aligned} \varepsilon \langle |\nabla \mathbf{D}(x)|^2 \exp(-\frac{2t}{\varepsilon}) \mathbf{d}_R^\varepsilon, \partial_t \mathbf{d}_R^\varepsilon \rangle &\lesssim \varepsilon |\nabla \mathbf{D}(x)|_{L^\infty}^2 |\mathbf{d}_R^\varepsilon|_{L^2} |\partial_t \mathbf{d}_R^\varepsilon|_{L^2} \\ &\lesssim \varepsilon |\mathbf{D}(x)|_{H^3}^2 |\mathbf{d}_R^\varepsilon|_{L^2} |\partial_t \mathbf{d}_R^\varepsilon|_{L^2}, \\ \langle (\nabla \mathbf{d}_0 \cdot \nabla \mathbf{D}(x)) \exp(-\frac{t}{\varepsilon}) \mathbf{d}_R^\varepsilon, \partial_t \mathbf{d}_R^\varepsilon \rangle &\lesssim |\nabla \mathbf{d}_0|_{L^\infty} |\nabla \mathbf{D}(x)|_{L^\infty} |\mathbf{d}_R^\varepsilon|_{L^2} |\partial_t \mathbf{d}_R^\varepsilon|_{L^2} \\ &\lesssim |\nabla \mathbf{d}_0|_{L_t^\infty H_x^2} |\mathbf{D}(x)|_{H^3} |\mathbf{d}_R^\varepsilon|_{L^2} |\partial_t \mathbf{d}_R^\varepsilon|_{L^2}, \\ \langle (\nabla \mathbf{d}_0 \cdot \nabla \mathbf{d}_R^\varepsilon) \mathbf{D}(x) \exp(-\frac{t}{\varepsilon}), \partial_t \mathbf{d}_R^\varepsilon \rangle &\lesssim |\nabla \mathbf{d}_0|_{L^\infty} |\mathbf{D}(x)|_{L^\infty} |\nabla \mathbf{d}_R^\varepsilon|_{L^2} |\partial_t \mathbf{d}_R^\varepsilon|_{L^2} \\ &\lesssim |\nabla \mathbf{d}_0|_{L_t^\infty H_x^2} |\mathbf{D}(x)|_{H^2} |\nabla \mathbf{d}_R^\varepsilon|_{L^2} |\partial_t \mathbf{d}_R^\varepsilon|_{L^2}, \end{aligned}$$

where we have used the Hölder inequality, the Sobolev embedding theorems and the bound  $\exp(-\frac{t}{\varepsilon}) \leq 1$ .

Similarly as before, we estimate the last two terms, as follows:

$$\begin{aligned} \langle (\nabla \mathbf{D}(x) \cdot \nabla \mathbf{d}_R^\varepsilon) \exp(-\frac{t}{\varepsilon}) \mathbf{d}_0, \partial_t \mathbf{d}_R^\varepsilon \rangle &\lesssim |\nabla \mathbf{D}(x)|_{L^\infty} |\nabla \mathbf{d}_R^\varepsilon|_{L^2} |\partial_t \mathbf{d}_R^\varepsilon|_{L^2} \\ &\lesssim |\mathbf{D}(x)|_{H^3} |\nabla \mathbf{d}_R^\varepsilon|_{L^2} |\partial_t \mathbf{d}_R^\varepsilon|_{L^2}, \\ \varepsilon \langle (\nabla \mathbf{D}(x) \cdot \nabla \mathbf{d}_R^\varepsilon) \exp(-\frac{2t}{\varepsilon}) \mathbf{D}(x), \partial_t \mathbf{d}_R^\varepsilon \rangle &\lesssim \varepsilon |\nabla \mathbf{D}(x)|_{L^\infty} |\mathbf{D}(x)|_{L^\infty} |\nabla \mathbf{d}_R^\varepsilon|_{L^2} |\partial_t \mathbf{d}_R^\varepsilon|_{L^2} \\ &\lesssim \varepsilon |\mathbf{D}(x)|_{H^3}^2 |\nabla \mathbf{d}_R^\varepsilon|_{L^2} |\partial_t \mathbf{d}_R^\varepsilon|_{L^2}. \end{aligned}$$

Combining the above estimates and using that  $\varepsilon \in (0, \frac{1}{2}]$ , we get

$$\begin{aligned} \langle \mathcal{R}_2, \partial_t \mathbf{d}_R^\varepsilon \rangle &\lesssim (1 + |\partial_t \mathbf{d}_0|_{L_t^\infty H_x^2} + |\nabla \mathbf{d}_0|_{L_t^\infty H_x^2} + |\mathbf{D}(x)|_{H^3}) (|\partial_t \mathbf{d}_0|_{L_t^\infty H_x^2} + |\mathbf{D}(x)|_{H^3}) \\ &\quad \times (|\mathbf{d}_R^\varepsilon|_{L^2} + |\nabla \mathbf{d}_R^\varepsilon|_{L^2} + |\partial_t \mathbf{d}_R^\varepsilon|_{L^2}) |\partial_t \mathbf{d}_R^\varepsilon|_{L^2}. \end{aligned} \quad (3.10)$$

The estimate of  $\langle \mathcal{R}_3, \partial_t \mathbf{d}_R^\varepsilon \rangle$ . One can easily derive the following estimates

$$\begin{aligned} \sqrt{\varepsilon} \langle (\partial_t \mathbf{d}_0 + \mathbf{D}(x) \exp(-\frac{t}{\varepsilon})) \cdot \partial_t \mathbf{d}_R^\varepsilon \mathbf{d}_R^\varepsilon, \partial_t \mathbf{d}_R^\varepsilon \rangle &\lesssim \sqrt{\varepsilon} (|\partial_t \mathbf{d}_0|_{L_t^\infty H_x^2} + |\mathbf{D}(x)|_{H^2}) |\mathbf{d}_R^\varepsilon|_{H^2} |\partial_t \mathbf{d}_R^\varepsilon|_{L^2}^2, \\ \sqrt{\varepsilon} \langle (\nabla \mathbf{D}(x) \cdot \nabla \mathbf{d}_R^\varepsilon) \exp(-\frac{t}{\varepsilon}) \mathbf{d}_R^\varepsilon, \partial_t \mathbf{d}_R^\varepsilon \rangle &\lesssim \sqrt{\varepsilon} |\mathbf{D}(x)|_{H^3} |\mathbf{d}_R^\varepsilon|_{H^2} |\nabla \mathbf{d}_R^\varepsilon|_{L^2} |\partial_t \mathbf{d}_R^\varepsilon|_{L^2}, \end{aligned}$$

by using the Hölder inequality and the Sobolev embedding theorems. Recalling that  $|\mathbf{d}_0| = 1$  and using the Sobolev embeddings  $H^1 \hookrightarrow L^4$  and  $H^2 \hookrightarrow L^\infty$  we get

$$\begin{aligned} \varepsilon^{\frac{3}{2}} \langle \mathbf{D}(x) \exp(-\frac{t}{\varepsilon}) |\partial_t \mathbf{d}_R^\varepsilon|^2, \partial_t \mathbf{d}_R^\varepsilon \rangle &\lesssim \varepsilon^{\frac{3}{2}} |\mathbf{D}(x)|_{L^\infty} |\partial_t \mathbf{d}_R^\varepsilon|_{L^4}^2 |\partial_t \mathbf{d}_R^\varepsilon|_{L^2} \\ &\lesssim \varepsilon^{\frac{3}{2}} |\mathbf{D}(x)|_{H^2} |\partial_t \mathbf{d}_R^\varepsilon|_{H^1}^2 |\partial_t \mathbf{d}_R^\varepsilon|_{L^2}, \\ \sqrt{\varepsilon} \langle |\partial_t \mathbf{d}_R^\varepsilon|^2 \mathbf{d}_0, \partial_t \mathbf{d}_R^\varepsilon \rangle &\lesssim \sqrt{\varepsilon} |\partial_t \mathbf{d}_R^\varepsilon|_{L^4}^2 |\partial_t \mathbf{d}_R^\varepsilon|_{L^2} \\ &\lesssim \sqrt{\varepsilon} |\partial_t \mathbf{d}_R^\varepsilon|_{H^1}^2 |\partial_t \mathbf{d}_R^\varepsilon|_{L^2}, \\ \sqrt{\varepsilon} \langle |\nabla \mathbf{d}_R^\varepsilon|^2 \mathbf{D}(x) \exp(-\frac{t}{\varepsilon}), \partial_t \mathbf{d}_R^\varepsilon \rangle &\lesssim \sqrt{\varepsilon} |\mathbf{D}(x)|_{L^\infty} |\nabla \mathbf{d}_R^\varepsilon|_{L^4}^2 |\partial_t \mathbf{d}_R^\varepsilon|_{L^2} \\ &\lesssim \sqrt{\varepsilon} |\mathbf{D}(x)|_{H^2} |\nabla \mathbf{d}_R^\varepsilon|_{H^1}^2 |\partial_t \mathbf{d}_R^\varepsilon|_{L^2}, \end{aligned}$$

The above estimates immediately give the bound on  $\langle \mathcal{R}_3, \partial_t \mathbf{d}_R^\varepsilon \rangle$ :

$$\begin{aligned} \langle \mathcal{R}_3, \partial_t \mathbf{d}_R^\varepsilon \rangle &\lesssim \sqrt{\varepsilon} (1 + |\partial_t \mathbf{d}_0|_{L_t^\infty H_x^2} + |\mathbf{D}(x)|_{H^3}) (|\partial_t \mathbf{d}_R^\varepsilon|_{H^1}^2 + |\nabla \mathbf{d}_R^\varepsilon|_{H^1}^2 \\ &\quad + |\mathbf{d}_R^\varepsilon|_{H^2} |\nabla \mathbf{d}_R^\varepsilon|_{L^2} + |\mathbf{d}_R^\varepsilon|_{H^2} |\partial_t \mathbf{d}_R^\varepsilon|_{L^2}) |\partial_t \mathbf{d}_R^\varepsilon|_{L^2} \\ &\lesssim \sqrt{\varepsilon} (1 + |\partial_t \mathbf{d}_0|_{L_t^\infty H_x^2} + |\mathbf{D}(x)|_{H^3}) (|\partial_t \mathbf{d}_R^\varepsilon|_{H^1}^2 + |\mathbf{d}_R^\varepsilon|_{H^2}^2) |\partial_t \mathbf{d}_R^\varepsilon|_{L^2}. \end{aligned} \quad (3.11)$$

The estimate of  $\langle \mathcal{R}_4, \partial_t \mathbf{d}_R^\varepsilon \rangle$ . The first term in  $\langle \mathcal{R}_4, \partial_t \mathbf{d}_R^\varepsilon \rangle$  can be bounded as

$$\begin{aligned} \varepsilon \langle |\partial_t \mathbf{d}_R^\varepsilon|^2 \mathbf{d}_R^\varepsilon, \partial_t \mathbf{d}_R^\varepsilon \rangle &\leq \varepsilon |\mathbf{d}_R^\varepsilon|_{L^\infty} |\partial_t \mathbf{d}_R^\varepsilon|_{L^4}^2 |\partial_t \mathbf{d}_R^\varepsilon|_{L^2} \\ &\lesssim \varepsilon |\mathbf{d}_R^\varepsilon|_{H^2} |\partial_t \mathbf{d}_R^\varepsilon|_{H^1}^2 |\partial_t \mathbf{d}_R^\varepsilon|_{L^2} \end{aligned}$$

by using the Hölder inequality and the Sobolev embedding theorems. The other term can be bounded in a similar way:

$$\begin{aligned} \langle |\nabla \mathbf{d}_R^\varepsilon|^2 \mathbf{d}_R^\varepsilon, \partial_t \mathbf{d}_R^\varepsilon \rangle &\lesssim |\mathbf{d}_R^\varepsilon|_{L^\infty} |\nabla \mathbf{d}_R^\varepsilon|_{L^4}^2 |\partial_t \mathbf{d}_R^\varepsilon|_{L^2} \\ &\lesssim |\mathbf{d}_R^\varepsilon|_{H^2} |\nabla \mathbf{d}_R^\varepsilon|_{H^1}^2 |\partial_t \mathbf{d}_R^\varepsilon|_{L^2}. \end{aligned}$$

Hence we obtain the estimate of  $\langle \mathcal{R}_4, \partial_t \mathbf{d}_R^\varepsilon \rangle$  as follows:

$$\langle \mathcal{R}_4, \partial_t \mathbf{d}_R^\varepsilon \rangle \lesssim (\varepsilon |\partial_t \mathbf{d}_R^\varepsilon|_{H^1}^2 + |\nabla \mathbf{d}_R^\varepsilon|_{H^1}^2) |\mathbf{d}_R^\varepsilon|_{H^2} |\partial_t \mathbf{d}_R^\varepsilon|_{L^2}. \quad (3.12)$$

Summing up the inequalities (3.9), (3.10), (3.11) and (3.12), we get

$$\begin{aligned} \langle \mathcal{R}(\mathbf{d}_R^\varepsilon), \partial_t \mathbf{d}_R^\varepsilon \rangle &\leq C_{12} \left[ \sqrt{\varepsilon} + |\mathbf{d}_R^\varepsilon|_{H^1} + |\partial_t \mathbf{d}_R^\varepsilon|_{L^2} + \sqrt{\varepsilon} |\partial_t \mathbf{d}_R^\varepsilon|_{H^1}^2 + \sqrt{\varepsilon} |\mathbf{d}_R^\varepsilon|_{H^2}^2 \right. \\ &\quad \left. + |\mathbf{d}_R^\varepsilon|_{H^2} (\varepsilon |\partial_t \mathbf{d}_R^\varepsilon|_{H^1}^2 + |\nabla \mathbf{d}_R^\varepsilon|_{H^1}^2) \right] |\partial_t \mathbf{d}_R^\varepsilon|_{L^2}, \end{aligned} \quad (3.13)$$

where

$$C_{12} = C(1 + |\partial_t \mathbf{d}_0|_{L_t^\infty H_x^2} + |\nabla \mathbf{d}_0|_{L_t^\infty H_x^2} + |\mathbf{D}(x)|_{H^3}^2) (1 + |\partial_t \mathbf{d}_0|_{L_t^\infty H_x^2} + |\mathbf{D}(x)|_{H^3}) > 0,$$

and  $C$  is a positive computable constant.

Therefore, plugging the estimates (3.8) and (3.13) into the equality (3.4), we have

$$\begin{aligned} &\frac{1}{2} \frac{d}{dt} (|\partial_t \mathbf{d}_R^\varepsilon|_{L^2}^2 + \frac{1}{\varepsilon} |\nabla \mathbf{d}_R^\varepsilon|_{L^2}^2) + \frac{1}{\varepsilon} |\partial_t \mathbf{d}_R^\varepsilon|_{L^2}^2 \\ &\leq C_1 \left\{ (1 + |\mathbf{d}_R^\varepsilon|_{H^2}^2 + |\frac{\mathbf{d}_R^\varepsilon}{\sqrt{\varepsilon}}|_{L^2} + |\frac{\nabla \mathbf{d}_R^\varepsilon}{\sqrt{\varepsilon}}|_{L^2}) |\frac{\partial_t \mathbf{d}_R^\varepsilon}{\sqrt{\varepsilon}}|_{L^2} + \left[ \sqrt{\varepsilon} + |\mathbf{d}_R^\varepsilon|_{H^1} + |\partial_t \mathbf{d}_R^\varepsilon|_{L^2} \right. \right. \\ &\quad \left. \left. + \sqrt{\varepsilon} |\partial_t \mathbf{d}_R^\varepsilon|_{H^1}^2 + \sqrt{\varepsilon} |\mathbf{d}_R^\varepsilon|_{H^2}^2 + |\mathbf{d}_R^\varepsilon|_{H^2} (\varepsilon |\partial_t \mathbf{d}_R^\varepsilon|_{H^1}^2 + |\nabla \mathbf{d}_R^\varepsilon|_{H^1}^2) \right] |\partial_t \mathbf{d}_R^\varepsilon|_{L^2} \right\}, \end{aligned} \quad (3.14)$$

where the constant

$$C_1 = C(1 + |\mathbf{D}(x)|_{H^3}^3 + |\partial_{tt} \mathbf{d}_0|_{L^\infty(0,T;L^2)} + |\partial_t \mathbf{d}_0|_{L^\infty(0,T;H^2)}^3 + |\nabla \mathbf{d}_0|_{L^\infty(0,T;H^2)}^4) > 0$$

for some computable positive constant  $C$ .

(III) *Estimates of the norm  $|\mathbf{d}_R^\varepsilon|_{L^2}$ :*

Observing that the norm  $|\mathbf{d}_R^\varepsilon|_{L^2}$  appearing on the right hand side of (3.14) is not yet controlled, we need additional work to estimate  $|\mathbf{d}_R^\varepsilon|_{L^2}$ . In order to do this it is natural to

multiply the equation of the reminder term (1.15) by  $\mathbf{d}_R^\epsilon$ , integrate over  $\mathbb{R}^3$  and by parts, and use the identity:

$$\begin{aligned} \langle \partial_t \mathbf{d}_R^\epsilon, \mathbf{d}_R^\epsilon \rangle &= \frac{d}{dt} \langle \partial_t \mathbf{d}_R^\epsilon, \mathbf{d}_R^\epsilon \rangle - |\partial_t \mathbf{d}_R^\epsilon|_{L^2}^2 \\ &= \frac{1}{2} \frac{d}{dt} (|\partial_t \mathbf{d}_R^\epsilon + \mathbf{d}_R^\epsilon|_{L^2}^2 - |\partial_t \mathbf{d}_R^\epsilon|_{L^2}^2 - |\mathbf{d}_R^\epsilon|_{L^2}^2) - |\partial_t \mathbf{d}_R^\epsilon|_{L^2}^2, \end{aligned}$$

in order to get

$$\begin{aligned} &\frac{1}{2} \frac{d}{dt} \left[ |\partial_t \mathbf{d}_R^\epsilon + \mathbf{d}_R^\epsilon|_{L^2}^2 + \left(\frac{1}{\epsilon} - 1\right) |\mathbf{d}_R^\epsilon|_{L^2}^2 - |\partial_t \mathbf{d}_R^\epsilon|_{L^2}^2 \right] - |\partial_t \mathbf{d}_R^\epsilon|_{L^2}^2 + \frac{1}{\epsilon} |\nabla \mathbf{d}_R^\epsilon|_{L^2}^2 \\ &= \langle \mathcal{S}(\mathbf{d}_R^\epsilon), \mathbf{d}_R^\epsilon \rangle + \langle \mathcal{R}(\mathbf{d}_R^\epsilon), \mathbf{d}_R^\epsilon \rangle. \end{aligned} \quad (3.15)$$

Using the estimates (3.8) and (3.13) previously derived for bounding the terms  $\langle \mathcal{S}(\mathbf{d}_R^\epsilon), \partial_t \mathbf{d}_R^\epsilon \rangle$  and  $\langle \mathcal{R}(\mathbf{d}_R^\epsilon), \partial_t \mathbf{d}_R^\epsilon \rangle$ , we can analogously estimate the terms  $\langle \mathcal{S}(\mathbf{d}_R^\epsilon), \mathbf{d}_R^\epsilon \rangle$  and  $\langle \mathcal{R}(\mathbf{d}_R^\epsilon), \mathbf{d}_R^\epsilon \rangle$  as follows:

$$\langle \mathcal{S}(\mathbf{d}_R^\epsilon), \mathbf{d}_R^\epsilon \rangle \leq C_{11} (1 + |\mathbf{d}_R^\epsilon|_{H^2}^2 + |\frac{\mathbf{d}_R^\epsilon}{\sqrt{\epsilon}}|_{L^2} + |\frac{\nabla \mathbf{d}_R^\epsilon}{\sqrt{\epsilon}}|_{L^2}) |\frac{\mathbf{d}_R^\epsilon}{\sqrt{\epsilon}}|_{L^2}, \quad (3.16)$$

and

$$\begin{aligned} \langle \mathcal{R}(\mathbf{d}_R^\epsilon), \mathbf{d}_R^\epsilon \rangle &\leq C_{12} \left[ \sqrt{\epsilon} + |\mathbf{d}_R^\epsilon|_{H^1} + |\partial_t \mathbf{d}_R^\epsilon|_{L^2} + \sqrt{\epsilon} |\partial_t \mathbf{d}_R^\epsilon|_{H^1}^2 + \sqrt{\epsilon} |\mathbf{d}_R^\epsilon|_{H^2}^2 \right. \\ &\quad \left. + |\mathbf{d}_R^\epsilon|_{H^2} (\epsilon |\partial_t \mathbf{d}_R^\epsilon|_{H^1}^2 + |\nabla \mathbf{d}_R^\epsilon|_{H^1}^2) \right] |\mathbf{d}_R^\epsilon|_{L^2}. \end{aligned} \quad (3.17)$$

So plugging (3.16) and (3.17) into (3.15) we obtain:

$$\begin{aligned} &\frac{1}{2} \frac{d}{dt} \left[ |\partial_t \mathbf{d}_R^\epsilon + \mathbf{d}_R^\epsilon|_{L^2}^2 + \left(\frac{1}{\epsilon} - 1\right) |\mathbf{d}_R^\epsilon|_{L^2}^2 - |\partial_t \mathbf{d}_R^\epsilon|_{L^2}^2 \right] - |\partial_t \mathbf{d}_R^\epsilon|_{L^2}^2 + \frac{1}{\epsilon} |\nabla \mathbf{d}_R^\epsilon|_{L^2}^2 \\ &\leq C_1 \left\{ (1 + |\mathbf{d}_R^\epsilon|_{H^2}^2 + |\frac{\mathbf{d}_R^\epsilon}{\sqrt{\epsilon}}|_{L^2} + |\frac{\nabla \mathbf{d}_R^\epsilon}{\sqrt{\epsilon}}|_{L^2}) |\frac{\mathbf{d}_R^\epsilon}{\sqrt{\epsilon}}|_{L^2} + \left[ \sqrt{\epsilon} + |\mathbf{d}_R^\epsilon|_{H^1} + |\partial_t \mathbf{d}_R^\epsilon|_{L^2} \right. \right. \\ &\quad \left. \left. + \sqrt{\epsilon} |\partial_t \mathbf{d}_R^\epsilon|_{H^1}^2 + \sqrt{\epsilon} |\mathbf{d}_R^\epsilon|_{H^2}^2 + |\mathbf{d}_R^\epsilon|_{H^2} (\epsilon |\partial_t \mathbf{d}_R^\epsilon|_{H^1}^2 + |\nabla \mathbf{d}_R^\epsilon|_{H^1}^2) \right] |\mathbf{d}_R^\epsilon|_{L^2} \right\}. \end{aligned} \quad (3.18)$$

Multiplying the inequality (3.18) by  $\frac{1}{2}$  and then adding it to the inequality (3.14), we get the  $L^2$ -energy estimate:

$$\begin{aligned} &\frac{1}{4} \frac{d}{dt} \left[ |\partial_t \mathbf{d}_R^\epsilon|_{L^2}^2 + |\partial_t \mathbf{d}_R^\epsilon + \mathbf{d}_R^\epsilon|_{L^2}^2 + \left(\frac{1}{\epsilon} - 1\right) |\mathbf{d}_R^\epsilon|_{L^2}^2 \right. \\ &\quad \left. + \frac{2}{\epsilon} |\nabla \mathbf{d}_R^\epsilon|_{L^2}^2 \right] + \left(\frac{1}{\epsilon} - \frac{1}{2}\right) |\partial_t \mathbf{d}_R^\epsilon|_{L^2}^2 + \frac{1}{2\epsilon} |\nabla \mathbf{d}_R^\epsilon|_{L^2}^2 \\ &\leq C_1 \left\{ (1 + |\mathbf{d}_R^\epsilon|_{H^2}^2 + |\frac{\mathbf{d}_R^\epsilon}{\sqrt{\epsilon}}|_{L^2} + |\frac{\nabla \mathbf{d}_R^\epsilon}{\sqrt{\epsilon}}|_{L^2}) \left( |\frac{\partial_t \mathbf{d}_R^\epsilon}{\sqrt{\epsilon}}|_{L^2} + |\frac{\mathbf{d}_R^\epsilon}{\sqrt{\epsilon}}|_{L^2} \right) \right. \\ &\quad \left. + \left[ \sqrt{\epsilon} + |\mathbf{d}_R^\epsilon|_{H^1} + |\partial_t \mathbf{d}_R^\epsilon|_{L^2} + \sqrt{\epsilon} |\partial_t \mathbf{d}_R^\epsilon|_{H^1}^2 + \sqrt{\epsilon} |\mathbf{d}_R^\epsilon|_{H^2}^2 \right. \right. \\ &\quad \left. \left. + |\mathbf{d}_R^\epsilon|_{H^2} (\epsilon |\partial_t \mathbf{d}_R^\epsilon|_{H^1}^2 + |\nabla \mathbf{d}_R^\epsilon|_{H^1}^2) \right] (|\partial_t \mathbf{d}_R^\epsilon|_{L^2} + |\mathbf{d}_R^\epsilon|_{L^2}) \right\}, \end{aligned} \quad (3.19)$$

where the constant

$$C_1 = C(1 + |\mathbf{D}(x)|_{H^3}^3 + |\partial_{tt} \mathbf{d}_0|_{L^\infty(0,T;L^2)} + |\partial_t \mathbf{d}_0|_{L^\infty(0,T;H^2)}^3 + |\nabla \mathbf{d}_0|_{L^\infty(0,T;H^2)}^4) > 0$$

for some computable positive constant  $C$ .

*Step 2. Higher order estimates.* In order to use the inequality (3.19) we also need a higher order estimate. To obtain this we take  $\nabla^k$  ( $k = 1, 2$ ) in the equation (1.15), we multiply it by  $\nabla^k \partial_t \mathbf{d}_R^\epsilon$ , integrate over  $\mathbb{R}^3$  and by parts, thus obtaining the following equality

$$\begin{aligned} &\frac{1}{2} \frac{d}{dt} (|\partial_t \nabla^k \mathbf{d}_R^\epsilon|_{L^2}^2 + \frac{1}{\epsilon} |\nabla^{k+1} \mathbf{d}_R^\epsilon|_{L^2}^2) + \frac{1}{\epsilon} |\partial_t \nabla^k \mathbf{d}_R^\epsilon|_{L^2}^2 \\ &= \langle \nabla^k \mathcal{S}(\mathbf{d}_R^\epsilon), \partial_t \nabla^k \mathbf{d}_R^\epsilon \rangle + \langle \nabla^k \mathcal{R}(\mathbf{d}_R^\epsilon), \partial_t \nabla^k \mathbf{d}_R^\epsilon \rangle \end{aligned} \quad (3.20)$$

(I) *Estimates of the singular terms  $\langle \nabla^k \mathcal{S}(\mathbf{d}_R^\epsilon), \partial_t \nabla^k \mathbf{d}_R^\epsilon \rangle$ :*

The singular terms can be divided into three parts:  $\langle \nabla^k \mathcal{S}_i, \partial_t \nabla^k \mathbf{d}_R^\epsilon \rangle (i = 1, 2, 3)$  which we estimate separately.

For the term  $\langle \nabla^k \mathcal{S}_1, \partial_t \nabla^k \mathbf{d}_R^\epsilon \rangle$ , by using the Hölder inequality and the Sobolev embedding theorems, we obtain:

$$\frac{1}{\sqrt{\epsilon}} \langle \nabla^k (\partial_{tt} \mathbf{d}_0 + \Delta \mathbf{D}(x) \exp(-\frac{t}{\epsilon})), \partial_t \nabla^k \mathbf{d}_R^\epsilon \rangle \lesssim (|\partial_{tt} \mathbf{d}_0|_{L_t^\infty H_x^2} + |\mathbf{D}(x)|_{H^4}) \left| \frac{\partial_t \mathbf{d}_R^\epsilon}{\sqrt{\epsilon}} \right|_{H^2}$$

and

$$\begin{aligned} & \frac{1}{\sqrt{\epsilon}} \langle \nabla^k (|\partial_t \mathbf{d}_0 + \mathbf{D}(x) \exp(-\frac{t}{\epsilon})|^2 \mathbf{d}_0), \partial_t \nabla^k \mathbf{d}_R^\epsilon \rangle \\ & \lesssim \frac{1}{\sqrt{\epsilon}} \sum_{\substack{i+j+e=k \\ e \geq 1}} |\langle \nabla^i (\partial_t \mathbf{d}_0 + \mathbf{D}(x)) \nabla^j (\partial_t \mathbf{d}_0 + \mathbf{D}(x)) \nabla^e \mathbf{d}_0, \partial_t \nabla^k \mathbf{d}_R^\epsilon \rangle| \\ & \quad + \frac{1}{\sqrt{\epsilon}} \sum_{i+j=k} |\langle \nabla^i (\partial_t \mathbf{d}_0 + \mathbf{D}(x)) \nabla^j (\partial_t \mathbf{d}_0 + \mathbf{D}(x)) \mathbf{d}_0, \partial_t \nabla^k \mathbf{d}_R^\epsilon \rangle| \\ & \lesssim (1 + |\nabla \mathbf{d}_0|_{L_t^\infty H_x^2}) (|\partial_t \mathbf{d}_0|_{L_t^\infty H_x^2}^2 + |\mathbf{D}(x)|_{H^2}^2) \left| \frac{\partial_t \mathbf{d}_R^\epsilon}{\sqrt{\epsilon}} \right|_{H^2}. \end{aligned}$$

Similarly as for estimating  $\frac{1}{\sqrt{\epsilon}} \langle \nabla^k (|\partial_t \mathbf{d}_0 + \mathbf{D}(x) \exp(-\frac{t}{\epsilon})|^2 \mathbf{d}_0), \partial_t \nabla^k \mathbf{d}_R^\epsilon \rangle$ , we can easily get the following estimates:

$$\begin{aligned} & \frac{1}{\sqrt{\epsilon}} \langle \nabla^k (|\nabla \mathbf{d}_0|^2 \mathbf{D}(x) \exp(-\frac{t}{\epsilon})), \partial_t \nabla^k \mathbf{d}_R^\epsilon \rangle \\ & \lesssim \frac{1}{\sqrt{\epsilon}} \sum_{i+j+e=k} |\langle \nabla^{i+1} \mathbf{d}_0 \nabla^{j+1} \mathbf{d}_0 \nabla^e \mathbf{D}(x), \partial_t \nabla^k \mathbf{d}_R^\epsilon \rangle| \\ & \lesssim |\nabla \mathbf{d}_0|_{L_t^\infty H_x^2}^2 |\mathbf{D}(x)|_{H^3} \left| \frac{\partial_t \mathbf{d}_R^\epsilon}{\sqrt{\epsilon}} \right|_{H^2} \end{aligned}$$

and

$$\begin{aligned} & \frac{1}{\sqrt{\epsilon}} \langle \nabla^k (\nabla \mathbf{d}_0 \cdot \nabla \mathbf{D}(x) \exp(-\frac{t}{\epsilon}) \mathbf{d}_0), \partial_t \nabla^k \mathbf{d}_R^\epsilon \rangle \\ & \lesssim \frac{1}{\sqrt{\epsilon}} \sum_{\substack{i+j+e=k \\ e \geq 1}} |\langle \nabla^{i+1} \mathbf{d}_0 \cdot \nabla^{j+1} \mathbf{D}(x) \nabla^e \mathbf{d}_0, \partial_t \nabla^k \mathbf{d}_R^\epsilon \rangle| \\ & \quad + \frac{1}{\sqrt{\epsilon}} \sum_{i+j=k} |\langle \nabla^{i+1} \mathbf{d}_0 \cdot \nabla^{j+1} \mathbf{D}(x) \mathbf{d}_0, \partial_t \nabla^k \mathbf{d}_R^\epsilon \rangle| \\ & \lesssim (1 + |\nabla \mathbf{d}_0|_{L_t^\infty H_x^2}) |\nabla \mathbf{d}_0|_{L_t^\infty H_x^3} |\mathbf{D}(x)|_{H^3} \left| \frac{\partial_t \mathbf{d}_R^\epsilon}{\sqrt{\epsilon}} \right|_{H^2}. \end{aligned}$$

Thus we have the estimate of  $\langle \nabla^k \mathcal{S}_1, \partial_t \nabla^k \mathbf{d}_R^\epsilon \rangle$  as follows:

$$\begin{aligned} \langle \nabla^k \mathcal{S}_1, \partial_t \nabla^k \mathbf{d}_R^\epsilon \rangle & \lesssim \left\{ (1 + |\nabla \mathbf{d}_0|_{L_t^\infty H_x^2}) (|\partial_t \mathbf{d}_0|_{L_t^\infty H_x^3}^2 + |\nabla \mathbf{d}_0|_{L_t^\infty H_x^3}^2 + |\mathbf{D}(x)|_{H^3}^2) \right. \\ & \quad \left. + |\partial_{tt} \mathbf{d}_0|_{L_t^\infty H_x^2} + |\mathbf{D}(x)|_{H^4} \right\} \left| \frac{\partial_t \mathbf{d}_R^\epsilon}{\sqrt{\epsilon}} \right|_{H^2}. \end{aligned} \quad (3.21)$$

For the term  $\langle \nabla^k \mathcal{S}_2, \partial_t \nabla^k \mathbf{d}_R^\epsilon \rangle$ , we can also use the Hölder inequality and Sobolev embedding theorems to get

$$\begin{aligned} & \frac{1}{\sqrt{\epsilon}} \langle \nabla^k ((\nabla \mathbf{d}_0 \cdot \nabla \mathbf{d}_R^\epsilon) \mathbf{d}_R^\epsilon), \partial_t \nabla^k \mathbf{d}_R^\epsilon \rangle \\ & \lesssim \frac{1}{\sqrt{\epsilon}} \sum_{\substack{i+j+e=k \\ e \geq 1}} |\langle \nabla^{i+1} \mathbf{d}_0 \cdot \nabla^{j+1} \mathbf{d}_R^\epsilon \nabla^e \mathbf{d}_R^\epsilon, \partial_t \nabla^k \mathbf{d}_R^\epsilon \rangle| \\ & \quad + \frac{1}{\sqrt{\epsilon}} \sum_{i+j=k} |\langle \nabla^{i+1} \mathbf{d}_0 \cdot \nabla^{j+1} \mathbf{d}_R^\epsilon \mathbf{d}_R^\epsilon, \partial_t \nabla^k \mathbf{d}_R^\epsilon \rangle| \\ & \lesssim |\nabla \mathbf{d}_0|_{L_t^\infty H_x^4} |\mathbf{d}_R^\epsilon|_{H^2} |\nabla \mathbf{d}_R^\epsilon|_{H^2} \left| \frac{\partial_t \mathbf{d}_R^\epsilon}{\sqrt{\epsilon}} \right|_{H^2}, \end{aligned}$$

and

$$\begin{aligned}
& \frac{1}{\sqrt{\varepsilon}} \langle \nabla^k (|\nabla \mathbf{d}_R^\varepsilon|^2 \mathbf{d}_0), \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle \\
& \lesssim \frac{1}{\sqrt{\varepsilon}} \sum_{\substack{i+j+e=k \\ e \geq 1}} |\langle \nabla^{i+1} \mathbf{d}_R^\varepsilon \nabla^{j+1} \mathbf{d}_R^\varepsilon \nabla^e \mathbf{d}_0, \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle| \\
& \quad + \frac{1}{\sqrt{\varepsilon}} \sum_{i+j=k} |\langle \nabla^{i+1} \mathbf{d}_R^\varepsilon \nabla^{j+1} \mathbf{d}_R^\varepsilon \mathbf{d}_0, \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle| \\
& \lesssim (1 + |\nabla \mathbf{d}_0|_{L_t^\infty H_x^3}) |\nabla \mathbf{d}_R^\varepsilon|_{H^2}^2 \left| \frac{\partial_t \mathbf{d}_R^\varepsilon}{\sqrt{\varepsilon}} \right|_{H^2}.
\end{aligned}$$

Summarizing, we obtain

$$\langle \nabla^k \mathcal{S}_2, \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle \lesssim (1 + |\nabla \mathbf{d}_0|_{L_t^\infty H_x^4}) (|\mathbf{d}_R^\varepsilon|_{H^2} + |\nabla \mathbf{d}_R^\varepsilon|_{H^2}) |\nabla \mathbf{d}_R^\varepsilon|_{H^2} \left| \frac{\partial_t \mathbf{d}_R^\varepsilon}{\sqrt{\varepsilon}} \right|_{H^2}. \quad (3.22)$$

For the estimate of  $\langle \nabla^k \mathcal{S}_3, \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle$ , we get the estimate of the first term by using again the Hölder inequality and the Sobolev embedding theorems:

$$\begin{aligned}
& \frac{1}{\varepsilon} \langle \nabla^k (|\nabla \mathbf{d}_0|^2 \mathbf{d}_R^\varepsilon), \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle \\
& = \frac{1}{\varepsilon} \sum_{i+j+e=k} \langle \nabla^{i+1} \mathbf{d}_0 \nabla^{j+1} \mathbf{d}_0 \nabla^e \mathbf{d}_R^\varepsilon, \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle \\
& \lesssim |\nabla \mathbf{d}_0|_{L_t^\infty H_x^4}^2 \left| \frac{\mathbf{d}_R^\varepsilon}{\sqrt{\varepsilon}} \right|_{H^2} \left| \frac{\partial_t \mathbf{d}_R^\varepsilon}{\sqrt{\varepsilon}} \right|_{H^2}.
\end{aligned}$$

Recalling that  $|\mathbf{d}_0| = 1$ , one can easily estimate the second term

$$\begin{aligned}
& \frac{1}{\varepsilon} \langle \nabla^k ((\nabla \mathbf{d}_0 \cdot \nabla \mathbf{d}_R^\varepsilon) \mathbf{d}_0), \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle \\
& \lesssim \frac{1}{\varepsilon} \sum_{\substack{i+j+e=k \\ e \geq 1}} |\langle \nabla^{i+1} \mathbf{d}_0 \cdot \nabla^{j+1} \mathbf{d}_R^\varepsilon \nabla^e \mathbf{d}_0, \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle| \\
& \quad + \frac{1}{\varepsilon} \sum_{i+j=k} |\langle \nabla^{i+1} \mathbf{d}_0 \cdot \nabla^{j+1} \mathbf{d}_R^\varepsilon \mathbf{d}_0, \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle| \\
& \lesssim (1 + |\nabla \mathbf{d}_0|_{L_t^\infty H_x^3}) |\nabla \mathbf{d}_0|_{L_t^\infty H_x^4} \left| \frac{\nabla \mathbf{d}_R^\varepsilon}{\sqrt{\varepsilon}} \right|_{H^2} \left| \frac{\partial_t \mathbf{d}_R^\varepsilon}{\sqrt{\varepsilon}} \right|_{H^2}.
\end{aligned}$$

Thus by the above two estimates we have

$$\langle \nabla^k \mathcal{S}_3, \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle \lesssim (1 + |\nabla \mathbf{d}_0|_{L_t^\infty H_x^4}) |\nabla \mathbf{d}_0|_{L_t^\infty H_x^4} \left( \left| \frac{\mathbf{d}_R^\varepsilon}{\sqrt{\varepsilon}} \right|_{H^2} + \left| \frac{\nabla \mathbf{d}_R^\varepsilon}{\sqrt{\varepsilon}} \right|_{H^2} \right) \left| \frac{\partial_t \mathbf{d}_R^\varepsilon}{\sqrt{\varepsilon}} \right|_{H^2}. \quad (3.23)$$

Then the inequalities (3.21), (3.22) and (3.23) give the following estimate

$$\langle \nabla^k \mathcal{S}(\mathbf{d}_R^\varepsilon), \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle \leq C_{k1} (1 + |\mathbf{d}_R^\varepsilon|_{H^2}^2 + |\nabla \mathbf{d}_R^\varepsilon|_{H^2}^2 + \left| \frac{\mathbf{d}_R^\varepsilon}{\sqrt{\varepsilon}} \right|_{H^2} + \left| \frac{\nabla \mathbf{d}_R^\varepsilon}{\sqrt{\varepsilon}} \right|_{H^2}) \left| \frac{\partial_t \mathbf{d}_R^\varepsilon}{\sqrt{\varepsilon}} \right|_{H^2}, \quad (3.24)$$

where the positive constant  $C_{k1}$  is

$$C_{k1} = C \left\{ (1 + |\nabla \mathbf{d}_0|_{L_t^\infty H_x^4}) (1 + |\partial_t \mathbf{d}_0|_{L_t^\infty H_x^3}^2 + |\nabla \mathbf{d}_0|_{L_t^\infty H_x^4}^2 + |\mathbf{D}(x)|_{H^3}^2) + |\partial_{tt} \mathbf{d}_0|_{L_t^\infty H_x^2} + |\mathbf{D}(x)|_{H^4} \right\}$$

for some computable positive constant  $C$ .

(II) *Estimates of the regular terms*  $\langle \nabla^k \mathcal{R}(\mathbf{d}_R^\varepsilon), \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle$ :

Finally, we turn to estimating the regular terms  $\langle \nabla^k \mathcal{R}(\mathbf{d}_R^\varepsilon), \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle$ , which are divided into four parts:  $\langle \nabla^k \mathcal{R}_i, \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle$  ( $i = 1, 2, 3$ ).

For the terms  $\langle \nabla^k \mathcal{R}_1, \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle$ , by the Hölder inequality and the Sobolev embedding theorems we have:

$$\begin{aligned}
& \sqrt{\varepsilon} \langle \nabla^k (|\partial_t \mathbf{d}_0 + \mathbf{D}(x) \exp(-\frac{t}{\varepsilon})|^2 \mathbf{D}(x) \exp(-\frac{t}{\varepsilon})), \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle \\
& \lesssim \sqrt{\varepsilon} \sum_{i+j+e=k} |\langle \nabla^i (\partial_t \mathbf{d}_0 + \mathbf{D}(x)) \nabla^j (\partial_t \mathbf{d}_0 + \mathbf{D}(x)) \nabla^e \mathbf{D}(x), \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle| \\
& \lesssim \sqrt{\varepsilon} |\mathbf{D}(x)|_{H^3} (|\mathbf{D}(x)|_{H^2}^2 + |\partial_t \mathbf{d}_0|_{L_t^\infty H_x^2}^2) |\partial_t \mathbf{d}_R^\varepsilon|_{H^2},
\end{aligned}$$

and

$$\begin{aligned}
& \sqrt{\varepsilon} \langle \nabla^k (|\nabla \mathbf{D}(x)|^2 \exp(-\frac{2t}{\varepsilon}) \mathbf{d}_0), \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle \\
& \lesssim \sqrt{\varepsilon} \sum_{\substack{i+j+e=k \\ e \geq 1}} |\langle \nabla^{i+1} \mathbf{D}(x) \nabla^{j+1} \mathbf{D}(x) \nabla^e \mathbf{d}_0, \partial_t \nabla^k \mathbf{d}_0 \rangle| \\
& \quad + \sqrt{\varepsilon} \sum_{i+j=k} |\langle \nabla^{i+1} \mathbf{D}(x) \nabla^{j+1} \mathbf{D}(x) \mathbf{d}_0, \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle| \\
& \lesssim \sqrt{\varepsilon} (1 + |\nabla \mathbf{d}_0|_{L_t^\infty H_x^2}) |\mathbf{D}(x)|_{H^3}^2 |\partial_t \mathbf{d}_R^\varepsilon|_{H^2}.
\end{aligned}$$

We can estimate the following two terms in a similar way, hence we get the following inequalities:

$$\begin{aligned}
& \varepsilon^{\frac{3}{2}} \langle \nabla^k (|\nabla \mathbf{D}(x)|^2 \exp(-\frac{3t}{\varepsilon}) \mathbf{D}(x)), \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle \\
& \lesssim \varepsilon^{\frac{3}{2}} \sum_{i+j+e=k} |\langle \nabla^{i+1} \mathbf{D}(x) \nabla^{j+1} \mathbf{D}(x) \nabla^e \mathbf{D}(x), \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle| \\
& \lesssim \varepsilon^{\frac{3}{2}} |\mathbf{D}(x)|_{H^3}^3 |\partial_t \mathbf{d}_R^\varepsilon|_{H^2}
\end{aligned}$$

and

$$\begin{aligned}
& \sqrt{\varepsilon} \langle \nabla^k (\nabla \mathbf{d}_0 \cdot \nabla \mathbf{D}(x)) \exp(-\frac{2t}{\varepsilon}) \mathbf{D}(x), \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle \\
& \lesssim \sqrt{\varepsilon} \sum_{i+j+e=k} |\langle \nabla^{i+1} \mathbf{d}_0 \nabla^{j+1} \mathbf{D}(x) \nabla^e \mathbf{D}(x), \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle| \\
& \lesssim \sqrt{\varepsilon} |\nabla \mathbf{d}_0|_{L_t^\infty H_x^4} |\mathbf{D}(x)|_{H^3}^2 |\partial_t \mathbf{d}_R^\varepsilon|_{H^2}.
\end{aligned}$$

So we have the estimate of  $\langle \nabla^k \mathcal{R}_1, \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle$  as follows:

$$\begin{aligned}
\langle \nabla^k \mathcal{R}_1, \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle & \lesssim \sqrt{\varepsilon} (1 + |\nabla \mathbf{d}_0|_{L_t^\infty H_x^4} + |\mathbf{D}(x)|_{H^3}) \\
& \quad \times (|\partial_t \mathbf{d}_0|_{L_t^\infty H_x^2}^2 + |\mathbf{D}(x)|_{H^3}^2) |\partial_t \mathbf{d}_R^\varepsilon|_{H^2}.
\end{aligned} \tag{3.25}$$

For the terms  $\langle \nabla^k \mathcal{R}_2, \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle$ , by using yet again the Hölder inequality and the Sobolev embedding theorems, we have

$$\begin{aligned}
& \varepsilon \langle \nabla^k (\partial_t \mathbf{d}_0 + \mathbf{D}(x) \exp(-\frac{t}{\varepsilon})) \partial_t \mathbf{d}_R^\varepsilon D(x) \exp(-\frac{t}{\varepsilon}), \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle \\
& \lesssim \varepsilon \sum_{\substack{i+j+e=k \\ e \geq 1}} |\langle (\nabla^i \partial_t \mathbf{d}_0 + \nabla^i \mathbf{D}(x) \exp(-\frac{t}{\varepsilon})) \nabla^j \partial_t \mathbf{d}_R^\varepsilon \nabla^e D(x) \exp(-\frac{t}{\varepsilon}), \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle| \\
& \quad + \varepsilon \sum_{i+j=k} |\langle (\nabla^i \partial_t \mathbf{d}_0 + \nabla^i \mathbf{D}(x) \exp(-\frac{t}{\varepsilon})) \nabla^j \partial_t \mathbf{d}_R^\varepsilon D(x) \exp(-\frac{t}{\varepsilon}), \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle| \\
& \lesssim \varepsilon (|\partial_t \mathbf{d}_0|_{L_t^\infty H_x^4} + |\mathbf{D}(x)|_{H^4}) |\mathbf{D}(x)|_{H^4} |\partial_t \mathbf{d}_R^\varepsilon|_{H^2}^2,
\end{aligned}$$

Similarly as for estimating the term  $\varepsilon \langle \nabla^k (\partial_t \mathbf{d}_0 + \mathbf{D}(x) \exp(-\frac{t}{\varepsilon})) \partial_t \mathbf{d}_R^\varepsilon D(x) \exp(-\frac{t}{\varepsilon}), \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle$  we can easily obtain the following estimates:

$$\begin{aligned}
\langle \nabla^k (|\partial_t \mathbf{d}_0 + \mathbf{D}(x) \exp(-\frac{t}{\varepsilon})|^2 \mathbf{d}_R^\varepsilon), \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle & \lesssim (|\partial_t \mathbf{d}_0|_{L_t^\infty H_x^4}^2 + |\mathbf{D}(x)|_{H^4}^2) |\mathbf{d}_R^\varepsilon|_{H^2} |\partial_t \mathbf{d}_R^\varepsilon|_{H^2}, \\
\langle \nabla^k (\nabla \mathbf{d}_0 \cdot \nabla \mathbf{d}_R^\varepsilon \mathbf{D}(x) \exp(-\frac{t}{\varepsilon})), \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle & \lesssim (|\nabla \mathbf{d}_0|_{L_t^\infty H_x^4}^2 + |\mathbf{D}(x)|_{H^4}^2) |\nabla \mathbf{d}_R^\varepsilon|_{H^2} |\partial_t \mathbf{d}_R^\varepsilon|_{H^2}, \\
\langle \nabla^k (\nabla \mathbf{d}_0 \cdot \nabla \mathbf{D}(x) \exp(-\frac{t}{\varepsilon}) \mathbf{d}_R^\varepsilon), \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle & \lesssim (|\nabla \mathbf{d}_0|_{L_t^\infty H_x^4}^2 + |\mathbf{D}(x)|_{H^4}^2) |\mathbf{d}_R^\varepsilon|_{H^2} |\partial_t \mathbf{d}_R^\varepsilon|_{H^2}.
\end{aligned}$$

Observing the structure of the terms  $\varepsilon \langle \nabla^k (|\nabla \mathbf{D}(x)|^2 \exp(-\frac{2t}{\varepsilon}) \mathbf{d}_R^\varepsilon), \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle$  and  $\varepsilon \langle \nabla^k (\nabla \mathbf{D}(x) \cdot \nabla \mathbf{d}_R^\varepsilon \exp(-\frac{2t}{\varepsilon}) \mathbf{D}(x)), \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle$  one can similarly estimate the following terms:

$$\begin{aligned}
\varepsilon \langle \nabla^k (|\nabla \mathbf{D}(x)|^2 \exp(-\frac{2t}{\varepsilon}) \mathbf{d}_R^\varepsilon), \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle & \lesssim \varepsilon |\mathbf{D}(x)|_{H^5}^2 |\mathbf{d}_R^\varepsilon|_{H^2} |\partial_t \mathbf{d}_R^\varepsilon|_{H^2}, \\
\varepsilon \langle \nabla^k (\nabla \mathbf{D}(x) \cdot \nabla \mathbf{d}_R^\varepsilon \exp(-\frac{2t}{\varepsilon}) \mathbf{D}(x)), \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle & \lesssim \varepsilon |\mathbf{D}(x)|_{H^5}^2 |\nabla \mathbf{d}_R^\varepsilon|_{H^2} |\partial_t \mathbf{d}_R^\varepsilon|_{H^2}.
\end{aligned}$$

Furthermore we get:

$$\begin{aligned}
& \langle \nabla^k [(\partial_t \mathbf{d}_0 + \mathbf{D}(x) \exp(-\frac{t}{\varepsilon})) \cdot \partial_t \mathbf{d}_R^\varepsilon \mathbf{d}_0], \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle \\
& \lesssim \sum_{\substack{i+j+e=k \\ e \geq 1}} |\langle \nabla^i (\partial_t \mathbf{d}_0 + \mathbf{D}(x)) \nabla^j \partial_t \mathbf{d}_R^\varepsilon \nabla^e \mathbf{d}_0, \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle| \\
& \quad + \sum_{i+j=k} |\langle \nabla^i (\partial_t \mathbf{d}_0 + \mathbf{D}(x)) \nabla^j \partial_t \mathbf{d}_R^\varepsilon \mathbf{d}_0, \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle| \\
& \lesssim (1 + |\nabla \mathbf{d}_0|_{L_t^\infty H_x^3}) (|\partial_t \mathbf{d}_0|_{L_t^\infty H_x^4} + |\mathbf{D}(x)|_{H^4}) |\partial_t \mathbf{d}_R^\varepsilon|_{H^2}^2
\end{aligned}$$

Similarly as for estimating  $\langle \nabla^k [(\partial_t \mathbf{d}_0 + \mathbf{D}(x) \exp(-\frac{t}{\varepsilon})) \cdot \partial_t \mathbf{d}_R^\varepsilon \mathbf{d}_0], \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle$ , one can also get:

$$\langle \nabla^k (\nabla \mathbf{D}(x) \cdot \nabla \mathbf{d}_R^\varepsilon \exp(-\frac{t}{\varepsilon}) \mathbf{d}_0), \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle \lesssim (1 + |\nabla \mathbf{d}_0|_{L_t^\infty H_x^3}) |\mathbf{D}(x)|_{H^5} |\nabla \mathbf{d}_R^\varepsilon|_{H^2} |\partial_t \mathbf{d}_R^\varepsilon|_{H^2}.$$

Thus we have the following estimate of  $\langle \nabla^k \mathcal{R}_2, \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle$ :

$$\begin{aligned}
\langle \nabla^k \mathcal{R}_2, \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle & \lesssim (|\partial_t \mathbf{d}_0|_{L_t^\infty H_x^4} + |\mathbf{D}(x)|_{H^5} + |\partial_t \mathbf{d}_0|_{L_t^\infty H_x^4}^2 + |\mathbf{D}(x)|_{H^5}^2 + |\nabla \mathbf{d}_0|_{L_t^\infty H_x^4}^2) \\
& \quad \times (|\mathbf{d}_R^\varepsilon|_{H^2} + |\nabla \mathbf{d}_R^\varepsilon|_{H^2} + |\partial_t \mathbf{d}_R^\varepsilon|_{H^2}) |\partial_t \mathbf{d}_R^\varepsilon|_{H^2}. \tag{3.26}
\end{aligned}$$

For the terms  $\langle \nabla^k \mathcal{R}_3, \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle$ , one can use an estimate similar to the one for the term  $\varepsilon \langle \overline{\nabla^k (\partial_t \mathbf{d}_0 \cdot \mathbf{D}(x) \exp(-\frac{t}{\varepsilon}) \partial_t \mathbf{d}_R^\varepsilon)}, \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle$  to get:

$$\begin{aligned}
\varepsilon^{\frac{3}{2}} \langle \nabla^k (\mathbf{D}(x) \exp(-\frac{t}{\varepsilon}) |\partial_t \mathbf{d}_R^\varepsilon|^2), \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle & \lesssim \varepsilon^{\frac{3}{2}} |\mathbf{D}(x)|_{H^4} |\partial_t \mathbf{d}_R^\varepsilon|_{H^2}^3, \\
\sqrt{\varepsilon} \langle \nabla^k (\nabla \mathbf{D}(x) \cdot \nabla \mathbf{d}_R^\varepsilon \exp(-\frac{t}{\varepsilon}) \mathbf{d}_R^\varepsilon), \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle & \lesssim \sqrt{\varepsilon} |\mathbf{D}(x)|_{H^5} (|\mathbf{d}_R^\varepsilon|_{H^2}^2 + |\nabla \mathbf{d}_R^\varepsilon|_{H^2}^2) |\partial_t \mathbf{d}_R^\varepsilon|_{H^2}, \\
\sqrt{\varepsilon} \langle \nabla^k (|\nabla \mathbf{d}_R^\varepsilon|^2 \mathbf{D}(x) \exp(-\frac{t}{\varepsilon})), \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle & \lesssim \sqrt{\varepsilon} |\mathbf{D}(x)|_{H^4} |\nabla \mathbf{d}_R^\varepsilon|_{H^2}^2 |\partial_t \mathbf{d}_R^\varepsilon|_{H^2}.
\end{aligned}$$

Reasoning analogously as in estimating  $\sqrt{\varepsilon} \langle \nabla^k [(\partial_t \mathbf{d}_0 + \mathbf{D}(x) \exp(-\frac{t}{\varepsilon})) \cdot \partial_t \mathbf{d}_R^\varepsilon \mathbf{d}_R^\varepsilon], \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle$  we have

$$\begin{aligned}
& \sqrt{\varepsilon} \langle \nabla^k [(\partial_t \mathbf{d}_0 + \mathbf{D}(x) \exp(-\frac{t}{\varepsilon})) \cdot \partial_t \mathbf{d}_R^\varepsilon \mathbf{d}_R^\varepsilon], \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle \\
& \lesssim \sqrt{\varepsilon} \sum_{\substack{i+j+e=k \\ e \geq 1}} |\langle \nabla^i (\partial_t \mathbf{d}_0 + \mathbf{D}(x)) \cdot \nabla^j \partial_t \mathbf{d}_R^\varepsilon \nabla^e \mathbf{d}_R^\varepsilon, \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle| \\
& \quad + \sqrt{\varepsilon} \sum_{i+j=k} |\langle \nabla^i (\partial_t \mathbf{d}_0 + \mathbf{D}(x)) \cdot \nabla^j \partial_t \mathbf{d}_R^\varepsilon \mathbf{d}_R^\varepsilon, \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle| \\
& \lesssim \sqrt{\varepsilon} (|\partial_t \mathbf{d}_0|_{L_t^\infty H_x^4} + |\mathbf{D}(x)|_{H^4}) (|\mathbf{d}_R^\varepsilon|_{H^2} + |\nabla \mathbf{d}_R^\varepsilon|_{H^2}) |\partial_t \mathbf{d}_R^\varepsilon|_{H^2}^2
\end{aligned}$$

and furthermore, using that  $|\mathbf{d}_0| = 1$ , it is easy to obtain

$$\sqrt{\varepsilon} \langle \nabla^k (|\partial_t \mathbf{d}_R^\varepsilon|^2 \mathbf{d}_0), \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle \lesssim \sqrt{\varepsilon} (1 + |\nabla \mathbf{d}_0|_{L_t^\infty H_x^3}) |\partial_t \mathbf{d}_R^\varepsilon|_{H^2}^3.$$

Summarizing, we get the estimate of  $\langle \nabla^k \mathcal{R}_3, \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle$  as follows:

$$\begin{aligned}
\langle \nabla^k \mathcal{R}_3, \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle & \lesssim \sqrt{\varepsilon} (1 + |\partial_t \mathbf{d}_0|_{L_t^\infty H_x^4} + |\nabla \mathbf{d}_0|_{L_t^\infty H_x^3} + |\mathbf{D}(x)|_{H^5}) \\
& \quad \times (1 + |\partial_t \mathbf{d}_R^\varepsilon|_{H^2}^2 + |\mathbf{d}_R^\varepsilon|_{H^2}^2 + |\nabla \mathbf{d}_R^\varepsilon|_{H^2}^2) |\partial_t \mathbf{d}_R^\varepsilon|_{H^2}. \tag{3.27}
\end{aligned}$$

For the terms  $\langle \nabla^k \mathcal{R}_4, \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle$ , we get, by the Hölder inequality and Sobolev embedding theorems:

$$\begin{aligned}
& \varepsilon \langle \nabla^k (|\partial_t \mathbf{d}_R^\varepsilon|^2 \mathbf{d}_R^\varepsilon), \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle \\
& = \varepsilon \sum_{\substack{i+j+e=k \\ e \geq 1}} \langle \nabla^i \partial_t \mathbf{d}_R^\varepsilon \nabla^j \partial_t \mathbf{d}_R^\varepsilon \nabla^e \mathbf{d}_R^\varepsilon, \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle \\
& \quad + \varepsilon \sum_{i+j=k} \langle \nabla^i \partial_t \mathbf{d}_R^\varepsilon \nabla^j \partial_t \mathbf{d}_R^\varepsilon \mathbf{d}_R^\varepsilon, \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle
\end{aligned}$$



$$\lesssim \varepsilon (|\mathbf{d}_R^\varepsilon|_{H^2} + |\nabla \mathbf{d}_R^\varepsilon|_{H^2}) |\partial_t \mathbf{d}_R^\varepsilon|_{H^2}^3$$

and then, similarly:

$$\langle \nabla^k (|\nabla \mathbf{d}_R^\varepsilon|^2 \mathbf{d}_R^\varepsilon), \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle \lesssim (|\mathbf{d}_R^\varepsilon|_{H^2} + |\nabla \mathbf{d}_R^\varepsilon|_{H^2}) |\nabla \mathbf{d}_R^\varepsilon|_{H^2}^2 |\partial_t \mathbf{d}_R^\varepsilon|_{H^2}.$$

So we obtain the estimate of  $\langle \nabla^k \mathcal{R}_4, \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle$  as follows:

$$\langle \nabla^k \mathcal{R}_4, \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle \lesssim (|\mathbf{d}_R^\varepsilon|_{H^2} + |\nabla \mathbf{d}_R^\varepsilon|_{H^2}) (|\nabla \mathbf{d}_R^\varepsilon|_{H^2}^2 + \varepsilon |\partial_t \mathbf{d}_R^\varepsilon|_{H^2}^2) |\partial_t \mathbf{d}_R^\varepsilon|_{H^2}. \quad (3.28)$$

Then the inequalities (3.25), (3.26), (3.27) and (3.28) give the estimate of the regular terms  $\langle \nabla^k \mathcal{R}(\mathbf{d}_R^\varepsilon), \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle$  as follows:

$$\begin{aligned} \langle \nabla^k \mathcal{R}(\mathbf{d}_R^\varepsilon), \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle &\leq C_{k2} [ (|\mathbf{d}_R^\varepsilon|_{H^2} + |\nabla \mathbf{d}_R^\varepsilon|_{H^2}) (1 + |\nabla \mathbf{d}_R^\varepsilon|_{H^2}^2 + \varepsilon |\partial_t \mathbf{d}_R^\varepsilon|_{H^2}^2) \\ &\quad + \sqrt{\varepsilon} (1 + |\partial_t \mathbf{d}_R^\varepsilon|_{H^2}^2 + |\mathbf{d}_R^\varepsilon|_{H^2}^2 + |\nabla \mathbf{d}_R^\varepsilon|_{H^2}^2) + |\partial_t \mathbf{d}_R^\varepsilon|_{H^2} ] |\partial_t \mathbf{d}_R^\varepsilon|_{H^2}, \end{aligned} \quad (3.29)$$

where the constant  $C_{k2}$  is

$$C_{k2} = C(1 + |\partial_t \mathbf{d}_0|_{L_t^\infty H_x^4} + |\nabla \mathbf{d}_0|_{L_t^\infty H_x^4} + |\mathbf{D}(x)|_{H^5}) (1 + |\partial_t \mathbf{d}_0|_{L_t^\infty H_x^4}^2 + |\mathbf{D}(x)|_{H^5}^2 + |\nabla \mathbf{d}_0|_{L_t^\infty H_x^4}^2) > 0,$$

and  $C$  is a computable positive constant.

Therefore, by substituting the inequalities (3.24) and (3.29) into (3.20) one has

$$\begin{aligned} &\frac{1}{2} \frac{d}{dt} (|\partial_t \nabla^k \mathbf{d}_R^\varepsilon|_{L^2}^2 + \frac{1}{\varepsilon} |\nabla^{k+1} \mathbf{d}_R^\varepsilon|_{L^2}^2) + \frac{1}{\varepsilon} |\partial_t \nabla^k \mathbf{d}_R^\varepsilon|_{L^2}^2 \\ &\leq C_k \left\{ (1 + |\mathbf{d}_R^\varepsilon|_{H^2}^2 + |\nabla \mathbf{d}_R^\varepsilon|_{H^2}^2 + \frac{|\mathbf{d}_R^\varepsilon|_{H^2}^2}{\sqrt{\varepsilon}} + \frac{|\nabla \mathbf{d}_R^\varepsilon|_{H^2}^2}{\sqrt{\varepsilon}}) \frac{|\partial_t \mathbf{d}_R^\varepsilon|_{H^2}}{\sqrt{\varepsilon}} \right. \\ &\quad + [ (|\mathbf{d}_R^\varepsilon|_{H^2} + |\nabla \mathbf{d}_R^\varepsilon|_{H^2}) (1 + |\nabla \mathbf{d}_R^\varepsilon|_{H^2}^2 + \varepsilon |\partial_t \mathbf{d}_R^\varepsilon|_{H^2}^2) \\ &\quad \left. + \sqrt{\varepsilon} (1 + |\partial_t \mathbf{d}_R^\varepsilon|_{H^2}^2 + |\mathbf{d}_R^\varepsilon|_{H^2}^2 + |\nabla \mathbf{d}_R^\varepsilon|_{H^2}^2) + |\partial_t \mathbf{d}_R^\varepsilon|_{H^2} ] |\partial_t \mathbf{d}_R^\varepsilon|_{H^2} \right\}, \end{aligned} \quad (3.30)$$

where the positive constant  $C_k$  is

$$C_k = C(1 + |\partial_{tt} \mathbf{d}_0|_{L^\infty(0,T;H^2)} + |\partial_t \mathbf{d}_0|_{L^\infty(0,T;H^4)}^3 + |\nabla \mathbf{d}_0|_{L^\infty(0,T;H^4)}^3 + |\mathbf{D}(x)|_{H^5}^3) > 0$$

and  $C > 0$  is a computable constant.

(III) For the estimate of  $|\nabla^k \mathbf{d}_R^\varepsilon|_{L^2}$  ( $k = 1, 2$ ):

Applying  $\nabla^k$  ( $k = 1, 2$ ) to the remainder equation (1.15), multiplying by  $\nabla^k \mathbf{d}_R^\varepsilon$ , integrating over  $\mathbb{R}^3$  and by parts, we have

$$\begin{aligned} &\frac{1}{2} \frac{d}{dt} [ |\nabla^k \partial_t \mathbf{d}_R^\varepsilon + \nabla^k \mathbf{d}_R^\varepsilon|_{L^2}^2 + (\frac{1}{\varepsilon} - 1) |\nabla^k \mathbf{d}_R^\varepsilon|_{L^2}^2 - |\partial_t \nabla^k \mathbf{d}_R^\varepsilon|_{L^2}^2 ] - |\partial_t \nabla^k \mathbf{d}_R^\varepsilon|_{L^2}^2 + \frac{1}{\varepsilon} |\nabla^{k+1} \mathbf{d}_R^\varepsilon|_{L^2}^2 \\ &= \langle \nabla^k \mathcal{S}(\mathbf{d}_R^\varepsilon), \nabla^k \mathbf{d}_R^\varepsilon \rangle + \langle \nabla^k \mathcal{R}(\mathbf{d}_R^\varepsilon), \nabla^k \mathbf{d}_R^\varepsilon \rangle. \end{aligned} \quad (3.31)$$

Similarly as in the estimates of the terms  $\langle \nabla^k \mathcal{S}(\mathbf{d}_R^\varepsilon), \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle$  and  $\langle \nabla^k \mathcal{R}(\mathbf{d}_R^\varepsilon), \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle$  in the inequalities (3.24) and (3.29), respectively, we can analogously estimate the terms  $\langle \nabla^k \mathcal{S}(\mathbf{d}_R^\varepsilon), \nabla^k \mathbf{d}_R^\varepsilon \rangle$  and  $\langle \nabla^k \mathcal{R}(\mathbf{d}_R^\varepsilon), \nabla^k \mathbf{d}_R^\varepsilon \rangle$  as follows:

$$\langle \nabla^k \mathcal{S}(\mathbf{d}_R^\varepsilon), \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle \leq C_{k1} (1 + |\mathbf{d}_R^\varepsilon|_{H^2}^2 + |\nabla \mathbf{d}_R^\varepsilon|_{H^2}^2 + \frac{|\mathbf{d}_R^\varepsilon|_{H^2}^2}{\sqrt{\varepsilon}} + \frac{|\nabla \mathbf{d}_R^\varepsilon|_{H^2}^2}{\sqrt{\varepsilon}}) \frac{|\mathbf{d}_R^\varepsilon|_{H^2}}{\sqrt{\varepsilon}}, \quad (3.32)$$

and

$$\begin{aligned} \langle \nabla^k \mathcal{R}(\mathbf{d}_R^\varepsilon), \partial_t \nabla^k \mathbf{d}_R^\varepsilon \rangle &\leq C_{k2} [ (|\mathbf{d}_R^\varepsilon|_{H^2} + |\nabla \mathbf{d}_R^\varepsilon|_{H^2}) (1 + |\nabla \mathbf{d}_R^\varepsilon|_{H^2}^2 + \varepsilon |\partial_t \mathbf{d}_R^\varepsilon|_{H^2}^2) \\ &\quad + \sqrt{\varepsilon} (1 + |\partial_t \mathbf{d}_R^\varepsilon|_{H^2}^2 + |\mathbf{d}_R^\varepsilon|_{H^2}^2 + |\nabla \mathbf{d}_R^\varepsilon|_{H^2}^2) + |\partial_t \mathbf{d}_R^\varepsilon|_{H^2} ] |\mathbf{d}_R^\varepsilon|_{H^2}. \end{aligned} \quad (3.33)$$

By plugging the inequalities (3.32) and (3.33) into the equality (3.31), we get the following estimate:

$$\begin{aligned}
& \frac{1}{2} \frac{d}{dt} \left[ |\nabla^k \partial_t \mathbf{d}_R^\epsilon + \nabla^k \mathbf{d}_R^\epsilon|_{L^2}^2 + \left(\frac{1}{\epsilon} - 1\right) |\nabla^k \mathbf{d}_R^\epsilon|_{L^2}^2 \right. \\
& \quad \left. - |\partial_t \nabla^k \mathbf{d}_R^\epsilon|_{L^2}^2 \right] - |\partial_t \nabla^k \mathbf{d}_R^\epsilon|_{L^2}^2 + \frac{1}{\epsilon} |\nabla^{k+1} \mathbf{d}_R^\epsilon|_{L^2}^2 \\
& \leq C_k \left\{ (1 + |\mathbf{d}_R^\epsilon|_{H^2}^2 + |\nabla \mathbf{d}_R^\epsilon|_{H^2}^2 + \left|\frac{\mathbf{d}_R^\epsilon}{\sqrt{\epsilon}}\right|_{H^2} + \left|\frac{\nabla \mathbf{d}_R^\epsilon}{\sqrt{\epsilon}}\right|_{H^2}) \left|\frac{\mathbf{d}_R^\epsilon}{\sqrt{\epsilon}}\right|_{H^2} \right. \\
& \quad + [ (|\mathbf{d}_R^\epsilon|_{H^2} + |\nabla \mathbf{d}_R^\epsilon|_{H^2}) (1 + |\nabla \mathbf{d}_R^\epsilon|_{H^2}^2 + \epsilon |\partial_t \mathbf{d}_R^\epsilon|_{H^2}^2) \\
& \quad \left. + \sqrt{\epsilon} (1 + |\partial_t \mathbf{d}_R^\epsilon|_{H^2}^2 + |\mathbf{d}_R^\epsilon|_{H^2}^2 + |\nabla \mathbf{d}_R^\epsilon|_{H^2}^2) + |\partial_t \mathbf{d}_R^\epsilon|_{H^2} \right] |\mathbf{d}_R^\epsilon|_{H^2} \left. \right\}. \tag{3.34}
\end{aligned}$$

Multiplying the inequality (3.34) by  $\frac{1}{2}$  and adding it to the inequality (3.30), we obtain the higher order estimate:

$$\begin{aligned}
& \frac{1}{4} \frac{d}{dt} \left[ |\partial_t \nabla^k \mathbf{d}_R^\epsilon|_{L^2}^2 + \left(\frac{1}{\epsilon} - 1\right) |\nabla^k \mathbf{d}_R^\epsilon|_{L^2}^2 + \frac{2}{\epsilon} |\nabla^{k+1} \mathbf{d}_R^\epsilon|_{L^2}^2 + |\nabla^k \partial_t \mathbf{d}_R^\epsilon + \nabla^k \mathbf{d}_R^\epsilon|_{L^2}^2 \right] \\
& \quad + \left(\frac{1}{\epsilon} - \frac{1}{2}\right) |\partial_t \nabla^k \mathbf{d}_R^\epsilon|_{L^2}^2 + \frac{1}{2\epsilon} |\nabla^{k+1} \mathbf{d}_R^\epsilon|_{L^2}^2 \\
& \leq \frac{3}{2} C_k \left\{ (1 + |\mathbf{d}_R^\epsilon|_{H^2}^2 + |\nabla \mathbf{d}_R^\epsilon|_{H^2}^2 + \left|\frac{\mathbf{d}_R^\epsilon}{\sqrt{\epsilon}}\right|_{H^2} + \left|\frac{\nabla \mathbf{d}_R^\epsilon}{\sqrt{\epsilon}}\right|_{H^2}) \left(\left|\frac{\mathbf{d}_R^\epsilon}{\sqrt{\epsilon}}\right|_{H^2} + \left|\frac{\partial_t \mathbf{d}_R^\epsilon}{\sqrt{\epsilon}}\right|_{H^2}\right) \right. \\
& \quad + [ (|\mathbf{d}_R^\epsilon|_{H^2} + |\nabla \mathbf{d}_R^\epsilon|_{H^2}) (1 + |\nabla \mathbf{d}_R^\epsilon|_{H^2}^2 + \epsilon |\partial_t \mathbf{d}_R^\epsilon|_{H^2}^2) + |\partial_t \mathbf{d}_R^\epsilon|_{H^2} \\
& \quad \left. + \sqrt{\epsilon} (1 + |\partial_t \mathbf{d}_R^\epsilon|_{H^2}^2 + |\mathbf{d}_R^\epsilon|_{H^2}^2 + |\nabla \mathbf{d}_R^\epsilon|_{H^2}^2) ] (|\mathbf{d}_R^\epsilon|_{H^2} + |\partial_t \mathbf{d}_R^\epsilon|_{H^2}) \left. \right\}. \tag{3.35}
\end{aligned}$$

Therefore, combining the  $L^2$ -estimate (3.19) and the  $k$ th-order estimate (3.35) for  $k = 1, 2$ , we obtain:

$$\begin{aligned}
& \frac{1}{4} \frac{d}{dt} \left[ |\partial_t \mathbf{d}_R^\epsilon|_{H^2}^2 + \left(\frac{1}{\epsilon} - 1\right) |\mathbf{d}_R^\epsilon|_{H^2}^2 + \frac{2}{\epsilon} |\nabla \mathbf{d}_R^\epsilon|_{H^2}^2 + |\partial_t \mathbf{d}_R^\epsilon + \mathbf{d}_R^\epsilon|_{H^2}^2 \right] \\
& \quad + \left(\frac{1}{\epsilon} - \frac{1}{2}\right) |\partial_t \mathbf{d}_R^\epsilon|_{H^2}^2 + \frac{1}{2\epsilon} |\nabla \mathbf{d}_R^\epsilon|_{H^2}^2 \\
& \leq \tilde{C}_k \left\{ (1 + |\mathbf{d}_R^\epsilon|_{H^2}^2 + |\nabla \mathbf{d}_R^\epsilon|_{H^2}^2 + \left|\frac{\mathbf{d}_R^\epsilon}{\sqrt{\epsilon}}\right|_{H^2} + \left|\frac{\nabla \mathbf{d}_R^\epsilon}{\sqrt{\epsilon}}\right|_{H^2}) \left(\left|\frac{\mathbf{d}_R^\epsilon}{\sqrt{\epsilon}}\right|_{H^2} + \left|\frac{\partial_t \mathbf{d}_R^\epsilon}{\sqrt{\epsilon}}\right|_{H^2}\right) \right. \\
& \quad + [ (|\mathbf{d}_R^\epsilon|_{H^2} + |\nabla \mathbf{d}_R^\epsilon|_{H^2}) (1 + |\nabla \mathbf{d}_R^\epsilon|_{H^2}^2 + \epsilon |\partial_t \mathbf{d}_R^\epsilon|_{H^2}^2) + |\partial_t \mathbf{d}_R^\epsilon|_{H^2} \\
& \quad \left. + \sqrt{\epsilon} (1 + |\partial_t \mathbf{d}_R^\epsilon|_{H^2}^2 + |\mathbf{d}_R^\epsilon|_{H^2}^2 + |\nabla \mathbf{d}_R^\epsilon|_{H^2}^2) ] (|\mathbf{d}_R^\epsilon|_{H^2} + |\partial_t \mathbf{d}_R^\epsilon|_{H^2}) \left. \right\}, \tag{3.36}
\end{aligned}$$

where the positive constant  $\tilde{C}_k$  is

$$\tilde{C}_k = C(1 + |\partial_{tt} \mathbf{d}_0|_{L^\infty(0,T;H^2)} + |\partial_t \mathbf{d}_0|_{L^\infty(0,T;H^4)}^3 + |\nabla \mathbf{d}_0|_{L^\infty(0,T;H^4)}^3 + |\mathbf{D}(x)|_{H^5}^3) > 0$$

and  $C > 0$  is a computable constant. Then, by the definition of the energy functionals  $E_\epsilon(t)$  and  $F_\epsilon(t)$ , and the condition  $0 < \epsilon < \frac{1}{2}$ , the  $H^2$ -estimate (3.36) implies that

$$\frac{d}{dt} E_\epsilon(t) + 4F_\epsilon(t) \leq C' \left\{ E_\epsilon^{\frac{1}{2}}(t) + E_\epsilon(t) + \epsilon^{\frac{1}{2}} E_\epsilon^{\frac{3}{2}}(t) + \epsilon^{\frac{3}{2}} E_\epsilon^2(t) + [1 + E_\epsilon^{\frac{1}{2}}(t) + \epsilon E_\epsilon(t)] F_\epsilon^{\frac{1}{2}}(t) \right\},$$

where  $C' > 0$ , which immediately implies the claimed inequality (3.1) by using Young's inequality. Consequently, the proof of Lemma 3.1 is completed.  $\square$

#### 4. THE PROOF OF THEOREM 1.1

In this section we will provide the proof of Theorem 1.1, by using the uniform energy bounds (3.1) in Section 3. Before doing this, we note that for any fixed inertia constant  $\epsilon > 0$  the well-posedness of the remainder system (1.15)-(1.16) can be stated as follows:

**Proposition 4.1.** *Given  $\mathbf{d}^{in} : \mathbb{R}^3 \rightarrow \mathbb{S}^2$  and  $\tilde{\mathbf{d}}^{in} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  satisfying  $\nabla \mathbf{d}^{in} \in H^6$ ,  $\tilde{\mathbf{d}}^{in} \in H^4$  with  $\mathbf{d}^{in} \cdot \tilde{\mathbf{d}}^{in} \equiv 0$ , we define  $\mathbf{D}(x) \equiv \tilde{\mathbf{d}}^{in}(x) - \Delta \mathbf{d}^{in}(x) - |\nabla \mathbf{d}^{in}(x)|^2 \mathbf{d}^{in}(x)$  and denote  $M = |\mathbf{D}|_{H^2}^2 + 2|\nabla \mathbf{D}|_{H^2}^2 < \infty$ .*

Then, for any fixed  $\varepsilon \in (0, \frac{1}{2})$ , there exists a time  $T^\varepsilon = \min\{T, \frac{1}{C} \ln(\frac{1+\varepsilon M}{\varepsilon(1+M)})\} > 0$ , where  $T, C > 0$  are provided in Proposition 1.1, Lemma 3.1, respectively, such that the remainder equation (1.15) with the initial conditions (1.16) admits a unique solution  $\mathbf{d}_R^\varepsilon \in C([0, T^\varepsilon]; H^3)$  and  $\partial_t \mathbf{d}_R^\varepsilon \in C([0, T^\varepsilon]; H^2)$ . Moreover, the solution  $\mathbf{d}_R^\varepsilon$  satisfies the inequality

$$|\partial_t \mathbf{d}_R^\varepsilon(t)|_{H^2}^2 + \frac{1}{\varepsilon} |\mathbf{d}_R^\varepsilon(t)|_{H^3}^2 \leq \frac{2Me^{Ct}}{1+\varepsilon M - \varepsilon(1+M)e^{Ct}} \quad (4.1)$$

for all  $t \in [0, T^\varepsilon]$ .

*Proof.* We employ a mollifier argument to prove this proposition. For any fixed  $\varepsilon > 0$  we can directly construct a system approximating (1.15)-(1.16) as follows:

$$\begin{cases} \varepsilon \partial_t \mathbf{w}_\eta^\varepsilon = -\mathcal{J}_\eta \mathbf{w}_\eta^\varepsilon + \mathcal{J}_\eta \Delta \mathbf{d}_{R,\eta}^\varepsilon + \varepsilon \mathcal{J}_\eta \mathcal{S}(\mathcal{J}_\eta \mathbf{d}_{R,\eta}^\varepsilon) + \varepsilon \mathcal{J}_\eta \mathcal{R}(\mathcal{J}_\eta \mathbf{d}_{R,\eta}^\varepsilon), \\ \partial_t \mathbf{d}_{R,\eta}^\varepsilon = \mathbf{w}_\eta^\varepsilon, \\ \mathbf{d}_{R,\eta}^\varepsilon|_{t=0} = \sqrt{\varepsilon} \mathcal{J}_\eta \mathbf{D}(x), \quad \mathbf{w}_\eta^\varepsilon|_{t=0} = 0, \end{cases} \quad (4.2)$$

where the mollifier operator  $\mathcal{J}_\eta$  is defined as

$$\mathcal{J}_\eta f = \mathcal{F}^{-1}(\mathbf{1}_{|\xi| \leq \frac{1}{\eta}} \mathcal{F}(f)(\xi)),$$

where the symbol  $\mathcal{F}$  denotes the standard Fourier transform operator and  $\mathcal{F}^{-1}$  is the inverse Fourier transform operator. By ODE theory in Hilbert spaces one can prove the existence and uniqueness of the approximate system (4.2) on the maximal time interval  $[0, T_\eta^\varepsilon]$ . Then by the fact  $\mathcal{J}_\eta^2 = \mathcal{J}_\eta$  and the uniqueness of (4.2) we know that  $\mathcal{J}_\eta \mathbf{d}_{R,\eta}^\varepsilon = \mathbf{d}_{R,\eta}^\varepsilon$  and  $\mathcal{J}_\eta \mathbf{w}_\eta^\varepsilon = \mathbf{w}_\eta^\varepsilon$ . Thus by the analogous energy estimate shown in Lemma 3.1 applied to the approximate system (4.2), one can obtain the following energy inequality for  $\mathbf{d}_{R,\eta}^\varepsilon$  and  $\mathbf{w}_\eta^\varepsilon$

$$\frac{d}{dt} E_{\varepsilon,\eta}(t) + 3F_{\varepsilon,\eta}(t) \leq C[1 + E_{\varepsilon,\eta}(t)][1 + \varepsilon E_{\varepsilon,\eta}(t)] \quad (4.3)$$

for all  $t \in [0, T_\eta^\varepsilon]$ , where the positive constant  $C > 0$  is independent of  $\varepsilon$  and  $\eta$ , and the energy functionals  $E_{\varepsilon,\eta}(t), F_{\varepsilon,\eta}(t)$  are of the same forms as  $E_\varepsilon(t), F_\varepsilon(t)$  defined in Section 3 (replacing  $\mathbf{d}_R^\varepsilon$  by  $\mathbf{d}_{R,\eta}^\varepsilon$ ), respectively.

Since  $\mathbf{d}_{R,\eta}^\varepsilon$  satisfies the initial conditions  $\partial_t \mathbf{d}_{R,\eta}^\varepsilon(0, x) = 0$  and  $\mathbf{d}_{R,\eta}^\varepsilon(0, x) = \sqrt{\varepsilon} \mathcal{J}_\eta \mathbf{D}(x)$ , we know that for  $\varepsilon \in (0, \frac{1}{2})$

$$\begin{aligned} E_{\varepsilon,\eta}(0) &= |\partial_t \mathbf{d}_{R,\eta}^\varepsilon(0, \cdot)|_{H^2}^2 + (\frac{1}{\varepsilon} - 1) |\mathbf{d}_{R,\eta}^\varepsilon(0, \cdot)|_{H^2}^2 \\ &\quad + \frac{2}{\varepsilon} |\nabla \mathbf{d}_{R,\eta}^\varepsilon(0, \cdot)|_{H^2}^2 + |\partial_t \mathbf{d}_{R,\eta}^\varepsilon(0, \cdot) + \mathbf{d}_{R,\eta}^\varepsilon(0, \cdot)|_{H^2}^2 \\ &= (\frac{1}{\varepsilon} - 1) |\sqrt{\varepsilon} \mathcal{J}_\eta \mathbf{D}|_{H^2}^2 + \frac{2}{\varepsilon} |\sqrt{\varepsilon} \nabla \mathcal{J}_\eta \mathbf{D}|_{H^2}^2 + |\sqrt{\varepsilon} \mathcal{J}_\eta \mathbf{D}|_{H^2}^2 \\ &\leq (1 - \varepsilon) |\mathbf{D}|_{H^2}^2 + 2 |\nabla \mathbf{D}|_{H^2}^2 + \varepsilon |\mathbf{D}|_{H^2}^2 \\ &= |\mathbf{D}|_{H^2}^2 + 2 |\nabla \mathbf{D}|_{H^2}^2 = M < \infty. \end{aligned} \quad (4.4)$$

Then, one can solve the ODE inequality (4.3) with the initial condition (4.4), obtaining that

$$\frac{1 + E_{\varepsilon,\eta}(t)}{1 + \varepsilon E_{\varepsilon,\eta}(t)} \leq \frac{1 + E_{\varepsilon,\eta}(0)}{1 + \varepsilon E_{\varepsilon,\eta}(0)} e^{C(1-\varepsilon)t} \leq \frac{1 + M}{1 + \varepsilon M} e^{Ct}$$

holds for all  $t \in [0, T_\eta^\varepsilon]$ . Consequently, for all  $t \in [0, \min\{T_\eta^\varepsilon, T, \frac{1}{C} \ln(\frac{1+\varepsilon M}{\varepsilon(1+M)})\}]$  we know that

$$E_{\varepsilon,\eta}(t) \leq \frac{(1 + M)e^{Ct}}{1 + \varepsilon M - \varepsilon(1 + M)e^{Ct}}. \quad (4.5)$$

Notice that the continuity of  $E_{\varepsilon,\eta}(t)$  and the maximality of  $T_\eta^\varepsilon > 0$  imply that

$$T_\eta^\varepsilon \geq \frac{1}{C} \ln\left(\frac{1+\varepsilon M}{\varepsilon(1+M)}\right) > 0.$$

Hence the inequality (4.5) holds for all  $t \in [0, T^\varepsilon]$  uniformly in  $\eta > 0$ , where

$$T^\varepsilon = \min\left\{T, \frac{1}{C} \ln\left(\frac{1+\varepsilon M}{\varepsilon(1+M)}\right)\right\} > 0.$$

Finally, we can finish the proof of this proposition by standard compactness methods and taking the limit as  $\eta \rightarrow 0$ . The uniqueness issue can be reduced to the uniqueness of the damped wave map system (1.8) which can be obtained by methods analogous to those in the book of Shatah and Struwe [15]. For convenience, we omit the details of the proof.  $\square$

**The Proof of Theorem 1.1.** Now, based on the energy estimate (4.1) in Proposition 4.1, we verify Theorem 1.1. We observe that the function

$$f(\varepsilon) := \frac{1}{C} \ln \left( \frac{1 + \varepsilon M}{\varepsilon(1 + M)} \right)$$

is strictly decreasing in  $\varepsilon \in (0, \frac{1}{2})$  and  $\lim_{\varepsilon \searrow 0} f(\varepsilon) = +\infty$ . Consequently, we can choose

$$\varepsilon_0 = \min \left\{ \frac{1}{2}, \frac{1}{(1+M)e^{CT}-M} \right\} \in (0, \frac{1}{2})$$

such that for any  $\varepsilon \in (0, \varepsilon_0)$

$$f(\varepsilon) = \frac{1}{C} \ln \left( \frac{1 + \varepsilon M}{\varepsilon(1 + M)} \right) > T.$$

As a result, for the number  $T^\varepsilon$  determined in Proposition 4.1, we have that  $T^\varepsilon \equiv T$  for all  $\varepsilon \in (0, \varepsilon_0)$ . Then, the inequality (4.1) in Proposition 4.1 implies that

$$|\partial_t \mathbf{d}_R^\varepsilon|_{L^\infty(0,T;H^2)}^2 + \frac{1}{\varepsilon} |\mathbf{d}_R^\varepsilon|_{L^\infty(0,T;H^3)}^2 \leq \frac{2Me^{CT}}{1+\varepsilon_0 M - \varepsilon_0(1+M)e^{CT}} := C_0 < \infty,$$

and the proof of Theorem 1.1 is completed.  $\square$

#### ACKNOWLEDGEMENT

The project of this paper was initialized when Ning Jiang visited Arghir Zarnescu at Basque Center for Applied Mathematics (BCAM) in March 2017. They appreciate the hospitality of BCAM. The activity of Ning Jiang on this work was supported by Chinese NSF grant 11471181.

The activity of Arghir Zarnescu on this work was partially supported by a grant of the Romanian National Authority for Scientific Research and Innovation, CNCS-UEFISCDI, project number PN-II-RU-TE-2014-4-0657; by the Project of the Spanish Ministry of Economy and Competitiveness with reference MTM2013-40824-P; by the Basque Government through the BERC 20142017 program; and by the Spanish Ministry of Economy and Competitiveness MINECO: BCAM Severo Ochoa accreditation SEV-2013-0323.

#### REFERENCES

- [1] R. Caffisch, The fluid dynamic limit of the nonlinear Boltzmann equation. *Comm. Pure Appl. Math.* **33** (1980), no. 5, 651-666.
- [2] R. Chill, and A. Haraux, An optimal estimate for the time singular limit of an abstract wave equation. *Funkcial. Ekvac.* **47** (2004), no. 2, 277-290.
- [3] F. De Anna and A. Zarnescu, Global well-posedness and twist-wave solutions for the inertial Qian-Sheng model of liquid crystals. *J. Differential Equations* **264** (2018), no. 2, 1080-1118.
- [4] J. Eells and J. H. Sampson, Harmonic mappings of Riemannian manifolds, *American Journal of Mathematics*, **86** (1964), no. 1, 109-160.
- [5] J. L. Ericksen, Conservation laws for liquid crystals. *Trans. Soc. Rheology* **5** 1961, 23-34.
- [6] J. L. Ericksen, Continuum theory of nematic liquid crystals. *Res. Mechanica* **21**, (1987), 381-392.
- [7] J. L. Ericksen, Liquid crystals with variable degree of orientation. *Arch. Rational Mech. Anal.* **113** (1990), no. 2, 97-120.
- [8] E. Feireisl, E. Rocca, G. Schimperna and A. Zarnescu, On a hyperbolic system arising in liquid crystals modeling. arXiv:1610.07828 (2016).
- [9] N. Jiang and Y. L. Luo, On well-posedness of Ericksen-Leslie's parabolic-hyperbolic liquid crystal model. arXiv:1709.06370 [math.AP].

- [10] F. M. Leslie, Some constitutive equations for liquid crystals. *Arch. Rational Mech. Anal.* **28** (1968), no. 4, 265-283.
- [11] F. M. Leslie, Theory of flow phenomena in liquid crystals. *The Theory of Liquid Crystals* Vol. 4, pp. 1-81. Academic Press, London-New York, 1979.
- [12] F.-H. Lin and C. Liu, Static and dynamic theories of liquid crystal. *J. Partial Differential Equations*, **14** (2001), no. 4, 289-330.
- [13] F.-H. Lin and C.-Y. Wang, *The analysis of harmonic maps and their heat flows*. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2008.
- [14] P.-L. Lions, *Mathematical Topics in Fluid Mechanics, Vol. 1: Incompressible Models*. Clarendon Press. Oxford, 1996.
- [15] J. Shatah and M. Struwe, *Geometric Wave Equations*. Courant Institute of Mathematical Sciences, New York University, New York and American Mathematical Society, Providence, Rhode Island, 2000.

(Ning Jiang)

SCHOOL OF MATHEMATICS AND STATISTICS, WUHAN UNIVERSITY, WUHAN, 430072, P. R. CHINA  
*E-mail address:* `njiang@whu.edu.cn`

(Yi-Long Luo)

SCHOOL OF MATHEMATICS AND STATISTICS, WUHAN UNIVERSITY, WUHAN, 430072, P. R. CHINA  
*E-mail address:* `yl-luo@whu.edu.cn`

(Shaojun Tang)

SCHOOL OF MATHEMATICS AND STATISTICS, WUHAN UNIVERSITY, WUHAN, 430072, P. R. CHINA  
*E-mail address:* `shaojun.tang@whu.edu.cn`

(Arghir Zarnescu)

BCAM, BASQUE CENTER FOR APPLIED MATHEMATICS, MAZARREDO 14, 48009 BILBAO, SPAIN  
*E-mail address:* `azarnescu@bcamath.org`