Multiscale approach for variational problem joint diffeomorphic image registration and intensity correction: theory and application*

Peng Chen†, Ke Chen‡, Huan Han§, and Daoping Zhang¶

Abstract. Image registration matches the features of two images, by minimizing the intensity difference between them, so that useful and complementary information can be extracted from the mapping. However, in real life problems, the images may be affected by the imaging environment, such as varying illumination and noise during the process of imaging acquisition. This may lead to the local intensity distortion, which makes it meaningless to minimize the intensity difference in traditional registration framework. To address this problem, we propose a variational model for joint image registration and intensity correction. Based on this model, the related greedy matching problem is solved by introducing a multiscale approach for joint image registration and intensity correction. An alternating direction method (ADM) is proposed to solve each multiscale step, and the convergence of the ADM method is proved. For the numerical implementation, a coarse-to-fine strategy is proposed to accelerate the numerical algorithm, and the convergence of the proposed coarse-to-fine strategy is proved. Several numerical tests are also performed to validate the efficiency of the proposed algorithm.

Key words. image registration, multiscale, diffeomorphism, multigrid, multi-resolution, from coarse to fine

AMS subject classifications. 68U10, 94A08, 65K10, 65M12

1. Introduction.

Image registration is to match the features of two images by keeping one image (target image) unchanged and deforming the other image (floating image). By comparing the deformed image with the target image, one can extract useful information from intensity differences. This is a fundamental process for image fusion and medical analysis. For an overview of image registration and related joint problems, one can refer to [1,3,5–7,10,14,15,23,24,32] for details.

Without loss of generality, in this paper, we mainly focus on 2D image registration, which is stated in the following way. Given two images $T(\mathbf{x}), D(\mathbf{x}) : \mathbf{x} \in \Omega \to \mathbb{R}$ and some bounded domain $\Omega \subset \mathbb{R}^2$, the goal of image registration is to find a mapping $\boldsymbol{\varphi}(\mathbf{x}) : \mathbf{x} \in \Omega \to \Omega$ such that $T \circ \boldsymbol{\varphi}(\cdot)$ looks like $D(\cdot)$ as much as possible. For each $\mathbf{x} \in \Omega$, $\boldsymbol{\varphi}(\mathbf{x})$ can be divided into the

*Submitted to the editors DATE.

Funding: P Chen was supported in part by Natural Science Foundation of Hubei Province of China (No. 2021CFB473); K Chen was supported in part by EPSRC(No. EP/N014499/1); H Han was supported in part by National Natural Science Foundation of China (No.11901443) and Natural Science Foundation of Hubei Province of China (No. 2022CFB379); D Zhang was supported in part by Natural Science Foundation of China (No. 12201320) and the Fundamental Research Funds for the Central Universities, Nankai University (No. 63221039).

†College of Science, China Three Gorges University, Yichang 443002, People's Republic of China (pengchen729@sina.com).

‡Liverpool Centre for Mathematics in Healthcare, Centre for Mathematical Imaging Techniques, and Department of Mathematical Sciences, The University of Liverpool, Liverpool L69 7ZL, United Kingdom (k.chen@liv.ac.uk).

§ Corresponding author. Department of Mathematics, Wuhan University of Technology, Wuhan 430070, People's Republic of China (hanhuan11@whut.edu.cn).

¶Department of Mathematical Sciences and LPMC, Nankai University, Tianjin 300071, People's Republic of China (daopingzhang@nankai.edu.cn).*All the authors contributed equally to this work.

identity part $\mathbf{x}$ and the displacement $\mathbf{u}(\mathbf{x})$, i.e., $\boldsymbol{\varphi}(\mathbf{x}) \triangleq \mathbf{x} + \mathbf{u}(\mathbf{x})$. Based on this assumption, the mono-modality image registration problem is formulated as follows:

$$\min_{\mathbf{u}\in\mathcal{A}} \ \lambda S(\mathbf{u}) + \mu R(\mathbf{u}), \tag{1.1}$$

where $\lambda, \mu > 0$, $\mathcal{A}$ is some proper set, $S(\mathbf{u}) = \int_\Omega [T(\mathbf{x}+\mathbf{u}(\mathbf{x})) - D(\mathbf{x})]^2 d\mathbf{x}$, $R(\mathbf{u})$ is a regularization to produce plausible solutions. For multi-modality diffeomorphic image registration [4,26], $S(\mathbf{u}) = \int_\Omega (f_1(T)(\mathbf{x}+\mathbf{u}(\mathbf{x})) - f_2(D)(\mathbf{x}))^2 d\mathbf{x}$, f_1, f_2 are two gray transform functions. The problem considered in this paper lies between the two types of registration problems because the given images appear in multi-modalities but the modelling must be done in mono-modality.

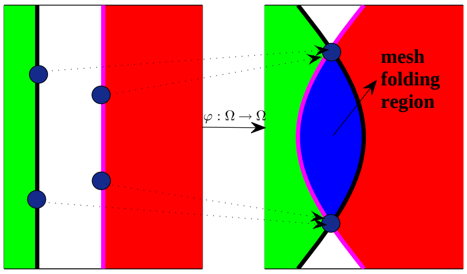
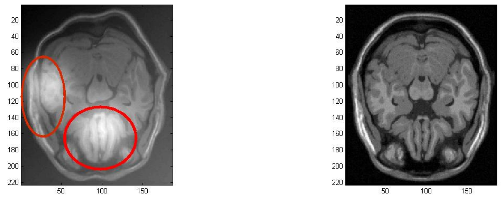


Figure 1. *Physical mesh folding caused by the deformation* $\boldsymbol{\varphi}$

Although image registration has achieved enormous success, it is still a challenging task. There are mainly two difficulties: (I) physical mesh folding; (II) illposeness of greedy matching. As shown in Fig 1, physical mesh folding is a phenomenon that points from different objects are mixed together after transformation. We can find that the essential reason for mesh folding is the non-bijection of the deformation mapping. Therefore, to eliminate mesh folding, it is necessary to guarantee that the Jacobian determinant of the deformation is larger than 0 for each pixel [13, 18, 19]. This is so called 'orientation-preserving registration'. Under this framework, several diffeomorphic image registration models have been proposed [8, 17–19, 21, 25, 28, 29, 34, 36, 37]. In the pioneering work [8], Lui introduced the quasi-conformal theory to control the mesh folding. Following this work, several models are proposed to improve the quasi-conformal model. In [39], Zhang and Chen proposed a diffeomorphic image registration model by restricting the deformation $\boldsymbol{\varphi}$ into a set which ensures $\det(\nabla\boldsymbol{\varphi}(\mathbf{x})) > 0$ for each $\mathbf{x} \in \Omega$. As a supplement, Zhang and Chen [35] introduced a diffeomorphic image registration model by controlling the modulus of Beltrami coefficient smaller than 1. Han, Wang and Zhang also gave a series of 2D/3D diffeomorphic image registration models and algorithms by restricting $\mathbf{u}$ into the 2D/3D conformal set [17–19, 21].

However, these above mentioned works are all based on the assumption that no intensity distortion (i.e., illumination and noise) occurs during the process of imaging acquisition. For example, in Fig 2, locally varying illumination occurs inside the region of the floating image $T(\cdot)$ and no illumination in the target image $D(\cdot)$. This leads to the intensity distortion in



(a) $T(\cdot)$ (b) $D(\cdot)$

Figure 2. *Local varying illumination in MRI image pair*

these two regions. In this case, models such as (1.1) by treating T, D as mono-modal images fail to register the two images. It is meaningless for latter applications, such as image fusion and image analysis, even if the sketch of the two objects are exactly matched as a multi-modal problem. Therefore, it is necessary to introduce some intensity correction steps during or after image registration. For this purpose, some variational models joint image registration and intensity correction are proposed [27,31]. By introducing the additive and multiplicative bias field for intensity correction simultaneously, the relationship between the true image $D^*(\mathbf{x}) = D_c(\mathbf{x})$ and the target image $D(\mathbf{x})$ is formulated as

$$D(\mathbf{x}) = m(\mathbf{x})D^*(\mathbf{x}) + s(\mathbf{x}),$$

where $s(\mathbf{x}) : \mathbf{x} \in \Omega \to \mathbb{R}$ and $m(\mathbf{x}) : \mathbf{x} \in \Omega \to \mathbb{R}^+$ are additive and multiplicative bias fields, respectively. Then Theljani and Chen [31] proposed a joint model for image registration and intensity correction:

$$\min_{\mathbf{u},m,s} \ \lambda S_c(\mathbf{u}, m, s) + \mu R(\mathbf{u}, m, s), \tag{1.2}$$

where $S_c(\mathbf{u}, m, s) = \int_\Omega (m(\mathbf{x})D(\mathbf{x}) + s(\mathbf{x}) - T(\mathbf{x} + \mathbf{u}(\mathbf{x})))^2 d\mathbf{x}$ and $R(\mathbf{u}, m, s)$ is a regularization on $\mathbf{u}, m$ and s. Viewing the solution of the variational model (1.2) as Nash game equilibrium, a novel numerical algorithm for joint image registration and intensity correction is also devised in [31]. However, the above mentioned mesh folding is not constrained (difficulty I) and the game solution is a 'perturbed solution', not a minimizer of the original variational functional. The other works on joint image registration and intensity correction can be found in [11,12].

The ultimate goal for joint image registration and intensity correction is to find the minimizer of the cost functional $S_c(\mathbf{u}, m, s)$. However, (1.2) aims to find the minimizer of $\lambda S_c(\mathbf{u}, m, s) + \mu R(\mathbf{u}, m, s)$. This raises a question of whether or not one can find the global minimizer of $S_c(\mathbf{u}, m, s)$ on some proper space without any prior estimate for $\mathbf{u}, m, s$? This is so called 'greedy matching'. Concerning this problem (difficulty II), Han, Wang and Zhang [20,21] gave an answer in case of $m(\mathbf{x}) \equiv 1$, $s(\mathbf{x}) \equiv 0$ and T, D having no bias (i.e., image registration without intensity correction) by introducing a multiscale approach and proved the equivalence between the proposed multiscale approach and 'greedy matching' with some suitable parameters. For the general cases, to the best of our knowledge, there seems to have no results. Motivated by [20,21], we aim to extend the work [20] to the case that m, s belong to some specific Banach spaces. For this purpose, we propose the following variational model for joint diffeomorphic image registration and intensity correction:

$$\min_{\mathbf{u}\in\mathcal{A}(\Omega)\setminus\mathcal{B}_\varepsilon(\Omega), m\in\mathcal{C}_\Omega, s\in SV_0(\Omega)} J(\mathbf{u},m,s) := \lambda S_{lc}(\mathbf{u},m,s) + \mu R(\mathbf{u},m,s), \tag{1.3}$$

where $S_{lc}(\mathbf{u},m,s) = \int_\Omega (m(\mathbf{x})+\ln D(\mathbf{x})-\ln(T(\mathbf{x}+\mathbf{u}(\mathbf{x}))-s(\mathbf{x})))^2 d\mathbf{x}$ and $R(\mathbf{u},m,s) = R_1(\mathbf{u})+R_2(m)+R_3(s)$, $R_1(\mathbf{u}) = \int_\Omega |\nabla^\alpha \mathbf{u}(\mathbf{x})|^2 d\mathbf{x}$, $R_2(m) = \int_\Omega |\nabla m(\mathbf{x})| d\mathbf{x}$, $R_3(s) = \int_\Omega |\nabla s(\mathbf{x})| d\mathbf{x}$. Note that here and in what follows, we assume that two images T, D map Ω onto the interval $[\kappa, \overline{M}] \subset \mathbb{R}^+$ for some $\overline{M} > \kappa > 0$. In addition, for the purpose of eliminating mesh folding, $\mathbf{u}$ is constrained into the set $\mathcal{A}(\Omega)\setminus\mathcal{B}_\varepsilon(\Omega)$, where $\mathcal{A}(\Omega)$ and $\mathcal{B}_\varepsilon(\Omega)$ are defined by

$$\mathcal{A}(\Omega) = \left\{ \mathbf{u} = (u_1,u_2)^T \in [H_0^\alpha(\Omega)]^2 : \frac{\partial u_1}{\partial x_1} = \frac{\partial u_2}{\partial x_2}, \frac{\partial u_1}{\partial x_2} = -\frac{\partial u_2}{\partial x_1} \right\}, \tag{1.4}$$

and

$$\mathcal{B}_\varepsilon(\Omega) = \{\mathbf{u} = (u_1,u_2)^T \in \mathcal{A}(\Omega) : \det(\nabla(\mathbf{x}+\mathbf{u}(\mathbf{x}))) < \varepsilon\}, \tag{1.5}$$

for $\alpha > 2$, small $\varepsilon > 0$ and $H_0^\alpha(\Omega)$ is fractional-order Sobelev space [16]. To control the intensity bias in practice, the multiplicative bias field m is constrained into the set

$$\mathcal{C}_\Omega = \{m \in BV_0(\Omega) : K_1 \le m \le K_2\}, \tag{1.6}$$

for some given K_1, K_2, and the additive bias field s is constrained into the set

$$SV_0(\Omega) = \{s \in BV_0(\Omega) : s(\mathbf{x}) < \kappa - \kappa_0 \text{ for } \forall \mathbf{x} \in \Omega\},$$

for some $\kappa > \kappa_0 > 0$ to ensure that $\ln(T(\mathbf{x}+\mathbf{u}(\mathbf{x})) - s(\mathbf{x}))$ is well-defined. Here, $BV_0(\Omega) = \{m \in BV(\Omega) : m(\mathbf{x})|_{\mathbf{x}\in\partial\Omega} = 0\}$ and the space $BV(\Omega)$ is as defined in [30].

Remark 1.1. By letting $m(\mathbf{x}) \equiv 1$, $s(\mathbf{x}) \equiv 0$, one can notice that the model (1.2) is reduced to (1.1), which means that the model (1.2) is much more general than the model (1.1).

Remark 1.2. By setting $m(\mathbf{x}) = \ln \bar{m}(\mathbf{x})$ for some positive function $\bar{m}(\mathbf{x})$, then $S_{lc}(\mathbf{u},m,s)$ from (1.2) becomes

$$S_{lc}(\mathbf{u},m,s) = \int_\Omega \left(\ln \frac{T(\mathbf{x}+\mathbf{u}(\mathbf{x})) - s(\mathbf{x})}{\bar{m}(\mathbf{x})D(\mathbf{x})} \right)^2 d\mathbf{x}. \tag{1.7}$$

That is, the problem from (1.3) $\arg\inf S_{lc}(\mathbf{u},m,s)$ is equivalent to $\arg\inf S_c(\mathbf{u},m,s)$ in (1.2). By using (1.7) as the data fidelity for (1.3), it has two advantages: (i) transforming the multiplicative bias field into additive bias field; (ii) eliminating the positive constraint $m(\mathbf{x}) > 0$ in the definition of $S_c(\mathbf{u},m,s)$. In addition, by using $S_{lc}(\mathbf{u},m,s)$ as the fidelity data, the final matched image for (1.3) should be calculated as $T_c(\cdot) = \frac{T(\cdot+\mathbf{u}(\cdot))-s(\cdot)}{e^{m(\cdot)}}$.

Based on the model (1.3), in this paper, we propose a multiscale approach for joint image registration and intensity correction, which aims to find the global minimizer of $S_{lc}(\mathbf{u},m,s)$ on $\mathcal{A}\times\mathcal{C}_\Omega\times SV_0(\Omega)$ for some given K_1, K_2 (see Section 2 for details). That is, $\inf_{(\mathbf{u},m,s)\in\mathcal{A}\times\mathcal{C}_\Omega\times SV_0(\Omega)} S_{lc}(\mathbf{u},m,s)$. This is so called 'greedy problem' for joint diffeomorphic image registration and intensity correction, which searches for the global minimizer of the similarity $S_{lc}(\mathbf{u},m,s)$ by placing the regularization into the constraint set $\mathcal{A}\times\mathcal{C}_\Omega\times SV_0(\Omega)$. The main contributions of the proposed multiscale approach contain the following three aspects:

- Propose a novel joint model for image registration and intensity correction;
- Address the greedy problem for joint image registration and intensity correction;
- Eliminate the intensity inhomogeneity by removing the bias.

The rest of this paper is organized as follows. In Section 2, we propose a multiscale approach for (1.3) to address the 'greedy problem'. In Section 3, an ADM method to solve the joint model for each scale is discussed and the convergence is also proved under some suitable assumptions. Then in Section 4, we propose a coarse-to-fine strategy for the multiscale approach to further accelerate the algorithm. In Section 5, some applications of the proposed multiscale approach are performed. Finally, we conclude our work and outline some problems for future research in Section 6.

2. Multiscale approach based on the model (1.3) and related greedy problem. Mesh folding may occur in large deformation registration. To control the mesh folding in large deformation registration, one can decompose the large deformation $\tilde{\boldsymbol{\varphi}}_n$ into the composition of several small deformations $\boldsymbol{\varphi}_i (i = 0, 1, 2, \cdots, n)$, where $\boldsymbol{\varphi}_i$ is the deformation induced by the joint model (1.3) under different scale parameters λ_i and ε_i. For example, by setting $\lambda_i = \lambda_0 \times a^i$, $\varepsilon_i = \frac{\varepsilon_0}{2^i}$ with $a > 1$, the parameter sequences $\{\lambda_n\}$ and $\{\varepsilon_n\}$ is initialized by some large number λ_0 and positive small number ε_0 (i.e, $\lambda_0 = 3000$, $\varepsilon_0 = 0.01$). In this way, the large diffeomorphism is achieved. Motivated by this idea, we propose the multiscale approach based on the model (1.3), to give an answer to the question of whether or not one can find the global minimizer of $S_{lc}(\mathbf{u}, m, s)$ on $\mathcal{L}(\Omega) = \mathcal{A} \times \mathcal{C}_\Omega \times SV_0(\Omega)$. The multiscale approach is divided into the following n steps:

Step 0. Searching for the solution of the following variational problem:

$$(\mathbf{u}_0, m_0, s_0) \in \underset{(\mathbf{u},m,s)\in\mathcal{L}_{\varepsilon_0}(\Omega)}{\arg\min} J_0(\mathbf{u}, m, s), \tag{2.1}$$

where $J_0(\mathbf{u}, m, s) = \lambda_0 \int_\Omega (m(\mathbf{x}) + \ln D(\mathbf{x}) - \ln(T(\mathbf{x}+\mathbf{u}(\mathbf{x})) - s(\mathbf{x})))^2 d\mathbf{x} + \mu R(\mathbf{u}, m, s)$, $\mathcal{L}_{\varepsilon_0}(\Omega) = (\mathcal{A}(\Omega) \setminus \mathcal{B}_{\varepsilon_0}(\Omega)) \times \mathcal{C}_\Omega \times SV_0(\Omega)$ and $\varepsilon_0 > 0$. Define $\tilde{\boldsymbol{\varphi}}_0(\mathbf{x}) = \boldsymbol{\varphi}_0(\mathbf{x}) = \mathbf{x} + \mathbf{u}_0(\mathbf{x})$.

Step 1. Searching for the solution of the following variational problem:

$$(\mathbf{u}_1, \delta m_1, \delta s_1) \in \underset{(\mathbf{u},m_0+m,s_0+s)\in\mathcal{L}_{\varepsilon_1}(\Omega)}{\arg\min} J_1(\mathbf{u}, m, s),$$

where $J_1(\mathbf{u}, m, s) = \lambda_1 \int_\Omega (m_0(\mathbf{x}) + m(\mathbf{x}) + \ln D(\mathbf{x}) - \ln(T \circ \tilde{\boldsymbol{\varphi}}_0(\mathbf{x} + \mathbf{u}(\mathbf{x})) - s_0(\mathbf{x}) - s(\mathbf{x})))^2 d\mathbf{x} + \mu R(\mathbf{u}, m, s)$, $\mathcal{L}_{\varepsilon_1}(\Omega) = (\mathcal{A}(\Omega) \setminus \mathcal{B}_{\varepsilon_1}(\Omega)) \times \mathcal{C}_\Omega \times SV_0(\Omega)$ and $\varepsilon_1 > 0$. Define $\boldsymbol{\varphi}_1(\mathbf{x}) = \mathbf{x} + \mathbf{u}_1(\mathbf{x})$, $\tilde{\boldsymbol{\varphi}}_1(\mathbf{x}) = \tilde{\boldsymbol{\varphi}}_0 \circ \boldsymbol{\varphi}_1(\mathbf{x})$, $m_1(\mathbf{x}) = m_0(\mathbf{x}) + \delta m_1(\mathbf{x})$ and $s_1(\mathbf{x}) = s_0(\mathbf{x}) + \delta s_1(\mathbf{x})$.

$$\vdots$$

Step n. By induction, for $n \geq 1$, searching for the solution of the following variational problem:

$$(\mathbf{u}_n, \delta m_n, \delta s_n) \in \underset{(\mathbf{u},m_{n-1}+m,s_{n-1}+s)\in\mathcal{L}_{\varepsilon_n}(\Omega)}{\arg\min} J_n(\mathbf{u}, m, s), \tag{2.2}$$

where $J_n(\mathbf{u}, m, s) = \lambda_n \int_\Omega (m_{n-1}(\mathbf{x}) + m(\mathbf{x}) + \ln D(\mathbf{x}) - \ln(T \circ \tilde{\boldsymbol{\varphi}}_{n-1}(\mathbf{x} + \mathbf{u}(\mathbf{x})) - s_{n-1}(\mathbf{x}) - s(\mathbf{x})))^2 d\mathbf{x} + \mu R(\mathbf{u}, m, s)$, $\mathcal{L}_{\varepsilon_n}(\Omega) = (\mathcal{A}(\Omega) \setminus \mathcal{B}_{\varepsilon_n}(\Omega)) \times \mathcal{C}_\Omega \times SV_0(\Omega)$ and $\varepsilon_n > 0$. Define $\boldsymbol{\varphi}_n(\mathbf{x}) = \mathbf{x} + \mathbf{u}_n(\mathbf{x})$, $\tilde{\boldsymbol{\varphi}}_n(\mathbf{x}) = \tilde{\boldsymbol{\varphi}}_{n-1} \circ \boldsymbol{\varphi}_n(\mathbf{x})$, $m_n(\mathbf{x}) = m_{n-1}(\mathbf{x}) + \delta m_n(\mathbf{x})$ and $s_n(\mathbf{x}) = s_{n-1}(\mathbf{x}) + \delta s_n(\mathbf{x})$.

Note that here the final deformation $\tilde{\boldsymbol{\varphi}}_n(\mathbf{x}) = \boldsymbol{\varphi}_0 \circ \boldsymbol{\varphi}_1 \circ \cdots \circ \boldsymbol{\varphi}_n(\mathbf{x})$, which implies that the multiscale approach (2.1)-(2.2) can simulate the large deformation well even if $\boldsymbol{\varphi}_i (i = 0, 1, \cdots, n)$ is small deformation. In addition, there are two key parameters λ_n, ε_n in the multiscale approach (2.1)-(2.2). These parameters determine whether or not one can find the global minimizer of $S_{lc}(\mathbf{u}, m, s)$ on $\mathcal{L}(\Omega)$. In practice, λ_n and ε_n are set to be large number and small positive number, respectively. However, it is still not enough. In order for the multiscale approach (2.1)-(2.2) to solve the greedy matching problem well, we shall give more precise condition shortly (see Theorem 2.5). Before that, concerning the existence of the solution for (2.2), we have the following result.

Theorem 2.1. *Assume* $\max_{\mathbf{x}\in\Omega} |T(\mathbf{x})| < M < +\infty$, $\max_{\mathbf{x}\in\Omega} |D(\mathbf{x})| < M < +\infty$ *and* $\triangle_T \triangleq \{\mathbf{x} : T(\mathbf{x})$ is discontinuous at $\mathbf{x}\}$ *is a zero measure set, then there exists at least one solution for (2.2).*

Proof. By selecting a minimizing sequence $\{(\mathbf{u}^k, \delta m^k, \delta s^k)\}$ of the functional $J_n(\mathbf{u}, \delta m, \delta s)$, one can conclude that $\mathbf{u}^k$, δm^k and δs^k are bounded on $[H^\alpha(\Omega)]^2$, $BV_0(\Omega)$ and $BV_0(\Omega)$, respectively, due to $J_n(\mathbf{u}, \delta m, \delta s) \leq J_n(\mathbf{0}, 0, 0)$.

Firstly, by the compactness of $H^\alpha(\Omega)$, there exists a subsequence of $\mathbf{u}^k$ which are still labelled by k and $\mathbf{u} \in [H^\alpha(\Omega)]^2$ such that $\mathbf{u}^k$ weakly converges to $\mathbf{u}$ with $R_1(\mathbf{u}) \leq \lim_{k\to\infty} \inf R_1(\mathbf{u}^k)$. By the compact embedding theorem (Theorem 4.58 in [9]), we know that $H_0^\alpha(\Omega) \hookrightarrow C^1(\Omega)$. Namely, there exists a subsequence of $\mathbf{u}^k$ which are still labelled by k and $\bar{\mathbf{u}} \in [C^1(\Omega)]^2$ such that $\mathbf{u}^k$ converges to $\bar{\mathbf{u}}$ in $[C^1(\Omega)]^2$. Moreover, by the uniqueness of the limitation, we get $\bar{\mathbf{u}} = \mathbf{u}$. That is, $\mathbf{u}^k \xrightarrow{k} \mathbf{u}$ in $[C^1(\Omega)]^2$. Therefore, we conclude $\mathbf{u} \in \mathcal{A}(\Omega) \setminus \mathcal{B}_{\varepsilon_n}(\Omega)$.

Secondly, by the compactness on $BV(\Omega)$, there exists a subsequence of δm^k which are still labelled by k and $\delta m \in BV(\Omega)$ such that δm^k weakly converges to δm with

$$(2.3) \qquad \|\delta m^k - \delta m\|_{L^1(\Omega)} \xrightarrow{k} 0 \text{ and } \int_\Omega \nabla \delta m^k \cdot \varphi d\mathbf{x} \xrightarrow{k} \int_\Omega \nabla \delta m \cdot \varphi d\mathbf{x}, \quad \forall \varphi \in C_0^\infty(\Omega),$$

where the first equation in (2.3) implies $m_{n-1} + \delta m \in C_\Omega$ and the second equation in (2.3) implies $R_2(m^k) \xrightarrow{k} R_2(m)$.

Similarly to the analysis on δm^k, one can conclude that there exists a subsequence of δs^k which are still labelled by k and $\delta s \in BV(\Omega)$ such that δs^k weakly converges to δs with

$$(2.4) \qquad \|\delta s^k - \delta s\|_{L^1(\Omega)} \xrightarrow{k} 0 \text{ and } \int_\Omega \nabla \delta s^k \cdot \varphi d\mathbf{x} \xrightarrow{k} \int_\Omega \nabla \delta s \cdot \varphi d\mathbf{x}, \quad \forall \varphi \in C_0^\infty(\Omega),$$

where the first equation in (2.4) implies $s_{n-1} + \delta m \in SV_0(\Omega)$ and the second equation in (2.4) implies $R_3(\delta s^k) \xrightarrow{k} R_3(\delta s)$.

Finally, by $\|\mathbf{u}^k - \mathbf{u}\|_{[C^1(\Omega)]^2} \xrightarrow{k} 0$, $\|\delta m^k - \delta m\|_{L^1(\Omega)} \xrightarrow{k} 0$ and $\|\delta s^k - \delta s\|_{L^1(\Omega)} \xrightarrow{k} 0$, we obtain $\int_\Omega (m_{n-1}(\mathbf{x}) + \delta m^k(\mathbf{x}) + \ln D(\mathbf{x}) - \ln(T \circ \tilde{\boldsymbol{\varphi}}_{n-1}(\mathbf{x} + \mathbf{u}^k(\mathbf{x})) - s_{n-1}(\mathbf{x}) - \delta s^k(\mathbf{x})))^2 d\mathbf{x} \xrightarrow{k}$

$\int_\Omega (m_{n-1}(\mathbf{x}) + \delta m(\mathbf{x}) + \ln D(\mathbf{x}) - \ln(T \circ \tilde{\boldsymbol{\varphi}}_{n-1}(\mathbf{x} + \mathbf{u}(\mathbf{x})) - s_{n-1}(\mathbf{x}) - \delta s(\mathbf{x})))^2 d\mathbf{x}$. Note here we use the fact $\tilde{\boldsymbol{\varphi}}_{n-1} \in [C^1(\Omega)]^2$, which implies that $T \circ \tilde{\boldsymbol{\varphi}}_{n-1}(\cdot)$ is continuous except for on some zero measure set. Therefore, $J_n(\mathbf{u}, \delta m, \delta s) \le \lim_{k\to\infty} \inf J_n(\mathbf{u}^k, \delta m^k, \delta s^k)$, which ensures the existence of solution for (2.2). ■

Then recall some important lemmas in [20], which are necessary for the proof of the convergence of the multiscale approach (2.1)-(2.2).

Lemma 2.2. *Assume $\mathbf{f}, \mathbf{g} : \Omega \to \Omega$, $\mathcal{W}(\mathbf{f}) = \mathbf{f} - \mathbf{I}$ and $\mathbf{I}$ is the identity mapping, then there holds*

(i) If $\mathcal{W}(\mathbf{f}) \in \mathcal{A}(\Omega) \setminus \mathcal{B}_{\varepsilon_1}(\Omega)$, $\mathcal{W}(\mathbf{g}) \in \mathcal{A}(\Omega) \setminus \mathcal{B}_{\varepsilon_2}(\Omega)$, then $\mathcal{W}(\mathbf{f} \circ \mathbf{g}) \in \mathcal{A}(\Omega) \setminus \mathcal{B}_{\varepsilon_1\varepsilon_2}(\Omega)$.

(ii) If $\mathcal{W}(\mathbf{f}) \in \mathcal{A}(\Omega) \setminus \mathcal{B}_{\varepsilon}(\Omega)$, then there exists $\mathbf{g} = \mathbf{f}^{-1} \in \mathcal{A}(\Omega)$.

(iii) Assume $\mathcal{W}(\mathbf{g}) \in \mathcal{A}(\Omega) \setminus \mathcal{B}_{\varepsilon}(\Omega)$, then there exists a constant C_1 such that $\int_\Omega \mathbf{f}(\mathbf{g}(\mathbf{x}))d\mathbf{x} \le C_1 R_1(\mathbf{g}^{-1}) \int_\Omega \mathbf{f}(\mathbf{y})d\mathbf{y}$.

Lemma 2.3. *Assume $\mathbf{p}(\mathbf{x}) = \mathbf{x} + \mathbf{u}(\mathbf{x})$ and $\mathcal{W}(\mathbf{q}) \in \mathcal{A}(\Omega) \setminus \mathcal{B}_{\varepsilon}(\Omega)$, then there exists a constant C_2 such that $R_1(\mathcal{W}(\mathbf{p} \circ \mathbf{q})) = 2(R_1(\mathcal{W}(\mathbf{q})) + C_2 R_1(\mathbf{q}^{-1}) R_1(\mathcal{W}(\mathbf{p})))$.*

Lemma 2.4. *Assume $\boldsymbol{\varphi}(\mathbf{x}) = \mathbf{x} + \mathbf{u}(\mathbf{x})$, $\mathbf{g}(\mathbf{x}) = \boldsymbol{\varphi}^{-1}(\mathbf{x}) = \mathbf{x} + \mathbf{v}(\mathbf{x})$ and $\mathbf{u}, \mathbf{v} \in \mathcal{A}(\Omega) \setminus \mathcal{B}_{\varepsilon}(\Omega)$, then there exists a constant C_3 such that $R_1(\mathbf{u}) = \int_\Omega \|\nabla^\alpha \mathbf{u}(\mathbf{x})\|^2 d\mathbf{x} \le C_3 R_1(\mathbf{g}) R_1(\mathcal{W}(\mathbf{g}))$.*

Based on these lemmas, we are now ready to give the result on the convergence of the multiscale approach (2.1)-(2.2).

By setting $m \equiv 0$, $s \equiv 0$ and $\mathbf{u} \equiv \mathbf{0}$, it follows from $J_n(\mathbf{u}_n, \delta m_n, \delta s_n) \le J_n(\mathbf{0}, 0, 0)$ in (2.2) that

$$\lambda_n S_{lc}^n(\mathbf{u}_n, m_n, s_n) + \mu R(\mathbf{u}_n, \delta m_n, \delta s_n) \le \lambda_n S_{lc}^{n-1}(\mathbf{u}_{n-1}, m_{n-1}, s_{n-1}),$$

where $S_{lc}^n(\mathbf{u}_n, m_n, s_n) = \int_\Omega (m_n(\mathbf{x}) + \ln D(\mathbf{x}) - \ln(T \circ \tilde{\boldsymbol{\varphi}}_n(\mathbf{x}) - s_n(\mathbf{x})))^2 d\mathbf{x}$ and $R(\mathbf{0}, 0, 0) = 0$. Hence, $S_{lc}^n(\mathbf{u}_n, m_n, s_n)$ is a decreasing sequence with lower bound, whose limitation is defined by

$$\delta = \lim_{n\to+\infty} S_{lc}^n(\mathbf{u}_n, m_n, s_n). \tag{2.5}$$

Define

$$\phi = \inf_{(\mathbf{u},m,s)\in\mathcal{L}(\Omega)} \int_\Omega (m(\mathbf{x}) + \ln D(\mathbf{x}) - \ln(T(\mathbf{x} + \mathbf{u}(\mathbf{x})) - s(\mathbf{x})))^2 d\mathbf{x}. \tag{2.6}$$

By proving $\delta = \phi$ under some suitable assumptions, we can give an answer to the problem of whether or not one can find the global minimizer of $S_{lc}(\mathbf{u}, m, s)$ on a proper set $\mathcal{L}(\Omega) = \mathcal{A}(\Omega) \times \mathcal{C}_\Omega \times SV_0(\Omega)$. Note (1.4)-(1.6) implies $\mathcal{A}(\Omega) \subseteq [H_0^\alpha(\Omega)]^2$, $\mathcal{C}_\Omega \subseteq BV(\Omega)$ and $SV_0(\Omega) \subseteq BV(\Omega)$. This ensures the greedy matching problem (2.6) is well regularized.

Theorem 2.5. *Let $\boldsymbol{\varphi}_n$, $\tilde{\boldsymbol{\varphi}}_n$, m_n and s_n be induced by the multiscale approach (2.1)-(2.2), and assume that $B = B(\Omega)$, M and λ_n are three positive numbers satisfying $\lim_{n\to+\infty} \frac{B^{4n-3}M^{4^n}}{\lambda_n} = 0$ and $\lim_{n\to+\infty} \varepsilon_n = 0$, where M is a positive number depending on $\mathbf{u}_0$, δm_0, δs_0, Ω, α and ϕ. Then there holds $\phi = \delta$.*

Proof. It is obvious that $\delta \geq \phi$. To show $\delta \leq \phi$, we use contradiction.

Assume $\delta > \phi$, then there exists a $C_1 \in (0,1)$ such that $\phi < C_1\delta < \delta$. By the definition of ϕ, there exists $\bar{\boldsymbol{\varphi}}(\mathbf{x}) = \mathbf{x} + \bar{\mathbf{u}}(\mathbf{x}) \in \mathcal{A}(\Omega)$, $\bar{m} \in \mathcal{C}_\Omega$ and $\bar{s} \in SV_0(\Omega)$ such that

$$\|\bar{m} + \ln D - \ln(T \circ \bar{\boldsymbol{\varphi}} - \bar{s})\|^2_{L^2(\Omega)} < C_1\delta. \tag{2.7}$$

Setting $\boldsymbol{\varphi} = \tilde{\boldsymbol{\varphi}}^{-1}_{n-1} \circ \bar{\boldsymbol{\varphi}}$, $m = \bar{m} - m_{n-1}$, $s = \bar{s} - s_{n-1}$, by Lemma 2.2, we obtain $\boldsymbol{\varphi} \in \mathcal{A}(\Omega)$, $m_{n-1} + m \in C_\Omega$ and $s_{n-1} + s \in SV_0(\Omega)$. By (2.2), (2.5) and (2.7), there holds

$$\begin{aligned}
&\lambda_n \int_\Omega (\tilde{m}_n(\mathbf{x}) + \ln D(\mathbf{x}) - \ln(T \circ \tilde{\boldsymbol{\varphi}}_n(\mathbf{x}) - \tilde{s}_n(\mathbf{x})))^2 d\mathbf{x} + \mu R(\mathbf{u}_n, \delta m_n, \delta s_n) \\
\leq &\lambda_n \|\bar{m} + \ln D - \ln(T \circ \boldsymbol{\varphi} - \bar{s})\|^2_{L^2(\Omega)} + \mu R(\mathcal{W}(\tilde{\boldsymbol{\varphi}}^{-1}_{n-1} \circ \bar{\boldsymbol{\varphi}}), \bar{m} - m_{n-1}, \bar{s} - s_{n-1}) \\
\leq &\lambda_n C_1\delta + \mu R(\mathcal{W}(\tilde{\boldsymbol{\varphi}}^{-1}_{n-1} \circ \bar{\boldsymbol{\varphi}}), \bar{m} - m_{n-1}, \bar{s} - s_{n-1}).
\end{aligned} \tag{2.8}$$

Then by (2.8), we further have

$$\lambda_n(1 - C_1)\delta + \mu R(\mathbf{u}_n, \delta m_n, \delta s_n) \leq \mu R(\mathcal{W}(\tilde{\boldsymbol{\varphi}}^{-1}_{n-1} \circ \bar{\boldsymbol{\varphi}}), \bar{m} - m_{n-1}, \bar{s} - s_{n-1}) \tag{2.9}$$

and

$$R(\mathbf{u}_n, \delta m_n, \delta s_n) \leq R(\mathcal{W}(\tilde{\boldsymbol{\varphi}}^{-1}_{n-1} \circ \bar{\boldsymbol{\varphi}}), \bar{m} - m_{n-1}, \bar{s} - s_{n-1}). \tag{2.10}$$

Recall $R(\mathbf{u}, m, s) = R_1(\mathbf{u}) + R_2(m) + R_3(s)$. Based on the inequality $|a + b| \leq |a| + |b|$ and the fact $m_{n-1} = m_{n-2} + \delta m_{n-1}$, $s_{n-1} = s_{n-2} + \delta s_{n-1}$, we obtain

$$R_2(\bar{m} - m_{n-1}) \leq R_2(\bar{m} - m_{n-2}) + R_2(\delta m_{n-1}),\ R_3(\bar{s} - s_{n-1}) \leq R_3(\bar{s} - s_{n-2}) + R_3(\delta s_{n-1}). \tag{2.11}$$

To estimate $R_1(\mathcal{W}(\tilde{\boldsymbol{\varphi}}^{-1}_{n-1} \circ \bar{\boldsymbol{\varphi}}))$, by Lemma 2.3, we obtain

$$R_1(\mathcal{W}(\tilde{\boldsymbol{\varphi}}^{-1}_{n-1} \circ \bar{\boldsymbol{\varphi}})) \leq 2R_1(\mathcal{W}(\tilde{\boldsymbol{\varphi}}^{-1}_{n-2} \circ \bar{\boldsymbol{\varphi}})) + 2CR_1((\tilde{\boldsymbol{\varphi}}^{-1}_{n-2} \circ \bar{\boldsymbol{\varphi}})^{-1}) R_1(\mathcal{W}(\boldsymbol{\varphi}^{-1}_{n-1})), \tag{2.12}$$

where we use the formula $\tilde{\boldsymbol{\varphi}}^{-1}_{n-1} \circ \bar{\boldsymbol{\varphi}} = \boldsymbol{\varphi}^{-1}_{n-1} \circ \tilde{\boldsymbol{\varphi}}^{-1}_{n-2} \circ \bar{\boldsymbol{\varphi}} = \boldsymbol{\varphi}^{-1}_{n-1} \circ (\tilde{\boldsymbol{\varphi}}^{-1}_{n-2} \circ \bar{\boldsymbol{\varphi}})$. Concerning the estimates on $R_1((\tilde{\boldsymbol{\varphi}}^{-1}_{n-2} \circ \bar{\boldsymbol{\varphi}})^{-1})$ and $R_1(\mathcal{W}(\boldsymbol{\varphi}^{-1}_{n-1}))$, we have

$$\begin{aligned}
R_1((\tilde{\boldsymbol{\varphi}}^{-1}_{n-2} \circ \bar{\boldsymbol{\varphi}})^{-1}) \leq &2R_1(\mathbf{x}) + 2R_1(\mathcal{W}((\tilde{\boldsymbol{\varphi}}^{-1}_{n-2} \circ \bar{\boldsymbol{\varphi}})^{-1})) \\
\leq &\tilde{c}_1 R_1(\tilde{\boldsymbol{\varphi}}^{-1}_{n-2} \circ \bar{\boldsymbol{\varphi}}) R_1(\mathcal{W}((\tilde{\boldsymbol{\varphi}}^{-1}_{n-2} \circ \bar{\boldsymbol{\varphi}}))) + \tilde{c}_2 \\
\leq &\tilde{c}_3 R_1^2(\mathcal{W}(\tilde{\boldsymbol{\varphi}}^{-1}_{n-2} \circ \bar{\boldsymbol{\varphi}})) + \tilde{c}_4 R_1(\mathcal{W}(\tilde{\boldsymbol{\varphi}}^{-1}_{n-2} \circ \bar{\boldsymbol{\varphi}})) + \tilde{c}_5 \\
\leq &B_1 \mathcal{M}^2((R_1(\mathcal{W}(\tilde{\boldsymbol{\varphi}}^{-1}_{n-2} \circ \bar{\boldsymbol{\varphi}}))))
\end{aligned} \tag{2.13}$$

and

$$\begin{aligned}
R_1(\mathcal{W}(\boldsymbol{\varphi}^{-1}_{n-1})) \leq &CR_1(\mathcal{W}(\boldsymbol{\varphi}_{n-1})) R_1(\boldsymbol{\varphi}_{n-1}) \\
\leq &CR_1(\mathcal{W}(\boldsymbol{\varphi}_{n-1}))(\bar{C} + R_1(\mathcal{W}(\boldsymbol{\varphi}_{n-1}))) \\
\leq &B_0 \mathcal{M}^2(R_1(\mathcal{W}(\boldsymbol{\varphi}_{n-1}))),
\end{aligned} \tag{2.14}$$

where for any $\xi \geq 0$,

$$\mathcal{M}(\xi) = \begin{cases} 1, & 0 \leq \xi \leq 1, \\ \xi, & \xi > 1. \end{cases}$$

Here the first and third inequality in (2.13) are based on the fact that $\mathbf{f}(\mathbf{x}) = \mathbf{x} + \mathcal{W}(\mathbf{f})$ for any deformation $\mathbf{f}$, and the second inequality in (2.13) is based on the conclusion $R_1(\mathcal{W}(\mathbf{g}^{-1})) \leq CR_1(\mathbf{g})R_1(\mathcal{W}(\mathbf{g}))$ in Lemma 2.4. Hence, by (2.12), (2.13) and (2.14), we get

$$R_1(\mathcal{W}(\tilde{\boldsymbol{\varphi}}_{n-1}^{-1} \circ \bar{\boldsymbol{\varphi}})) \leq 2R_1(\mathcal{W}(\tilde{\boldsymbol{\varphi}}_{n-2}^{-1} \circ \bar{\boldsymbol{\varphi}})) + B\mathcal{M}^2(R_1(\mathcal{W}(\boldsymbol{\varphi}_{n-1})))\mathcal{M}^2(R_1(\mathcal{W}(\tilde{\boldsymbol{\varphi}}_{n-2}^{-1} \circ \bar{\boldsymbol{\varphi}}))). \tag{2.15}$$

Furthermore, by (2.10), (2.11) and (2.15), we have

$$\begin{aligned} &R(\mathcal{W}(\tilde{\boldsymbol{\varphi}}_{n-1}^{-1} \circ \bar{\boldsymbol{\varphi}}), \bar{m} - m_{n-1}, \bar{s} - s_{n-1}) \\ \leq &2R(\mathcal{W}(\tilde{\boldsymbol{\varphi}}_{n-2}^{-1} \circ \bar{\boldsymbol{\varphi}}), \bar{m} - m_{n-2}, \bar{s} - s_{n-2}) + \bar{B}\mathcal{M}^4[R(\mathcal{W}(\tilde{\boldsymbol{\varphi}}_{n-2}^{-1} \circ \bar{\boldsymbol{\varphi}}), \bar{m} - m_{n-2}, \bar{s} - s_{n-2})] \\ \leq &B\mathcal{M}^4[R(\mathcal{W}(\tilde{\boldsymbol{\varphi}}_{n-2}^{-1} \circ \bar{\boldsymbol{\varphi}}), \bar{m} - m_{n-2}, \bar{s} - s_{n-2})] \\ \leq &\cdots \\ \leq &B^{4n-3}\mathcal{M}^{4^n}[R(\mathcal{W}(\tilde{\boldsymbol{\varphi}}_0^{-1} \circ \bar{\boldsymbol{\varphi}}), \bar{m} - m_0, \bar{s} - s_0)]. \end{aligned} \tag{2.16}$$

Define $M \triangleq \mathcal{M}[R(\mathcal{W}(\tilde{\boldsymbol{\varphi}}_0^{-1} \circ \bar{\boldsymbol{\varphi}}), \bar{m} - m_0, \bar{s} - s_0)]$. By (2.9) and (2.16), we then obtain $1 - C_1 \leq 0$ as $n \to +\infty$, which contradicts $C_1 \in (0, 1)$. Therefore, $\delta = \phi$. ■

Remark 2.6. By Theorem 2.5, the multiscale approach (2.1)-(2.2) provides a solution to the following 'greedy problem':

$$\inf_{(\mathbf{u},m,s) \in \mathcal{L}(\Omega)} S_{lc}(\mathbf{u}, m, s). \tag{2.17}$$

Here a key point is that the regularization in (2.17) is reflected on $\mathcal{L}(\Omega)$. Otherwise, the trivial solution (i.e., $\mathbf{u} = \mathbf{0}$, $m = \frac{T}{D}$, $s = 0$) may occur. In our method, some constraints (i.e., $\mathbf{u} \in \mathcal{A}(\Omega)$, $m \in \mathcal{C}_\Omega$, $s \in SV_0(\Omega)$) are additionally added in (1.4)-(1.6). Compared with the greedy problem in [20] that has nothing to do with parameters, the result of the greedy problem (2.17) is affected by two parameters K_1 and K_2. In applications, a practitioner needs to give some estimates on the intensity of varying illumination and set suitable K_1, K_2 (i.e., K_1, K_2 are suggested to be set near zero if no varying illumination in image pairs), then the multiscale approach (2.1)-(2.2) can work well to produce some expected solutions.

3. Alternating direction method for (2.2).

In this section, we mainly focus on the numerical implementation of the proposed multiscale approach (2.1)-(2.2) with λ_n and ε_n chosen by Theorem 2.5. To address the non-convexity of $S_{lc}(\mathbf{u}, m, s)$, an auxiliary variable $\mathbf{v}$ is additionally introduced and (2.2) is reformulated as follows:

$$(\mathbf{v}_n, \mathbf{u}_n, \delta m_n, \delta s_n) \in \mathop{\arg\min}_{(\mathbf{v},\mathbf{u},m_{n-1}+m,s_{n-1}+s) \in \overline{\mathcal{L}}_{\varepsilon_n}(\Omega)} E_n(\mathbf{v}, \mathbf{u}, m, s), \tag{3.1}$$

where $E_n(\mathbf{v},\mathbf{u},m,s)=\lambda_n\int_\Omega(m_{n-1}(\mathbf{x})+m(\mathbf{x})+\ln D(\mathbf{x})-\ln(T\circ\tilde{\boldsymbol{\varphi}}_{n-1}(\mathbf{x}+\mathbf{v}(\mathbf{x}))-s_{n-1}(\mathbf{x})-s(\mathbf{x})))^2d\mathbf{x}+\mu R(\mathbf{u},m,s)+\Theta R_c(\mathbf{u})+\frac{1}{2\theta_n}\int_\Omega|\mathbf{v}-\mathbf{u}|^2d\mathbf{x}$, $\overline{\mathcal{L}}_{\varepsilon_n}(\Omega)=L^2(\Omega)\times\mathcal{L}_{\varepsilon_n}(\Omega)$, $\theta_n>0$ is a small number, $\Theta>0$ is a large number and $R_c(\mathbf{u})=\int_\Omega(\frac{\partial u_1}{\partial x_1}-\frac{\partial u_2}{\partial x_2})^2+(\frac{\partial u_1}{\partial x_2}+\frac{\partial u_2}{\partial x_1})^2d\mathbf{x}$.

Then setting an initialization $\mathbf{v}_n^0=\mathbf{0}$, $\mathbf{u}_n^0=\mathbf{0}$, $\delta m_n^0=0$, $\delta s_n^0=0$ for some given scale n, (3.1) can be split into the following four subproblems:

$$\mathbf{v}_n^{k+1}\in\mathop{\arg\min}_{\mathbf{v}\in[L^2(\Omega)]^2}E_n(\mathbf{v},\mathbf{u}_n^k,\delta m_n^k,\delta s_n^k),\tag{3.2}$$

$$\mathbf{u}_n^{k+1}\in\mathop{\arg\min}_{\mathbf{u}\in[H_0^\alpha(\Omega)]^2}E_n(\mathbf{v}_n^{k+1},\mathbf{u},\delta m_n^k,\delta s_n^k),\tag{3.3}$$

$$\delta m_n^{k+1}=\mathop{\arg\min}_{m_{n-1}+m\in\mathcal{C}_\Omega}E_n(\mathbf{v}_n^{k+1},\mathbf{u}_n^{k+1},m,\delta s_n^k),\tag{3.4}$$

$$\delta s_n^{k+1}=\mathop{\arg\min}_{s_{n-1}+s\in SV_0(\Omega)}E_n(\mathbf{v}_n^{k+1},\mathbf{u}_n^{k+1},\delta m_n^{k+1},s),\tag{3.5}$$

for $k=0,1,2,\cdots$.

Concerning the convergence of $(\mathbf{v}_n^k,\mathbf{u}_n^k,\delta m_n^k,\delta s_n^k)$, here, we assume $\alpha>3.5$ for the technical demand to ensure $\boldsymbol{\varphi}\in[C^2(\Omega)]^2$. Before showing the convergence result, we give some lemmas for subproblems (3.2)-(3.5), which will be used in the later proof.

Lemma 3.1. *Suppose $\alpha>3.5$, $T(\cdot)$ is twice differentiable with $\operatorname*{ess\,sup}_{\mathbf{x}\in\Omega}|T(\mathbf{x})|<\bar{M}<+\infty$, $\operatorname*{ess\,sup}_{\mathbf{x}\in\Omega}|D(\mathbf{x})|<\bar{M}<+\infty$, $0<\theta_n<\frac{(\kappa-\kappa_0)^2}{10\overline{M}^2(\lambda_n\widetilde{M})^n}$, $\operatorname*{ess\,sup}_{\mathbf{x}\in\Omega}|\nabla T(\mathbf{x})|<\bar{M}<+\infty$ and $\operatorname*{ess\,sup}_{\mathbf{x}\in\Omega}|\nabla^2T(\mathbf{x})|<\bar{M}<+\infty$, where $\widetilde{M}\triangleq\widetilde{M}(\Omega,\alpha)=2C|\Omega|[K^2+\ln^2(\overline{M}/\kappa)^2]>0$, $K=\max\{|K_1|,|K_2|\}$ and $C=C(\Omega,\alpha)$ is a positive constant (see Lemma 3.2 and Lemma 3.3 in [16] for details). Then for subproblem (3.2), there exists a constant $c>0$ such that*

$$-2\lambda_n(m_{n-1}+m_n^k+\ln D-\ln T_n^{k+1})\frac{\nabla_{\mathbf{v}}T\circ\tilde{\boldsymbol{\varphi}}_{n-1}(\mathbf{x}+\mathbf{v}_n^{k+1})}{T_n^{k+1}}+\frac{1}{\theta_n}(\mathbf{v}_n^{k+1}-\mathbf{u}_n^k)=0\tag{3.6}$$

and

$$E_n(\mathbf{v}_n^k,\mathbf{u}_n^k,\delta m_n^k,\delta s_n^k)-E_n(\mathbf{v}_n^{k+1},\mathbf{u}_n^k,\delta m_n^k,\delta s_n^k)\geq c\|\mathbf{v}_n^{k+1}-\mathbf{v}_n^k\|^2_{[L^2(\Omega)]^2}.$$

Here and in what follow, $T_n^{k+1}=T\circ\tilde{\boldsymbol{\varphi}}_{n-1}(\mathbf{x}+\mathbf{v}_n^{k+1})-s_{n-1}-\delta s_n^k$.

Proof. The first-order variation of (3.2) is

$$\begin{aligned}-2\lambda_n\int_\Omega(m_{n-1}+m_n^k+\ln D-\ln T_n^{k+1})\frac{\nabla_{\mathbf{v}}T\circ\tilde{\boldsymbol{\varphi}}_{n-1}(\mathbf{x}+\mathbf{v}_n^{k+1})}{T_n^{k+1}}\cdot\mathbf{z}(\mathbf{x})d\mathbf{x}&\\+\frac{1}{\theta_n}\int_\Omega(\mathbf{v}_n^{k+1}-\mathbf{u}_n^k)\cdot\mathbf{z}(\mathbf{x})d\mathbf{x}=0,&\end{aligned}\tag{3.7}$$

where $\mathbf{z}$ is the test function. By the variational principle, this concludes (3.6).

Letting $\mathbf{z} = \mathbf{v}_n^k - \mathbf{v}_n^{k+1}$ in (3.7), it yields,

$$L(\mathbf{v}_n^k, \mathbf{v}_n^{k+1}) = \int_\Omega (-2\lambda_n(m_{n-1} + \delta m_n^k + \ln D - \ln T_n^{k+1}) \frac{\nabla_{\mathbf{v}} T \circ \tilde{\boldsymbol{\varphi}}_{n-1}(\mathbf{x} + \mathbf{v}_n^{k+1})}{T_n^{k+1}} + \frac{1}{\theta_n}(\mathbf{v}_n^{k+1} - \mathbf{u}_n^k)) \cdot (\mathbf{v}_n^k - \mathbf{v}_n^{k+1}) d\mathbf{x} = 0.$$

Then we have

(3.8)
$$\begin{aligned} &E_n(\mathbf{v}_n^k, \mathbf{u}_n^k, \delta m_n^k, \delta s_n^k) - E_n(\mathbf{v}_n^{k+1}, \mathbf{u}_n^k, \delta m_n^k, \delta s_n^k) \\ =&\lambda_n \int_\Omega \ln \frac{T \circ \tilde{\boldsymbol{\varphi}}_{n-1}(\mathbf{x} + \mathbf{v}_n^{k+1}) - s_{n-1} - \delta s_n^k}{T \circ \tilde{\boldsymbol{\varphi}}_{n-1}(\mathbf{x} + \mathbf{v}_n^k) - s_{n-1} - \delta s_n^k} \cdot (2m_{n-1} + 2\delta m_n^k + 2\ln D - 2\ln T_n^{k+1} \\ &+ \ln \frac{T \circ \tilde{\boldsymbol{\varphi}}_{n-1}(\mathbf{x} + \mathbf{v}_n^{k+1}) - s_{n-1} - \delta s_n^k}{T \circ \tilde{\boldsymbol{\varphi}}_{n-1}(\mathbf{x} + \mathbf{v}_n^k) - s_{n-1} - \delta s_n^k}) d\mathbf{x} + \frac{1}{2\theta_n} \int_\Omega (\mathbf{v}_n^k - \mathbf{v}_n^{k+1}) \cdot (\mathbf{v}_n^k + \mathbf{v}_n^{k+1} - 2\mathbf{u}_n^k) d\mathbf{x}. \end{aligned}$$

By using the Taylor's formula, we get

$$\ln(T \circ \tilde{\boldsymbol{\varphi}}_{n-1}(\mathbf{x} + \mathbf{v}_n^{k+1}) - s_{n-1} - \delta s_n^k) = \ln(T \circ \tilde{\boldsymbol{\varphi}}_{n-1}(\mathbf{x} + \mathbf{v}_n^k) - s_{n-1} - \delta s_n^k) + A + B, \tag{3.9}$$

where $A = \frac{\nabla_{\mathbf{v}} T \circ \tilde{\boldsymbol{\varphi}}_{n-1}(\mathbf{x} + \mathbf{v}_n^k) \cdot (\mathbf{v}_n^{k+1} - \mathbf{v}_n^k)}{T \circ \tilde{\boldsymbol{\varphi}}_{n-1}(\mathbf{x} + \mathbf{v}_n^k) - s_{n-1} - \delta s_n^k}$, $B = (\mathbf{v}_n^k - \mathbf{v}_n^{k+1}) H(\sigma) (\mathbf{v}_n^k - \mathbf{v}_n^{k+1})^T$, and $H(\sigma)$ is the Hessian matrix of function $\ln(T \circ \tilde{\boldsymbol{\varphi}}_{n-1}(\mathbf{x} + \mathbf{v}_n^k) - s_{n-1} - \delta s_n^k)$ on point σ between $\mathbf{v}_n^k$ and $\mathbf{v}_n^{k+1}$.

Hence, by (3.8) and (3.9), there holds

$$\begin{aligned} &E_n(\mathbf{v}_n^k, \mathbf{u}_n^k, \delta m_n^k, \delta s_n^k) - E_n(\mathbf{v}_n^{k+1}, \mathbf{u}_n^k, \delta m_n^k, \delta s_n^k) \\ \geq &\lambda_n \int_\Omega (A + B)^2 d\mathbf{x} + c \|\mathbf{v}_n^{k+1} - \mathbf{v}_n^k\|^2_{[L^2(\Omega)]^2} + L(\mathbf{v}_n^k, \mathbf{v}_n^{k+1}) \geq c \|\mathbf{v}_n^{k+1} - \mathbf{v}_n^k\|^2_{[L^2(\Omega)]^2}, \end{aligned}$$

where $c = \frac{1}{2\theta_n} - c_0 > 0$ and $c_0 = \|H(\sigma)\|_{L^\infty(\Omega)} = \frac{10\overline{M}^2 (\lambda_n \widetilde{M})^n}{(\kappa - \kappa_0)^2}$ (see Appendix A for details). ∎

Lemma 3.2. *For subproblem* (3.3)*, there holds*

(3.10)
$$\begin{aligned} &2\mu \int_\Omega \nabla^\alpha \mathbf{u}_n^{k+1} \cdot \nabla^\alpha \mathbf{w} d\mathbf{x} + 2\Theta \int_\Omega \left(\frac{\partial u_{n,1}^{k+1}}{\partial x_1} - \frac{\partial u_{n,2}^{k+1}}{\partial x_2} \right) \cdot \left(\frac{\partial w_1^{k+1}}{\partial x_1} - \frac{\partial w_2^{k+1}}{\partial x_2} \right) d\mathbf{x} \\ &+ 2\Theta \int_\Omega \left(\frac{\partial u_{n,1}^{k+1}}{\partial x_2} + \frac{\partial u_{n,2}^{k+1}}{\partial x_1} \right) \cdot \left(\frac{\partial w_1^{k+1}}{\partial x_2} + \frac{\partial w_2^{k+1}}{\partial x_1} \right) d\mathbf{x} - \frac{1}{\theta_n} \int_\Omega (\mathbf{v}_n^{k+1} - \mathbf{u}_n^{k+1}) \cdot \mathbf{w} d\mathbf{x} = 0, \end{aligned}$$

for any $\mathbf{w} \in [C_0^\infty(\Omega)]^2$*, and*

$$E_n(\mathbf{v}_n^{k+1}, \mathbf{u}_n^k, \delta m_n^k, \delta s_n^k) - E_n(\mathbf{v}_n^{k+1}, \mathbf{u}_n^{k+1}, \delta m_n^k, \delta s_n^k) \geq c_1 \|\mathbf{u}_n^k - \mathbf{u}_n^{k+1}\|^2_{[H_0^\alpha(\Omega)]^2},$$

for some $c_1 > 0$.

Proof. The first-order variation of (3.3) is

$$2\mu\int_\Omega \nabla^\alpha \mathbf{u}_n^{k+1}\cdot\nabla^\alpha \mathbf{w}d\mathbf{x} + 2\Theta\int_\Omega\left(\frac{\partial u_{n,1}^{k+1}}{\partial x_1}-\frac{\partial u_{n,2}^{k+1}}{\partial x_2}\right)\cdot\left(\frac{\partial w_1^{k+1}}{\partial x_1}-\frac{\partial w_2^{k+1}}{\partial x_2}\right)d\mathbf{x}$$
$$+2\Theta\int_\Omega\left(\frac{\partial u_{n,1}^{k+1}}{\partial x_2}+\frac{\partial u_{n,2}^{k+1}}{\partial x_1}\right)\cdot\left(\frac{\partial w_1^{k+1}}{\partial x_2}+\frac{\partial w_2^{k+1}}{\partial x_1}\right)d\mathbf{x} - \frac{1}{\theta_n}\int_\Omega(\mathbf{v}_n^{k+1}-\mathbf{u}_n^{k+1})\cdot\mathbf{w}d\mathbf{x}=0,$$

where $\mathbf{w}=(w_1,w_2)^T$ is a test function. This concludes (3.10).

Letting $\mathbf{w}=\mathbf{u}_n^k-\mathbf{u}_n^{k+1}$, then there holds

$$\begin{aligned}\tilde{L}(\mathbf{u}_n^k,\mathbf{u}_n^{k+1}) =&2\mu\int_\Omega \nabla^\alpha\mathbf{u}_n^{k+1}\cdot\nabla^\alpha(\mathbf{u}_n^k-\mathbf{u}_n^{k+1})d\mathbf{x}+\frac{1}{\theta_n}\int_\Omega(\mathbf{u}_n^{k+1}-\mathbf{v}_n^{k+1})\cdot(\mathbf{u}_n^k-\mathbf{u}_n^{k+1})d\mathbf{x}\\ &+2\Theta\int_\Omega\left(\frac{\partial u_{n,1}^{k+1}}{\partial x_1}-\frac{\partial u_{n,2}^{k+1}}{\partial x_2}\right)\left(\frac{\partial u_{n,1}^{k}}{\partial x_1}-\frac{\partial u_{n,2}^{k}}{\partial x_2}\right)d\mathbf{x}\\ &+2\Theta\int_\Omega\left(\frac{\partial u_{n,1}^{k+1}}{\partial x_2}+\frac{\partial u_{n,2}^{k+1}}{\partial x_1}\right)\left(\frac{\partial u_{n,1}^{k}}{\partial x_2}+\frac{\partial u_{n,2}^{k}}{\partial x_1}\right)d\mathbf{x}\\ &-2\Theta\int_\Omega\left(\frac{\partial u_{n,1}^{k+1}}{\partial x_1}-\frac{\partial u_{n,2}^{k+1}}{\partial x_2}\right)^2d\mathbf{x}-2\Theta\int_\Omega\left(\frac{\partial u_{n,1}^{k+1}}{\partial x_2}+\frac{\partial u_{n,2}^{k+1}}{\partial x_1}\right)^2d\mathbf{x}=0.\end{aligned}\tag{3.11}$$

Therefore, based on (3.11), we obtain

$$\begin{aligned}&E_n(\mathbf{v}_n^{k+1},\mathbf{u}_n^k,\delta m_n^k,\delta s_n^k)-E_n(\mathbf{v}_n^{k+1},\mathbf{u}_n^{k+1},\delta m_n^k,\delta s_n^k)\\ \ge&\mu\|\nabla^\alpha(\mathbf{u}_n^k-\mathbf{u}_n^{k+1})\|^2_{[L^2(\Omega)]^2}+\frac{1}{2\theta_n}\|\mathbf{u}_n^k-\mathbf{u}_n^{k+1}\|^2_{[L^2(\Omega)]^2}+\tilde{L}(\mathbf{u}_n^k,\mathbf{u}_n^{k+1})\\ \ge&c_1\|\mathbf{u}_n^k-\mathbf{u}_n^{k+1}\|^2_{[H_0^\alpha(\Omega)]^2}.\end{aligned}\tag{3.12}$$

■

Lemma 3.3. *For subproblems (3.4) and (3.5), there hold*

$$2\lambda_n(m_{n-1}+\delta m_n^{k+1}+\ln D-\ln T_n^{k+1})-\mu\mathrm{div}\left(\frac{\delta m_n^{k+1}}{|\delta m_n^{k+1}|}\right)=0,\tag{3.13}$$

$$2\lambda_n(m_{n-1}+\delta m_n^{k+1}+\ln D-\ln \tilde{T}_n^{k+1})\frac{1}{\tilde{T}_n^{k+1}}-\mu\mathrm{div}\left(\frac{\nabla\delta s_n^{k+1}}{|\nabla\delta s_n^{k+1}|}\right)=0,\tag{3.14}$$

$$E_n(\mathbf{v}_n^{k+1},\mathbf{u}_n^{k+1},\delta m_n^k,\delta s_n^k)-E_n(\mathbf{v}_n^{k+1},\mathbf{u}_n^{k+1},\delta m_n^{k+1},\delta s_n^k)\ge\lambda_n\|\delta m_n^k-\delta m_n^{k+1}\|^2_{L^2(\Omega)},\tag{3.15}$$

and

$$E_n(\mathbf{v}_n^{k+1},\mathbf{u}_n^{k+1},\delta m_n^{k+1},\delta s_n^k)-E_n(\mathbf{v}_n^{k+1},\mathbf{u}_n^{k+1},\delta m_n^{k+1},\delta s_n^{k+1})\ge\lambda_n\|\delta s_n^k-\delta s_n^{k+1}\|^2_{L^2(\Omega)},\tag{3.16}$$

where $\tilde{T}_n^{k+1}=T\circ\tilde{\boldsymbol{\varphi}}_{n-1}(\mathbf{x}+\mathbf{v}_n^{k+1})-s_{n-1}-\delta s_n^{k+1}$.

Proof. The first-order variation of (3.4) is

$$2\lambda_n \int_\Omega (m_{n-1} + \delta m_n^{k+1} + \ln D - \ln T_n^{k+1}) \cdot p d\mathbf{x} + \mu \int_\Omega \frac{\nabla \delta m_n^{k+1} \cdot \nabla p}{|\nabla \delta m_n^{k+1}|} d\mathbf{x} = 0,$$

where p is the test function. By using the integration-by-parts formula [13], we get (3.13).

Letting $p = \delta m_n^k - \delta m_n^{k+1}$, we have

$$\begin{aligned}\hat{L} =& 2\lambda_n \int_\Omega [m_{n-1} + \delta m_n^{k+1} + \ln D - \ln T_n^{k+1}] \cdot (\delta m_n^k - \delta m_n^{k+1}) d\mathbf{x} \\ &+ \mu \int_\Omega \frac{\nabla \delta m_n^{k+1} \cdot \nabla(\delta m_n^k - \delta m_n^{k+1})}{|\nabla \delta m_n^{k+1}|} d\mathbf{x} = 0.\end{aligned}$$

Then we obtain

$$\begin{aligned}&E_n(\mathbf{v}_n^{k+1}, \mathbf{u}_n^{k+1}, \delta m_n^k, \delta s_n^k) - E_n(\mathbf{v}_n^{k+1}, \mathbf{u}_n^{k+1}, \delta m_n^{k+1}, \delta s_n^k) \\ \geq& \lambda_n \|\delta m_n^k - \delta m_n^{k+1}\|_{L^2(\Omega)}^2 + \hat{L} = \lambda_n \|\delta m_n^k - \delta m_n^{k+1}\|_{L^2(\Omega)}^2.\end{aligned}$$

Further, by giving similar analysis on subproblem (3.5), we conclude (3.14) and (3.16). ■

Now, based on Lemmas 3.1-3.3, we can give a convergence result of the sequence $\{(\mathbf{v}_n^k, \mathbf{u}_n^k, \delta m_n^k, \delta s_n^k)\}$.

Theorem 3.4. *Suppose that the conditions in Lemmas 3.1-3.3 are satisfied. Then the sequence $\{(\mathbf{v}_n^k, \mathbf{u}_n^k, \delta m_n^k, \delta s_n^k)\}$ generated by (3.2)-(3.5) converges to the solution of (3.1) when $k \to +\infty$.*

Proof. First, we claim that there exists $(\mathbf{v}_n, \mathbf{u}_n, \delta m_n, \delta s_n) \in [L^2(\Omega)]^2 \times [H_0^\alpha(\Omega)]^2 \times \mathcal{C}_\Omega \times SV_0(\Omega)$ such that

$$\begin{aligned}&\mathbf{v}_n^k \xrightarrow{k} \mathbf{v}_n \quad \text{in} \quad [L^2(\Omega)]^2, \quad \mathbf{u}_n^k \xrightarrow{k} \mathbf{u}_n \quad \text{in} \quad [H_0^\alpha(\Omega)]^2, \\ &\delta m_n^k \xrightarrow{k} \delta m_n \quad \text{in} \quad \mathcal{C}_\Omega, \quad \delta s_n^k \xrightarrow{k} \delta s_n \quad \text{in} \quad SV_0(\Omega).\end{aligned} \tag{3.17}$$

By Lemmas 3.1-3.3, we obtain that

$$\begin{aligned}&E_n(\mathbf{v}_n^k, \mathbf{u}_n^k, \delta m_n^k, \delta s_n^k) - E_n(\mathbf{v}_n^{k+1}, \mathbf{u}_n^{k+1}, \delta m_n^{k+1}, \delta s_n^{k+1}) \\ \geq& c\|\mathbf{v}_n^k - \mathbf{v}_n^{k+1}\|_{[L^2(\Omega)]^2}^2 + c_1 \|\mathbf{u}_n^k - \mathbf{u}_n^{k+1}\|_{[H_0^\alpha(\Omega)]^2}^2 \\ &+ \lambda_n \|\delta m_n^k - \delta m_n^{k+1}\|_{L^2(\Omega)}^2 + \lambda_n \|\delta s_n^k - \delta s_n^{k+1}\|_{L^2(\Omega)}^2.\end{aligned} \tag{3.18}$$

Note that $E_n(\mathbf{v}_n^k, \mathbf{u}_n^k, \delta m_n^k, \delta s_n^k)$ is a decreasing sequence with a lower bound, which implies that the left side of (3.18) converges to zero when $k \to +\infty$. Hence, we have,

$$\begin{aligned}&\|\mathbf{v}_n^k - \mathbf{v}_n^{k+1}\|_{[L^2(\Omega)]^2}^2 \xrightarrow{k} 0, \quad \|\mathbf{u}_n^k - \mathbf{u}_n^{k+1}\|_{[H_0^\alpha(\Omega)]^2}^2 \xrightarrow{k} 0, \\ &\|\delta m_n^k - \delta m_n^{k+1}\|_{L^2(\Omega)}^2 \xrightarrow{k} 0, \quad \|\delta s_n^k - \delta s_n^{k+1}\|_{L^2(\Omega)}^2 \xrightarrow{k} 0,\end{aligned} \tag{3.19}$$

as $k \to +\infty$. Then by the compactness of Banach space $L^2(\Omega)$, $H_0^\alpha(\Omega)$, there exists $(\mathbf{v}_n, \mathbf{u}_n, \delta m_n, \delta s_n) \in [L^2(\Omega)]^2 \times [H_0^\alpha(\Omega)]^2 \times L^2(\Omega) \times L^2(\Omega)$ such that

$$\begin{aligned} &\mathbf{v}_n^k \xrightarrow{k} \mathbf{v}_n \quad \text{in} \quad [L^2(\Omega)]^2, \quad \mathbf{u}_n^k \xrightarrow{k} \mathbf{u}_n \quad \text{in} \quad [H_0^\alpha(\Omega)]^2, \\ &\delta m_n^k \xrightarrow{k} \delta m_n \quad \text{in} \quad L^2(\Omega), \quad \delta s_n^k \xrightarrow{k} \delta s_n \quad \text{in} \quad L^2(\Omega). \end{aligned}$$

In addition, since δm_n^k is bounded in $BV(\Omega)$, there exists a subsequence of δm_n^k which are still labelled with δm_n^k and $\delta \bar{m}_n \in BV_0(\Omega)$ such that

$$\|\delta m_n^k - \delta \bar{m}_n\|_{L^1(\Omega)} \xrightarrow{k} 0, \int_\Omega \nabla \delta m_n^k \cdot w d\mathbf{x} \xrightarrow{k} \int_\Omega \nabla \delta \bar{m}_n \cdot w d\mathbf{x}, \tag{3.20}$$

for any $w \in C_0^\infty(\Omega)$. By (3.19), (3.20) and the uniqueness of the limitation for m_n^k, there holds $\delta m_n = \delta \bar{m}_n \in BV_0(\Omega)$. So we have $m_{n-1} + \delta m_n \in \mathcal{C}_\Omega$. Similarly, we also have $s_{n-1} + \delta s_n \in SV_0(\Omega)$. Therefore, we obtain the claim (3.17).

Next, we claim that $(\mathbf{v}_n, \mathbf{u}_n, \delta m_n, \delta s_n)$ is a minimizer of (3.1).

By (3.6), (3.10) and (3.17), we know that

$$-2\lambda_n[m_{n-1} + m_n + \ln D - \ln T_n]\frac{\nabla_{\mathbf{v}} T \circ \tilde{\boldsymbol{\varphi}}_{n-1}(\mathbf{x} + \mathbf{v}_n)}{T_n} + \frac{1}{\theta_n}(\mathbf{v}_n - \mathbf{u}_n) = 0 \tag{3.21}$$

and

$$\begin{aligned} &2\mu \int_\Omega \nabla^\alpha \mathbf{u}_n \cdot \nabla^\alpha \mathbf{z} d\mathbf{x} + 2\Theta \int_\Omega (\frac{\partial u_{n,1}}{\partial x_1} - \frac{\partial u_{n,2}}{\partial x_2}) \cdot (\frac{\partial z_1}{\partial x_1} - \frac{\partial z_2}{\partial x_2}) d\mathbf{x} \\ &+2\Theta \int_\Omega (\frac{\partial u_{n,1}}{\partial x_2} + \frac{\partial u_{n,2}}{\partial x_1}) \cdot (\frac{\partial z_1}{\partial x_2} + \frac{\partial z_2}{\partial x_1}) d\mathbf{x} - \frac{1}{\theta_n} \int_\Omega (\mathbf{v}_n - \mathbf{u}_n) \cdot \mathbf{z} d\mathbf{x} = 0, \end{aligned} \tag{3.22}$$

where $T_n = T \circ \tilde{\boldsymbol{\varphi}}_{n-1}(\mathbf{x} + \mathbf{v}_n) - s_{n-1} - \delta s_n$. By (3.13), we also obtain that

$$2\lambda_n[m_{n-1} + \delta m_n^{k+1} + \ln D - \ln T_n^{k+1}] + \mathcal{H}(\delta m_n^{k+1}) = 0,$$

where $\mathcal{H}(m_n^{k+1}) = -\mu \mathrm{div}(\frac{\nabla \delta m_n^{k+1}}{|\nabla \delta m_n^{k+1}|})$. This implies,

$$\mathcal{H}(\delta m_n^{k+1}) \xrightarrow{k} -2\lambda_n[m_{n-1} + \delta m_n + \ln D - \ln T_n] \triangleq \mathcal{T}.$$

Note that $\mathcal{H}$ is a monotone operator because $\mathcal{H}$ is the derivative of a convex functional, which shows

$$\int_\Omega [\mathcal{H}(\delta m_n^{k+1}) - \mathcal{H}(\omega)] \cdot (\delta m_n^{k+1} - \omega) d\mathbf{x} \geq 0 \quad \forall \ \omega \in BV(\Omega).$$

Furthermore, there holds

$$\int_\Omega \mathcal{H}(\delta m_n^{k+1}) \cdot \delta m_n^{k+1} d\mathbf{x} \xrightarrow{k} \int_\Omega \mathcal{T} \cdot \delta m_n d\mathbf{x},$$

and

$$\int_\Omega \mathcal{H}(\omega) \cdot \delta m_n^{k+1} d\mathbf{x} \xrightarrow{k} \int_\Omega \mathcal{H}(\omega) \cdot \delta m_n d\mathbf{x}.$$

So we get

$$\int_\Omega [\mathcal{T} - \mathcal{H}(\omega)] \cdot (\delta m_n - \omega) d\mathbf{x} \geq 0.$$

Let $\omega = \delta m_n + h\psi$ for any $\psi \in C_0^\infty(\Omega)$. Then

$$\int_\Omega [\mathcal{T} - \mathcal{H}(\delta m_n + h\psi)] \cdot \psi d\mathbf{x} \leq 0.$$

Besides,

$$\int_\Omega \mathcal{H}(\delta m_n + h\psi) \cdot \psi d\mathbf{x} \xrightarrow{k} \int_\Omega \mathcal{H}(\delta m_n) \cdot \psi d\mathbf{x}$$

as $h \to 0$. Therefore,

$$\int_\Omega \mathcal{T} \cdot \psi d\mathbf{x} \leq \int_\Omega \mathcal{H}(\delta m_n) \cdot \psi d\mathbf{x}.$$

So we have $\mathcal{H}(\delta m_n) = \mathcal{T}$ and

$$2\lambda_n[m_{n-1} + \delta m_n + \ln D - \ln T_n] + \mathcal{H}(\delta m_n) = 0. \tag{3.23}$$

In a similar way, there holds,

$$2\lambda_n[m_{n-1} + \delta m_n + \ln D - \ln T_n]\frac{1}{T_n} - \mu \mathrm{div}(\frac{\nabla \delta s_n}{|\nabla \delta s_n|}) = 0. \tag{3.24}$$

Then by (3.21), (3.22), (3.23) and (3.24), we conclude that $(\mathbf{v}_n, \mathbf{u}_n, \delta m_n, \delta s_n)$ is a minimizer of (3.1). ∎

At the end of this section, we focus on the numerical implementation of the subproblem (3.2)-(3.5). For some given domain $\Omega = (0, a) \times (0, a)$ and scale number n, we define $h = \frac{a}{N_S}$ for some given $N_S \in \mathbb{N}^+$. Here, we also define $(x_1)_i = ih$, $(x_2)_j = jh$ for $i, j = 0, 1, 2, \cdots, N_S$ and $\mathbf{P}_{i,j} = ((x_1)_i, (x_2)_j)$ for $i, j = 0, 1, 2, \cdots, N_S$.

v-problem: Define $r(\mathbf{v}) = m_{n-1} + \delta m_n^k + \ln D - \ln(T \circ \tilde{\boldsymbol{\varphi}}_{n-1}(\mathbf{x} + \mathbf{v}(\mathbf{x})) - s_{n-1} - \delta s_n^k)$. By using the Taylor's formula, there holds

$$r(\mathbf{v}_n^{k+1}) \approx r(\mathbf{u}_n^k) - \mathbf{L}^k \cdot (\mathbf{v}_n^{k+1} - \mathbf{u}_n^k), \tag{3.25}$$

where $\mathbf{L}^k = (L_x, L_y)^T = \frac{1}{T \circ \tilde{\boldsymbol{\varphi}}_{n-1}(\mathbf{x} + \mathbf{u}_n^k(\mathbf{x})) - s_{n-1} - \delta s_n^k} \nabla_{\mathbf{u}} T \circ \tilde{\boldsymbol{\varphi}}_{n-1}(\mathbf{x} + \mathbf{u}_n^k(\mathbf{x}))$. Substituting (3.25) into (3.2), we obtain the following Euler-Lagrange equation for (3.2):

$$\mathbf{G}\mathbf{v}_n^{k+1} = \boldsymbol{\phi}(\mathbf{u}_n^k), \qquad \forall \mathbf{x} \in \Omega, \tag{3.26}$$

where

$$\mathbf{G} = \begin{pmatrix} 1+2\lambda_{n-1}\theta_n L_x^2, & 2\lambda_{n-1}\theta_n L_x L_y \\ 2\lambda_{n-1}\theta_n L_x L_y, & 1+2\lambda_{n-1}\theta_n L_y^2 \end{pmatrix}, \ \mathbf{v}_n^{k+1} = \begin{pmatrix} v_{n,1}^{k+1} \\ v_{n,2}^{k+1} \end{pmatrix},$$

$$\boldsymbol{\phi}(\mathbf{u}_n^k) = \begin{pmatrix} u_{n,1}^k + 2\lambda_{n-1}\theta_n[r(\mathbf{u}_n^k)L_x + L_x^2 u_{n,1}^k + L_x L_y u_{n,2}^k] \\ u_{n,2}^k + 2\lambda_{n-1}\theta_n[r(\mathbf{u}_n^k)L_y + L_x L_y u_{n,1}^k + L_y^2 u_{n,2}^k] \end{pmatrix},$$

and $\mathbf{u}_n^k \triangleq (u_{n,1}^k, u_{n,2}^k)^T$. By solving the linear system (3.26), one gets the updated $\mathbf{v}_n^{k+1}$ for each $\mathbf{x} \in \Omega$.

u-problem: Based on the updated $\mathbf{v}_n^{k+1}$, the Euler-Lagrange equation for $\mathbf{u}$ from (3.3) is formulated as follows:

$$2\mu\theta_n \mathrm{div}^{\alpha*}(\nabla^\alpha \mathbf{u}_n^{k+1}) - 2\Theta\theta_n \triangle \mathbf{u}_n^{k+1} + \mathbf{u}_n^{k+1} = \mathbf{v}_n^{k+1} \qquad \forall \mathbf{x} \in \Omega. \tag{3.27}$$

Concerning the numerical computation of (3.27), the multigrid method is used to accelerate the algorithm. The details of the multigrid method for (3.27) are listed in Appendix B.

m-**problem:** By ignoring the constant term, (3.4) is essentially equivalent to the following convex optimization problem:

$$\delta m_n^{k+1} = \underset{m_{n-1}+m \in \mathcal{C}_\Omega}{\arg\min} \ \frac{1}{2}\int_\Omega |g-m|^2 d\mathbf{x} + \frac{\mu}{2\lambda_{n-1}}\int_\Omega |\nabla m| d\mathbf{x} \ \forall \mathbf{x} \in \Omega, \tag{3.28}$$

where $g(\mathbf{x}) = \ln(T \circ \tilde{\boldsymbol{\varphi}}_{n-1}(\mathbf{x} + \mathbf{v}^{k+1}) - s_{n-1}(\mathbf{x}) - \delta s_n^k(\mathbf{x})) - m_{n-1}(\mathbf{x}) - \ln D(\mathbf{x})$. Without the constraint $K_1 \le m \le K_2$, (3.28) is essentially a standard form of total variation (TV) minimization. The solution of (3.28) (without constraint $K_1 \le m \le K_2$) is,

$$m = g - \mathcal{P}_{\lambda K}(g). \tag{3.29}$$

Note that here and in what follows, $\mathcal{P}_A(v)$ denotes the element in A which minimizes the distance between v and all the elements in A. Here we use the Chambolle Projection algorithm [2] to compute $\mathcal{P}_{\lambda K}(g)$. By giving the initial value $\mathbf{p}^0 = (0,0)$, $0 < \tau < \frac{1}{8}$ and the following iterative sequence

$$\mathbf{p}_{i,j}^{l+1} = \frac{\mathbf{p}^l + \tau\nabla(\mathrm{div}\mathbf{p}^l - g/\lambda)_{i,j}}{1 + \tau|\nabla(\mathrm{div}\mathbf{p}^l - g/\lambda)_{i,j}|},$$

$\lambda \mathrm{div}\mathbf{p}^l \to \mathcal{P}_{\lambda K}(g)$ with $l \to +\infty$ [2]. Then, based on (3.29), we get the solution of (3.28) by projecting the solution of (3.29) onto the set $\mathcal{C}_\Omega$:

$$(\delta m_n^{k+1})_{i,j} = \begin{cases} [g - \mathcal{P}_{\lambda K}(g)]_{i,j}, & K_1 \le [g - \mathcal{P}_{\lambda K}(g)]_{i,j} \le K_2, \\ K_1, & [g - \mathcal{P}_{\lambda K}(g)]_{i,j} < K_1, \\ K_2, & [g - \mathcal{P}_{\lambda K}(g)]_{i,j} > K_2, \end{cases} \tag{3.30}$$

for $i, j = 0, 1, 2, \cdots, N_S$.

s-problem: Define $\mathcal{G}(\Omega) = \{s \in BV_0(\Omega) | s(\mathbf{x}) \leq \kappa - \kappa_0 - s_{n-1}(\mathbf{x}) \text{ for } \forall\, \mathbf{x} \in \Omega\}$. Then $\mathcal{G}(\Omega)$ is a closed and convex set. Assume that δs_n is a solution of (3.3). Then for any $r \in \mathcal{G}(\Omega)$, there holds $\delta s_n + \tau(r - \delta s_n) = (1-\tau)\delta s_n + \tau r \in \mathcal{G}(\Omega)$ for $0 \leq \tau \leq 1$. Next, we define $J(\tau) = E_n(\mathbf{v}_n^{k+1}, \mathbf{u}_n^{k+1}, \delta m_n^{k+1}, \delta s_n + \tau(r - \delta s_n))$, which yields

$$J(0) \leq J(\tau) \qquad \forall \tau \in [0,1].$$

Therefore,

$$0 \leq J'(0) = \int_\Omega \mathcal{F}(\delta s_n) \cdot (r - \delta s_n) d\mathbf{x}, \qquad \text{for } \forall\ r \in \mathcal{G}, \tag{3.31}$$

where $\mathcal{F}(\delta s_n) = \frac{2\lambda_n(m_{n-1} + \delta m_n^{k+1} + \ln D - \ln(T \circ \tilde{\boldsymbol{\varphi}}_{n-1}(\mathbf{x} + \mathbf{v}_n^{k+1}) - s_{n-1} - \delta s_n))}{T \circ \tilde{\boldsymbol{\varphi}}_{n-1}(\mathbf{x} + \mathbf{v}_n^{k+1}) - s_{n-1} - \delta s_n} - \mu \mathrm{div}\left(\frac{\nabla \delta s_n}{|\nabla \delta s_n|}\right)$. Note that $\mathcal{F}$ is a monotone operator [22] and (3.31) is equivalent to $\delta s_n = \mathcal{P}_\mathcal{G}[\delta s_n - \varrho \mathcal{F}(\delta s_n)]$, which induces the following iterative method for (3.3):

$$\delta s_n^{l+1} = \mathcal{P}_\mathcal{G}[\delta s_n^l - \varrho \mathcal{F}(\delta s_n^l)], \quad l = 0, 1, 2, \cdots, \tag{3.32}$$

with $\varrho > 0$. Concerning the projection in (3.32), it is essentially to solve the following optimization problem:

$$(\delta s_n^{l+1})_{i,j} = \arg\min_{w_{i,j}} \|[\delta s_n^l - \varrho \mathcal{F}(\delta s_n^l)]_{i,j} - w_{i,j}\|^2,$$

subject to $w_{i,j} \leq \kappa - \kappa_0 - (s_{n-1})_{i,j}$ for $i, j = 0, 1, 2, \cdots, N_S$. That is,

$$(\delta s_n^{l+1})_{i,j} = \begin{cases} [\delta s_n^l - \varrho \mathcal{F}(\delta s_n^l)]_{i,j}, & [\delta s_n^l - \varrho \mathcal{F}(s_n^l)]_{i,j} \leq \kappa - \kappa_0 - (s_{n-1})_{i,j}, \\ \kappa - \kappa_0 - (s_{n-1})_{i,j}, & [\delta s_n^l - \varrho \mathcal{F}(\delta s_n^l)]_{i,j} > \kappa - \kappa_0 - (s_{n-1})_{i,j}, \end{cases} \tag{3.33}$$

for $i, j = 0, 1, 2, \cdots, N_S$.

To summarize, the ADM algorithm for solving (3.1) is listed in Algorithm 3.1. Furthermore, based on Algorithm 3.1, we propose Algorithm 3.2 to implement the multiscale approach (2.1)-(2.2), which will be refined next based on the view of the multi-resolution.

Algorithm 3.1 ADM for (3.1)

Initialization: $k = 0$, $\mathbf{u}_n^0 = \mathbf{0}$, $\mathbf{v}_n^0 = \mathbf{0}$, $m_n^0 = 0$, $s_n^0 = 0$, Ω and maximum iteration times K.
while $k \leq K$
  Step 1. Update $\mathbf{v}_n^{k+1}$ using (3.26);
  Step 2. Update $\mathbf{u}_n^{k+1}$ using (3.27);
  Step 3. Update δm_n^{k+1} using (3.30);
  Step 4. Update δs_n^{k+1} using (3.33);
  Set $k = k + 1$;
endwhile
Output: $\mathbf{u}_n = \mathbf{u}_n^K$,$\mathbf{v}_n = \mathbf{v}_n^K$,$\delta m_n = \delta m_n^K$, $\delta s_n = \delta s_n^K$.

Algorithm 3.2 Multiscale algorithm for (2.1)-(2.2)

Initialization: $n = 0$, $\mathbf{u}_n^0 = \mathbf{0}$, $\mathbf{v}_n^0 = \mathbf{0}$, $\delta m_n^0 = 0$, $\delta s_n^0 = 0$, λ_n, $\theta_n(n = 0, 1, 2, \cdots, N)$, Θ and maximum scale N.
while $n \leq N$
  Step 1. Use Algorithm 3.1 to compute $\mathbf{u}_n$, $\mathbf{v}_n$, δm_n and δs_n on Ω;
  Step 2. Compute $\tilde{\varphi}_n$, m_n and s_n on Ω;
  Set $n = n + 1$;
endwhile
Output: $\tilde{\varphi}_N$,m_N, s_N.

4. Coarse-to-fine strategy for the multiscale approach. To solve the multiscale approach (2.1)-(2.2), one needs to iteratively solve the subproblem (3.2)-(3.5) for each scale n (See Algorithm 3.2 for details). This strategy is not yet efficient. Based on the view of the multi-resolution, we now propose a modified coarse-to-fine strategy for the numerical implementation of the multiscale approach (2.1)-(2.2). This strategy contains following two steps (the flow chart of the proposed coarse-to-fine strategy is shown in Fig 3. Note that here and in what follows, $\Omega \downarrow 2^n$ denotes the downsampling of the region Ω with size 2^n. For example, given the region $\Omega = (0, 128) \times (0, 128)$, $\Omega \downarrow 2^1$ denotes the region $(0, 64) \times (0, 64)$):

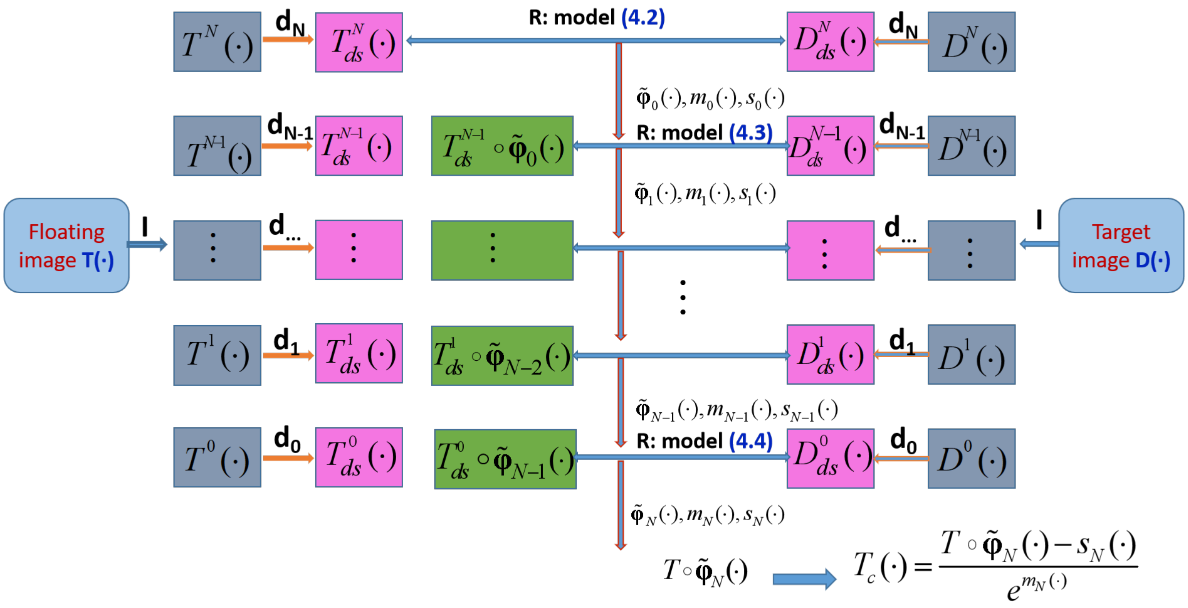


Figure 3. *The flow chart of the proposed coarse-to-fine strategy for the diffeomorphic image registration joint intensity correction. Note that here* **I** *denotes the image decomposition process;* **R** *denotes the image registration process;* **d** *denotes the downsampling process.*

(I) Image decomposition: To improve the resolution of downsampled images, image decomposition process is additionally introduced. Here, the decomposition model we used is the canonical multiscale image decomposition model developed by [30, 33]. This model is essentially the following definite partial differential equation (PDE) problem:

$$\text{(4.1)} \quad \begin{cases} \dfrac{\partial \xi(\mathbf{x},t)}{\partial t} = \rho(t)\mathrm{div}\left(\dfrac{\delta(\mathbf{x})\nabla\xi(\mathbf{x},t)}{|\nabla\xi(\mathbf{x},t)|}\right), & \mathbf{x}\in\Omega, t>0, \\ \xi(\mathbf{x},0) = f(\mathbf{x}), & \mathbf{x}\in\overline{\Omega}, \\ \xi(\mathbf{x},t)\mid_{\mathbf{x}\in\partial\Omega} = 0, & t>0, \end{cases}$$

where we set $\rho(t) = 1.05^t$, $\delta(\mathbf{x}) = \frac{1}{\sqrt{1+|\nabla(G_\iota * f)(\mathbf{x})|^2/\beta^2}}$, $\beta = 0.07$ and G_ι is a Gaussian kernel with a small standard deviation ι. By choosing $N+1$ different time points $0 = t_0 < t_1 < \cdots < t_N$ and setting $f = T$ or $f = D$, we obtain the image decomposition results: T^N, T^{N-1}, $\cdots$, T^0 and D^N, D^{N-1}, $\cdots$, D^0, respectively. Concerning the numerical implementation of (4.1), one can refer to [33] for details. Therefore, we downsample the decomposed images $T^n(\cdot), D^n(\cdot)$ $(n = 0,1,2,\cdots,N)$ with size 2^n to obtain the downsampled images $T^n_{ds}(\cdot)$ and $D^n_{ds}(\cdot)$, respectively.

(II) Image registration: The coarse-to-fine strategy for a multiscale approach of problem joint image registration and intensity correction model is divided into the following $N+1$ steps, here and in what follows, $\Omega_n = \Omega \downarrow 2^{N-n}(n = 0,1,2,\cdots,N)$:

Step 0. By taking $T^N_{ds}(\cdot)$ and $D^N_{ds}(\cdot)$ as the floating image and the target image, respectively, we solve the following variational problem on Ω_0:

$$\text{(4.2)} \quad (\mathbf{u}_0, \delta m_0, \delta s_0) \in \operatorname*{arg\,min}_{(\mathbf{u},m,s)\in\mathcal{L}_{\varepsilon_0}(\Omega_0)} \widetilde{J}_0(\mathbf{u},m,s),$$

where $\widetilde{J}_0(\mathbf{u},m,s) = \lambda_0\int_{\Omega_0}(m(\mathbf{x}) + \ln D^N_{ds}(\mathbf{x}) - \ln(T^N_{ds}(\mathbf{x}+\mathbf{u}(\mathbf{x})) - s(\mathbf{x})))^2 d\mathbf{x} + \mu R_{\Omega_0}(\mathbf{u},m,s)$, $\mathcal{L}_{\varepsilon_0}(\Omega_0) = (\mathcal{A}(\Omega_0)\setminus\mathcal{B}_{\varepsilon_0}(\Omega_0))\times\mathcal{C}_{\Omega_0}\times SV_0(\Omega_0)$ and $\varepsilon_0 > 0$. $R_{\Omega_n}(\mathbf{u},m,s)$ is defined by replacing Ω with Ω_n in (1.3). Define $\tilde{\boldsymbol{\varphi}}_0(\mathbf{x}) = \boldsymbol{\varphi}_0(\mathbf{x}) = \mathbf{x}+\mathbf{u}_0(\mathbf{x})$, $m_0(\mathbf{x}) = \delta m_0(\mathbf{x})$ and $s_0(\mathbf{x}) = \delta s_0(\mathbf{x})$ for each $\mathbf{x}\in\Omega_0$.

Step 1. Scale $\tilde{\boldsymbol{\varphi}}_0(\mathbf{x})$, $m_0(\mathbf{x})$ and $s_0(\mathbf{x})$ to Ω_1 and solve the following variational problem on Ω_1 (note that here $|\Omega_1| = 4|\Omega_0|$):

$$\text{(4.3)} \quad (\mathbf{u}_1, \delta m_1, \delta s_1) \in \operatorname*{arg\,min}_{(\mathbf{u},m_0+m,s_0+s)\in\mathcal{L}_{\varepsilon_1}(\Omega_1)} \widetilde{J}_1(\mathbf{u},m,s),$$

where $\widetilde{J}_1(\mathbf{u},m,s) = \lambda_1\int_{\Omega_1}(m_0(\mathbf{x}) + m(\mathbf{x}) + \ln D^{N-1}_{ds}(\mathbf{x}) - \ln(T^{N-1}_{ds}\circ\tilde{\boldsymbol{\varphi}}_0(\mathbf{x}+\mathbf{u}(\mathbf{x})) - s_0(\mathbf{x}) - s(\mathbf{x})))^2 d\mathbf{x} + \mu R_{\Omega_1}(\mathbf{u},m,s)$, $\mathcal{L}_{\varepsilon_1}(\Omega_1) = (\mathcal{A}(\Omega_1)\setminus\mathcal{B}_{\varepsilon_1}(\Omega_1))\times\mathcal{C}_{\Omega_1}\times SV_0(\Omega_1)$ and $\varepsilon_1 > 0$. Define $\boldsymbol{\varphi}_1(\mathbf{x}) = \mathbf{x}+\mathbf{u}_1(\mathbf{x})$, $\tilde{\boldsymbol{\varphi}}_1(\mathbf{x}) = \tilde{\boldsymbol{\varphi}}_0\circ\boldsymbol{\varphi}_1(\mathbf{x})$, $m_1(\mathbf{x}) = m_0(\mathbf{x})+\delta m_1(\mathbf{x})$ and $s_1(\mathbf{x}) = s_0(\mathbf{x})+\delta s_1(\mathbf{x})$ for each $\mathbf{x}\in\Omega_1$.

$$\vdots$$

Step N. Scale $\tilde{\boldsymbol{\varphi}}_{N-1}(\mathbf{x})$, $m_{N-1}(\mathbf{x})$ and $s_{N-1}(\mathbf{x})$ to Ω_N and solve the following variational problem on Ω_N (Note that $\Omega_N = \Omega$):

$$\text{(4.4)} \quad (\mathbf{u}_N, \delta m_N, \delta s_N) \in \operatorname*{arg\,min}_{(\mathbf{u},m_{N-1}+m,s_{N-1}+s)\in\mathcal{L}_{\varepsilon_N}(\Omega_N)} \widetilde{J}_N(\mathbf{u},m,s),$$

where $\widetilde{J}_N(\mathbf{u}, m, s) = \lambda_n \int_{\Omega_N} (m_{N-1}(\mathbf{x}) + m(\mathbf{x}) + \ln D(\mathbf{x}) - \ln(T \circ \tilde{\varphi}_{N-1}(\mathbf{x} + \mathbf{u}(\mathbf{x})) - s_{N-1}(\mathbf{x}) - s(\mathbf{x})))^2 d\mathbf{x} + \mu R_{\Omega_N}(\mathbf{u}, m, s)$, $\mathcal{L}_{\varepsilon_N}(\Omega_N) = (\mathcal{A}(\Omega_N) \setminus \mathcal{B}_{\varepsilon_N}(\Omega_N)) \times \mathcal{C}_{\Omega_N} \times SV_0(\Omega_N)$ and $\varepsilon_N > 0$. Define $\boldsymbol{\varphi}_N(\mathbf{x}) = \mathbf{x} + \mathbf{u}_N(\mathbf{x})$, $\tilde{\boldsymbol{\varphi}}_N(\mathbf{x}) = \tilde{\boldsymbol{\varphi}}_{N-1} \circ \boldsymbol{\varphi}_N(\mathbf{x})$, $m_N(\mathbf{x}) = m_{N-1}(\mathbf{x}) + \delta m_N(\mathbf{x})$ and $s_N(\mathbf{x}) = s_{N-1}(\mathbf{x}) + \delta s_N(\mathbf{x})$.

To show the convergence of the proposed coarse-to-fine strategy, we introduce some notations. In the coarse-to-fine strategy, one needs to scale the functions $\boldsymbol{\varphi}^{\Omega_n} : \Omega_n \to \Omega_n$, $m^{\Omega_n} : \Omega_n \to \mathbb{R}$, $s^{\Omega_n} : \Omega_n \to \mathbb{R}$ and $\mathbf{u}^{\Omega_n} : \Omega_n \to \mathbb{R}$ to the functions $\boldsymbol{\varphi} : \Omega \to \Omega$, $m : \Omega \to \mathbb{R}$, $s : \Omega \to \mathbb{R}$ and $\mathbf{u} : \Omega \to \mathbb{R}$, respectively. By the principle of scaling, there holds $\boldsymbol{\varphi}(\mathbf{y}) = \boldsymbol{\varphi}^{\Omega_n}(\frac{\mathbf{y}}{2^{N-n}})$, $\mathbf{u}(\mathbf{y}) = 2^{N-n}\mathbf{u}^{\Omega_n}(\frac{\mathbf{y}}{2^{N-n}})$, $m(\mathbf{y}) = m^{\Omega_n}(\frac{\mathbf{y}}{2^{N-n}})$, and $s(\mathbf{y}) = s^{\Omega_n}(\frac{\mathbf{y}}{2^{N-n}})$, where $\mathbf{y} \in \Omega$ and $\mathbf{x} = \mathbf{y}/2^{N-n} \in \Omega_n$. Here, functions $f_{\Omega_n}(f = \boldsymbol{\varphi}, m, s, \mathbf{u})$ denote the version of the function f on the domain Ω_n. In addition, there also holds $T^n_{ds}(\frac{\mathbf{y}}{2^{N-n}}) = T(\mathbf{y})$ and $D^n_{ds}(\frac{\mathbf{y}}{2^{N-n}}) = D(\mathbf{y})$.

Based on these notations, we have the following results.

Theorem 4.1. *For any $n \le N-1$, the coarse level registration problem*

$$(\mathbf{u}^{\Omega_n}_n, \delta m^{\Omega_n}_n, \delta s^{\Omega_n}_n) \in \underset{(\mathbf{u}^{\Omega_n}, m^{\Omega_n}, s^{\Omega_n}) \in \mathcal{L}_{\varepsilon_n}(\Omega_n)}{\arg\min} \widetilde{E}_n(\mathbf{u}^{\Omega_n}, m^{\Omega_n}, s^{\Omega_n}) \tag{4.5}$$

is equivalent to the following variational problem

$$(\mathbf{u}_n, \delta m_n, \delta s_n) \in \underset{(\mathbf{u}, m, s) \in \mathcal{L}_{\varepsilon_n}(\Omega)}{\arg\min} \overline{E}_n(\mathbf{u}, m, s), \tag{4.6}$$

where $\widetilde{E}_n(\mathbf{u}, m, s) = \lambda_n \int_{\Omega_n} (m^{\Omega_n}_{n-1}(\mathbf{x}) + m^{\Omega_n}(\mathbf{x}) + \ln D^{N-n}_{ds}(\mathbf{x}) - \ln(T^{N-n}_{ds} \circ \tilde{\boldsymbol{\varphi}}^{\Omega_n}_{n-1}(\mathbf{x} + \mathbf{u}^{\Omega_n}(\mathbf{x})) - s^{\Omega_n}_{n-1}(\mathbf{x}) - s^{\Omega_n}(\mathbf{x})))^2 d\mathbf{x} + \mu R_{\Omega_n}(\mathbf{u}^{\Omega_n}, m^{\Omega_n}, s^{\Omega_n})$, $\overline{E}_n(\mathbf{u}, m, s) = 4^{N-n}\lambda_n \int_\Omega (m_{n-1}(\mathbf{x}) + m(\mathbf{x}) + \ln D^{N-n}_{ds}(\mathbf{x}) - \ln(T^{N-n}_{ds} \circ \tilde{\boldsymbol{\varphi}}_{n-1}(\mathbf{x} + \mathbf{u}(\mathbf{x})) - s_{n-1}(\mathbf{x}) - s(\mathbf{x})))^2 d\mathbf{x} + \mu(R_{1,\Omega}(\mathbf{u}) + 2^{N-n}(R_{2,\Omega}(m) + R_{3,\Omega}(s)))$.

Proof. By letting $\mathbf{y} = 2^{N-n}\mathbf{x} \in \Omega$ for any $\mathbf{x} \in \Omega_n$, we get

$$\begin{aligned} &\widetilde{E}_n(\mathbf{u}^{\Omega_n}, m^{\Omega_n}, s^{\Omega_n}) \\ =&\frac{1}{4^{N-n}}\lambda_n \int_\Omega (m_{n-1}(\mathbf{y}) + m(\mathbf{y}) + \ln D^{N-n}_{ds}(\mathbf{y}) - \ln(T^{N-n}_{ds} \circ \tilde{\boldsymbol{\varphi}}_{n-1}(\mathbf{y} + \mathbf{u}(\mathbf{y})) \\ &- s_{n-1}(\mathbf{y}) - s(\mathbf{y})))^2 d\mathbf{y} + \frac{\mu}{16^{N-n}}(R_{1,\Omega}(\mathbf{u}) + 2^{N-n}(R_{2,\Omega}(m) + R_{3,\Omega}(s))). \end{aligned}$$

Therefore, (4.5) is equivalent to (4.6). ∎

By Theorem 4.1, the variational problem (4.2)-(4.4) on each coarse grid is equivalent to the following variational problem

$$(\mathbf{u}_n, \delta m_n, \delta s_n) \in \underset{(\mathbf{u}, m, s) \in \mathcal{L}_{\varepsilon_n}(\Omega)}{\arg\min} \overline{E}_n(\mathbf{u}, m, s), \quad n = 0, 1, 2, \cdots, N. \tag{4.7}$$

Then based on Theorems 4.1 and 2.5, we give the following convergence result of the proposed coarse-to-fine strategy (4.2)-(4.4).

Theorem 4.2. *Let $\tilde{\boldsymbol{\varphi}}_n$ and m_n, s_n $(n = 0, 1, 2, \cdots, N)$ be induced by the multiscale approach* (4.2)-(4.4). *Assume three large numbers $B = B(\Omega)$, M, λ_N satisfy $\lim\limits_{n\to N_-, N\to+\infty} \frac{B^{4n-3}M^{4^n}}{4^{N-n}\lambda_n} = 0$, where M is a positive number depending on $\mathbf{u}_0$, δm_0, δs_0, Ω, α and ϕ. Then there holds $\phi = \delta$, i.e., the modified coarse to fine strategy* (4.2)-(4.4) *is also equivalent to the original greedy matching problem* (2.17).

Proof. Based on 4.1, we can transform the variational problems (4.2)-(4.4) into an equivalent problem (4.7), which is defined on Ω. Based on (4.7), one can notice that (4.7) is equivalent to (2.2) with $n \to N_-$. Therefore, we can use Theorem 2.5 to show $\phi = \delta$. ∎

Based on Algorithm 3.1, the proposed coarse-to-fine strategy for the multiscale approach (4.2)-(4.4) is summarized in Algorithm 4.1.

Algorithm 4.1 Coarse-to-fine algorithm for the multiscale approach (4.2)-(4.4)

Initialization: $n = 0$, $\mathbf{u}_n^0 = \mathbf{0}$, $\mathbf{v}_n^0 = \mathbf{0}$, $m_n^0 = 0$, $s_n^0 = 0$, λ_n, θ_n $(n = 0, 1, 2, \cdots, N)$, Θ and maximum scale N.

I: Image decomposition:

Solve the image decomposition model (4.1) by setting $f = T$ and D to obtain the decomposition result; Downsample the decomposed images T^n, D^n $(n = 0, 1, 2, \cdots, N)$ with size 2^n to obtain the downsampled images $T_{ds}^n(\cdot)$ and $D_{ds}^n(\cdot)$, respectively.

II: Image registration:

while $n \le N$

Step 1. Use Algorithm 3.1 to compute $\mathbf{u}_n$, $\mathbf{v}_n$, δm_n and δs_n on Ω_n and replace $T(\cdot)$, $D(\cdot)$ with $T_{sd}^{N-n}(\cdot)$, $D_{sd}^{N-n}(\cdot)$, respectively;

Step 2. Compute $\tilde{\boldsymbol{\varphi}}_n$, m_n and s_n on Ω_n;

Step 3. Scale the definition of $\tilde{\boldsymbol{\varphi}}_n$, m_n and s_n onto a finer domain Ω_{n+1};

Set $n = n + 1$;

endwhile

Output: $\tilde{\boldsymbol{\varphi}}_N$,$m_N$, s_N and $T_c(\cdot) = \frac{T\circ\tilde{\boldsymbol{\varphi}}_N(\cdot) - s_N(\cdot)}{e^{m_N(\cdot)}}$.

Remark 4.3. Algorithm 4.1 is a multi-resolution modification for the Algorithm 3.2. In fact, one needs to solve the variational problem on Ω in Algorithm 3.2 while only needs to solve the same problem on $\Omega_n (n = 0, 1, 2, \cdots, N)$ in Algorithm 4.1. In fact, $|\Omega_n| = \frac{1}{4^{N-n}}|\Omega|$. This implies the Algorithm 4.1 accelerates the Algorithm 3.2, which will be validated in the numerical tests in Section 5.

5. Applications for the proposed multiscale approach.

In this section, we perform three different kinds of numerical tests to validate the theoretical results and Algorithms in Section 2-4. The content of this section contains: In Test 5.2, we perform the comparison between the proposed coarse-to-fine Algorithm 4.1 and M2FDIR in [20] to show that Algorithm 4.1 is more efficient on addressing the image registration problem with local varying illumination. In Test 5.3, a comparison between Algorithm 4.1 and Algorithm 3.2 is performed to show that the proposed Algorithm 4.1 has advantage on reducing the CPU consumption. In Test 5.4, the proposed Algorithm 4.1 is compared with some state-of-art image registration algorithms, like 1DFDIM [17], DFIRA [18], LDDMM [24] and FBNE [31]. All the numerical tests are performed under Windows 7 and MATLAB R2012b with Intel core i7-6700 CPU @3.40 GHz and 8GB memory. For the quantitative comparison, we choose the following two indexes:

- Relative sum of squared differences (Re_SSD for short) which is defined by

$$\mathrm{Re\_SSD}(T, D, \mathbf{u}) = \frac{\mathrm{SSD}(T(\mathbf{x}+\mathbf{u}), D)}{\mathrm{SSD}(T, D)},$$

  where $\mathrm{SSD}(T, D) = \frac{1}{2}\sum\limits_{i,j}(T_{i,j} - D_{i,j})^2$;
- Mesh folding number (MFN for short) which is defined by

$$\mathrm{MFN}(\mathbf{u}) = \sharp\left(\det \overline{J}(\mathbf{u}) \leq 0\right),$$

  where $\det \overline{J}(\mathbf{u}) = \left(1 + \frac{\partial u_1}{\partial x_1}\right)\left(1 + \frac{\partial u_2}{\partial x_2}\right) - \frac{\partial u_1}{\partial x_2}\frac{\partial u_2}{\partial x_1}$ and for any set A, $\sharp(A)$ denotes the number of elements in A.

5.1. Sensitivity test for parameters λ_n and μ in Algorithm 4.1. λ_n and μ are two key parameters for Algorithm 4.1. To show the sensitivity for the sequence $\{\lambda_n\}$ and the parameter μ, the synthetic image pair (pair I) is used as testing data. For pair I, the floating image and target image are defined as follows:

$$T(\mathbf{x}) = 255\chi_{\bar{\Gamma}_1}(\mathbf{x}) + 0.01,\ \ D(\mathbf{x}) = 255\chi_{\bar{\Gamma}_2}(\mathbf{x}) + 180\chi_{\bar{\Gamma}_3}(\mathbf{x}) + 0.01,$$

where $\Omega = (0, 128) \times (0, 128)$, $\bar{\Gamma}_1 = \{\mathbf{x} = (x_1, x_2)^T : (x_1 - 65)^2 + (x_2 - 65)^2 \leq 40^2\}$, $\bar{\Gamma}_2 = \{\mathbf{x} = (x_1, x_2)^T : (x_1 - 65)^2 + (x_2 - 65)^2 < 20^2\}$, $\bar{\Gamma}_3 = \{\mathbf{x} = (x_1, x_2)^T : 20^2 \leq (x_1 - 65)^2 + (x_2 - 65)^2 \leq 30^2\}$ and χ is an indicator function. The original synthetic image pair is shown in Fig 5.

By setting $\lambda_n = \lambda_0 \times 4^n (n = 0, 1, 2, \cdots)$ and $\mu \in [0.01, 1000]$ and $\lambda_0 \in [3000, 5500]$, we use the Algorithm 4.1 to perform the registration for image pair I by giving 546 different groups(only 169 groups are shown on Fig 4 to make the vision more plausible) of λ_0 and μ. By viewing the final $\mathrm{Re\_SSD}(T, D, \mathbf{u})$ as the heat value, the heat map for λ_0 and μ is shown in Fig 4.

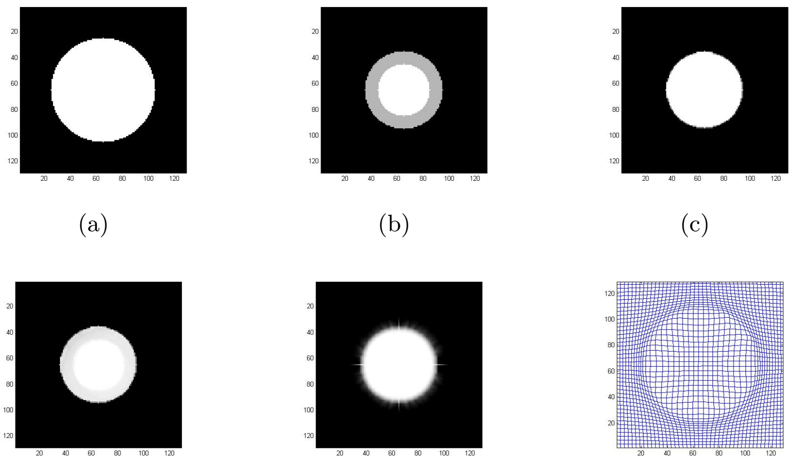
By Fig 4, we find that the final $\mathrm{Re\_SSD}(T, D, \mathbf{u})$ is not affected by the parameters λ_n and μ. This validates the fact that the multiscale approach (2.1)-(2.2) provides a solution to the greedy matching problem (2.17) which has nothing to do with the parameters λ_n and μ.

5.2. Comparison between the proposed coarse-to-fine Algorithm 4.1 and M2FDIR in [20]. To show that the proposed model via (4.2)-(4.4) properly treats the locally varying illumination, we compare the proposed Algorithm 4.1 with the multiscale M2FDIR in [20], which does not take the locally varying illumination into consideration.

The test pair in this part contains synthetic image pair I and two brain MRI image pairs (II-III) with local varying illumination.

For pair I, One can notice from Fig 5 that there is a shadow on the outer ring of the circle in the target image $D(\cdot)$, while no shadow appears in the floating image $T(\cdot)$. By using image pair I, we use the proposed Algorithm 4.1 and M2FDIR in [20] for registration. The final

Figure 4. *The heat map for λ_0 and μ*



(a) (b) (c)

(d) (e) (f)

Figure 5. *Comparison on pair I: (a) floating image $T(\cdot)$; (b) target image $D(\cdot)$; (c) $T \circ \tilde{\boldsymbol{\varphi}}_N(\cdot)$ in Algorithm 4.1,* Re_SSD=6.33%*; (d) $T_c \circ \tilde{\boldsymbol{\varphi}}_N(\cdot)$ in Algorithm 4.1,*Re_SSD=3.40%*; (e) $T \circ \tilde{\boldsymbol{\varphi}}_{K_M}(\cdot)$ in M2FDIR [20],* Re_SSD=5.14%*; (f) mesh grid of the deformation $\tilde{\boldsymbol{\varphi}}_N(\cdot)$ in Algorithm 4.1*

Table 1

Quantitative comparison between registration results of Algorithm 4.1 and M2FDIR (Test 5.2)

Data	Algorithm	Re-SSD(%)	MFN	CPU/s
Pair I	M2FDIR [20]	5.14	0	321.8
	Algorithm 4.1	**3.40**	**0**	**31.2**
Pair II	M2FDIR [20]	46.86	0	536.1
	Algorithm 4.1	**11.82**	**0**	**36.1**
Pair III	M2FDIR [20]	9.91	0	661.3
	Algorithm 4.1	**3.11**	**0**	**43.1**

registration results and quantitative comparison are listed in Fig 5 and Table 1. By Fig 5(f), one can notice that the proposed Algorithm 4.1 produces a diffeomorphic deformation $\boldsymbol{\varphi}$. It follows from Fig 5(d) that the registration result of Algorithm 4.1 matches the shadow ring of the target image $D(\cdot)$ well, while the final result of M2FDIR has trouble in matching the shadow ring. This shows that the proposed Algorithm 4.1 addresses the image registration with local illumination well.

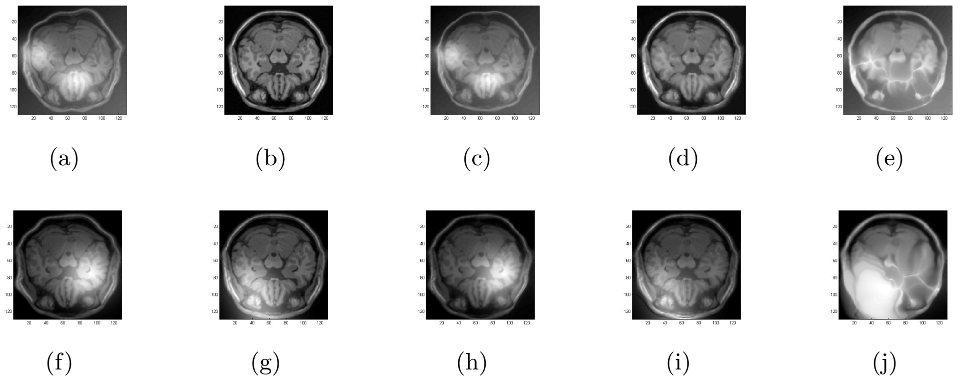


Figure 6. *Comparison on pair II (First row): (a) floating image $T(\cdot)$; (b) target image $D(\cdot)$; (c) $T \circ \tilde{\boldsymbol{\varphi}}_N(\cdot)$ in Algorithm 4.1,* Re_SSD=85.31%*; (d) $T_c \circ \tilde{\boldsymbol{\varphi}}_N(\cdot)$ in Algorithm 4.1,*Re_SSD=11.82%*; (e) $T \circ \tilde{\boldsymbol{\varphi}}_{K_M}(\cdot)$ in M2FDIR [20],* Re_SSD=46.86%*. Comparison on pair III (Second row): (f) floating image $T(\cdot)$; (g) target image $D(\cdot)$; (h) $T \circ \tilde{\boldsymbol{\varphi}}_N(\cdot)$ in Algorithm 4.1,* Re_SSD=86.04%*; (i) $T_c \circ \tilde{\boldsymbol{\varphi}}_N(\cdot)$ in Algorithm 4.1,* Re_SSD=3.11%*; (j) $T \circ \tilde{\boldsymbol{\varphi}}_{K_M}(\cdot)$ in M2FDIR [20],* Re_SSD=9.91%*;*

For pair II (see the first row of Fig 6), there are two domains which suffer from local varying illumination on the lower left of the floating image $T(\cdot)$, while no illumination appears in the target image $D(\cdot)$. For pair III (see the second row of Fig 6), there is local illumination on the right side of the floating image $T(\cdot)$, while local illumination appears on the opposite side of the target image $D(\cdot)$. We use Algorithm 4.1 and M2FDIR for pairs II and III. The results are shown in Fig 6 and the quantitative comparison result are listed in Table 1.

By Fig 6(e), we see that the registration result on pair II of M2FDIR is disturbed by the local varying illumination and leads to an unexpected result. In addition, one can notice from Fig 6(d) that the proposed Algorithm 4.1 addresses the local varying illumination well. This validates that the proposed Algorithm 4.1 has advantage on addressing the registration with local varying illumination over M2FDIR. This is also the main motivation for us to study the problem joint diffeomorphic image registration and intensity correction. Concerning the comparison on pair III, it follows from Fig 6(j) that the registration result is seriously bad in the region with local varying illumination, while the proposed Algorithm 4.1 can accurately correct the intensity distortion caused by the local illumination (see Fig 6(i) for details). This validates the conclusion that the proposed Algorithm 4.1 addresses intensity distorted registration well.

5.3. Comparison between the proposed coarse-to-fine Algorithm 4.1 and Algorithm 3.2 (without coarse-to-fine process). To solve the proposed multiscale approach (2.1))-(2.2)), one has two choices: (1) Use the proposed Algorithm 3.2 without coarse-to-fine strategy. For this choice, one is expected to implement the ADM process (4.2)-(4.4) on Ω for each scale

n. (2) Use the proposed coarse-to-fine Algorithm 4.1. For this choice, one only needs to solve the ADM process (4.2)-(4.4) on Ω_n for each scale n. Note that Ω_n is a domain smaller than Ω, which indicates that the proposed coarse-to-fine strategy (4.2)-(4.4) has advantage on reducing the CPU consumption over Algorithm 3.2.

To numerically validate this theoretical result, we perform the comparison between Algorithm 4.1 and Algorithm 3.2. Both these two algorithms aim to find the solution of the multiscale approach (2.1)-(2.2), where the coarse-to-fine strategy is introduced in Algorithm 4.1, while no multi-resolution based coarse-to-fine strategy is used in Algorithm 3.2. The data set used for the test are labelled IV-VI. For pair IV, it contains five image pairs which are collected at two different time from one patient (No.1) who suffers from mouth cavity lymphoma. Similarly, data V and VI contain the same content from some other two patients (No.2 and No.3). By registering these image pairs, clinicians can extract useful information from the difference between the deformed image $T_c \circ \tilde{\varphi}_N(\cdot)$ and the target image $D(\cdot)$. Furthermore, by analysing the difference, some evaluation for the severeness of the tumor is made. Therefore, the accuracy of the image registration result is of vital importance for the evaluation. In this part, we use Algorithm 4.1 and Algorithm 3.2 to register these 15 image pairs. The registration result for IV-VI are listed on Fig 7-Fig 9 and Table 2, where Re_SSD is represented by the mean value $\pm$ standard deviation of five different image pairs for each patient, and the CPU is represented in a similar way.

Table 2

Quantitative comparison between registration results of Algorithm 4.1 and Algorithm 3.2 (Test 5.3)

Data	Algorithm	Re-SSD(%)	MFN	CPU/s
data IV	**Algorithm 4.1**	10.67 ± 2.47	**0**	38.5 ± 5.3
	Algorithm 3.2	11.19 ± 3.44	0	518.5 ± 31.9
data V	**Algorithm 4.1**	9.82 ± 2.47	**0**	36.1 ± 6.1
	Algorithm 3.2	13.16 ± 3.51	0	436.7 ± 35.6
data VI	**Algorithm 4.1**	8.96 ± 1.68	**0**	43.1 ± 3.8
	Algorithm 3.2	12.76 ± 0.72	0	621.7 ± 45.6

By Table 2, we see that the registration result of the proposed Algorithm 4.1 is similar to (though a bit better than) Algorithm 3.2. However, the CPU consumption of Algorithm 4.1 is greatly reduced compared with Algorithm 3.2. This shows the efficiency of the proposed coarse-to-fine Algorithm 4.1.

5.4. Comparison between Algorithm 4.1 and some other image registration algorithms. In this part, to further validate the efficiency of the proposed coarse-to-fine Algorithm 4.1, we perform some comparison between Algorithm 4.1 and 1DFDIM [17], DFIRA [18], LDDMM [24] and FBNE [31]. For this purpose, we use these five algorithms to match three different medical image pairs which are labelled with VII-IX. Here, to show the efficiency of the proposed multiscale approach, VII-IX are kept the same state with data set used in [31]. These three image pairs are introduced as follows. For image pair VII, the floating image $T(\cdot)$ contains highly contrasted region in the middle of the region. By viewing the contrast as bias field relative to the target image $D(\cdot)$, the elimination of this kind of bias field provides a

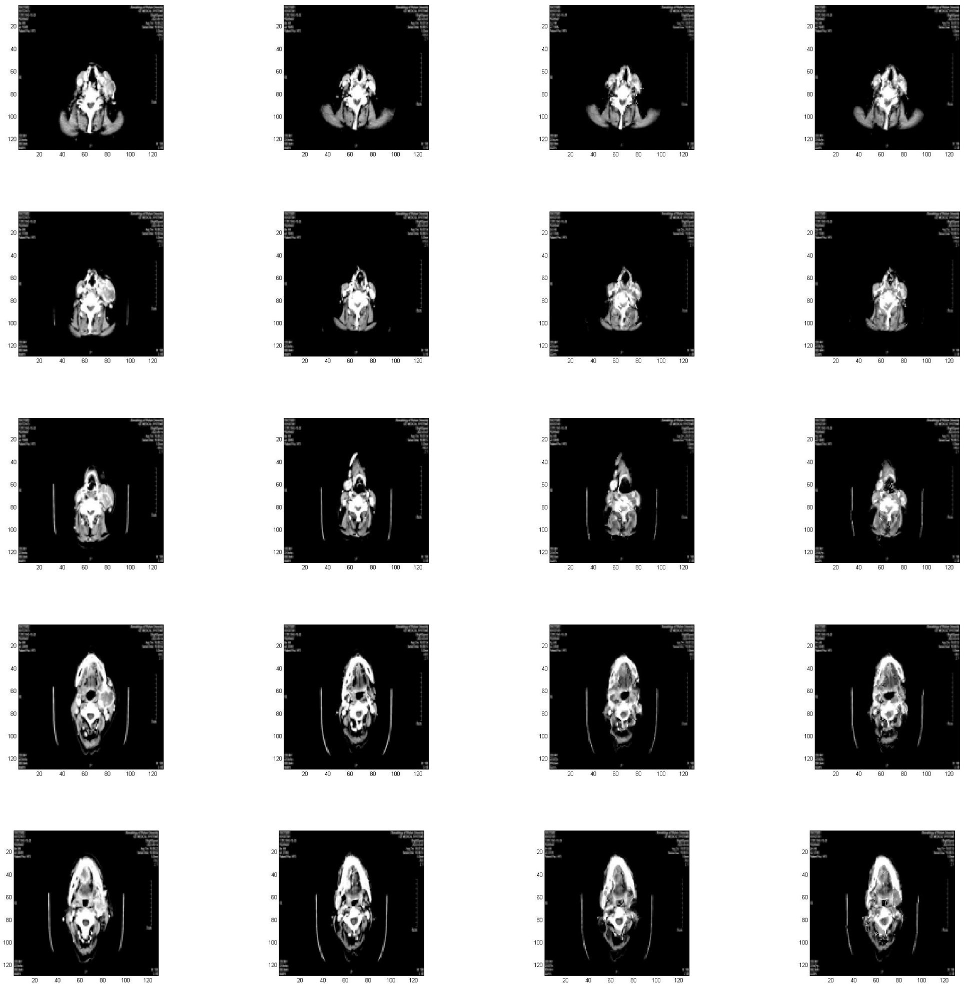


Figure 7. *Comparison on IV: The first column is the floating image $T(\cdot)$ for each image pair; The second column is the floating image $D(\cdot)$ for each image pair; The third and fourth columns are the image registration results of Algorithm 4.1 and Algorithm 3.2 for each image pair, respectively.*

strong evidence that the proposed multiscale approach for the variational model joint image registration and intensity correction has advantage on addressing the diffeomorphic image with local varying illumination. This is the main reason why these image pairs are selected for the numerical comparison. The quantitative comparison results for image pair VII are listed in Fig 10 and Table 3. One can notice from Fig 10 that only the proposed algorithm and the FBNE in [31] eliminate the bias field in the middle of the region well. The other three algorithms which do not take intensity correction into consideration lead to a narrow white bias field. This phenomenon occurs due to the minimization of the similarity $S(\mathbf{u})$. However, these solutions are not expected in image registration of image pair VII. This shows the necessity for introducing the intensity correction process in the proposed Algorithm 4.1. Note that [31] pursuits a minimizer of the cost functional with three different regularizations, while

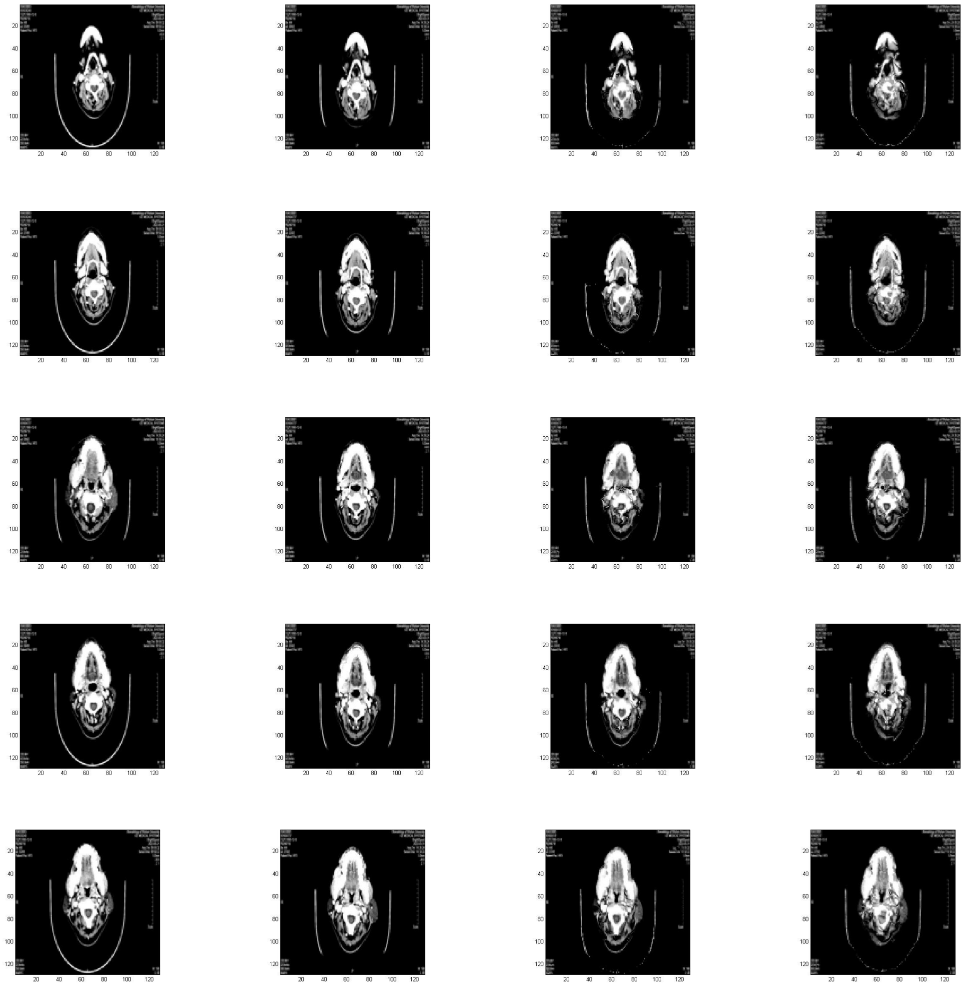


Figure 8. *Comparison on V: The first column is the floating image $T(\cdot)$ for each image pair; The second column is the floating image $D(\cdot)$ for each image pair; The third and fourth columns are the image registration results of Algorithm 4.1 and Algorithm 3.2 for each image pair, respectively.*

Algorithm 4.1 searches for the minimizer of $S(\mathbf{u}, m, s)$ without any regularization. The comparison between Algorithm 4.1 and FBNE algorithm in [31] further validates the advantage of greedy matching. However, without proper multiscale consideration, greedy matching without regularization is not expected to work well. This is the main reason why the multiscale approach is introduced in this paper.

For image pairs VIII, there is a low contrast in some local region of the floating image $T(\cdot)$, which may make it ineffective for some image registration models without intensity correction process. The registration result for image pair VIII is listed in Fig 11 and Table 3. This image with a low contrast in floating image $T(\cdot)$, the proposed algorithm and FBNE [31] successfully recover the low contrast region and lead to a final result with more details on the tissue. This shows the importance of intensity correction in the registration for these image pairs with a

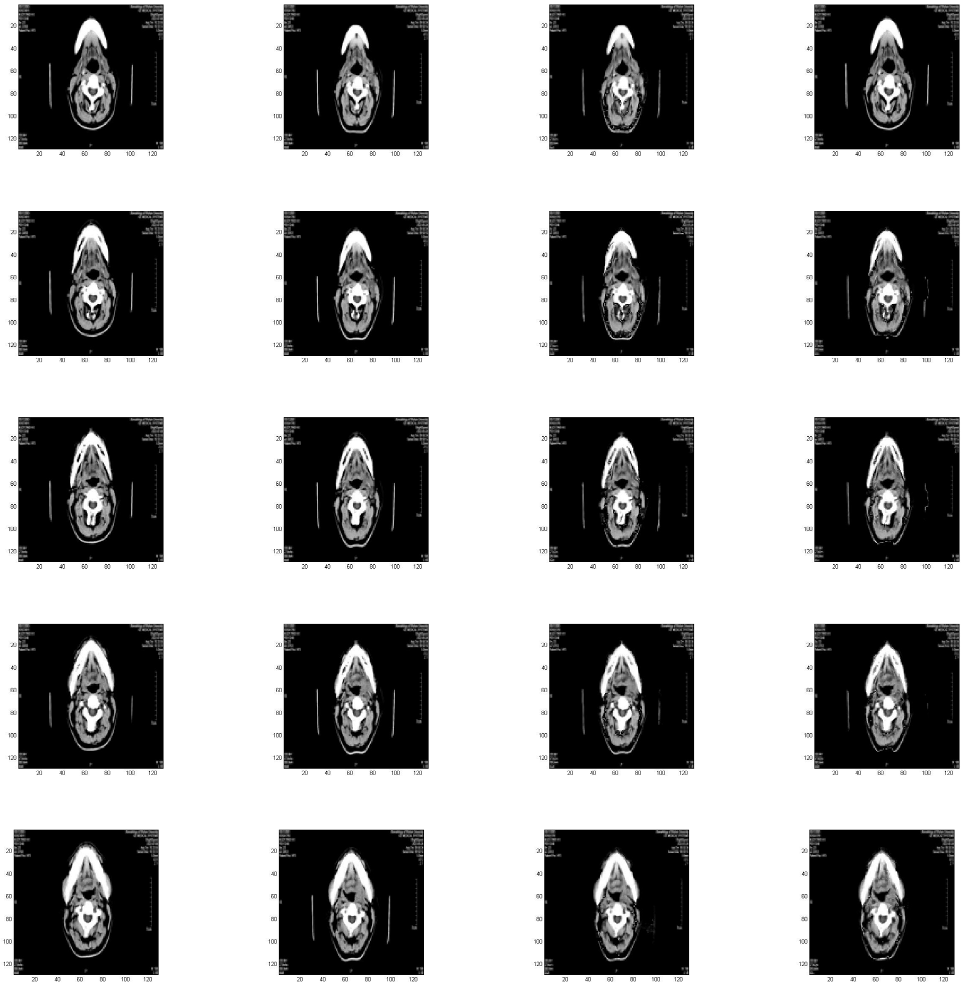


Figure 9. *Comparison on VI: The first column is the floating image $T(\cdot)$ for each image pair; The second column is the floating image $D(\cdot)$ for each image pair; The third and fourth columns are the image registration results of Algorithm 4.1 and Algorithm 3.2 for each image pair, respectively.*

low contrast floating image. In this view, it is helpful to use the proposed algorithm to register the image pairs which contains at least one high resolution image and one low contrast image. In addition, one can notice from Table 3 that the proposed Algorithm 4.1 achieves the best results for image pair VIII.

For image pair IX, the floating image contains bias field and varying illumination on different regions of the domain. Compared with image pair II used in Test 5.2, there is a square shadow surrounding the brain. This may affect the registration result. By Fig 12, we see that the local bias and square shadow are well eliminated in the final result of the proposed Algorithm 4.1 and FBNE. This is an advantage led by bias correction in the proposed Algorithm 4.1 and FBNE. Moreover, by the quantitative comparison on Table 3, one can see that the proposed Algorithm 4.1 achieves a smaller Re-SSD than FBNE.

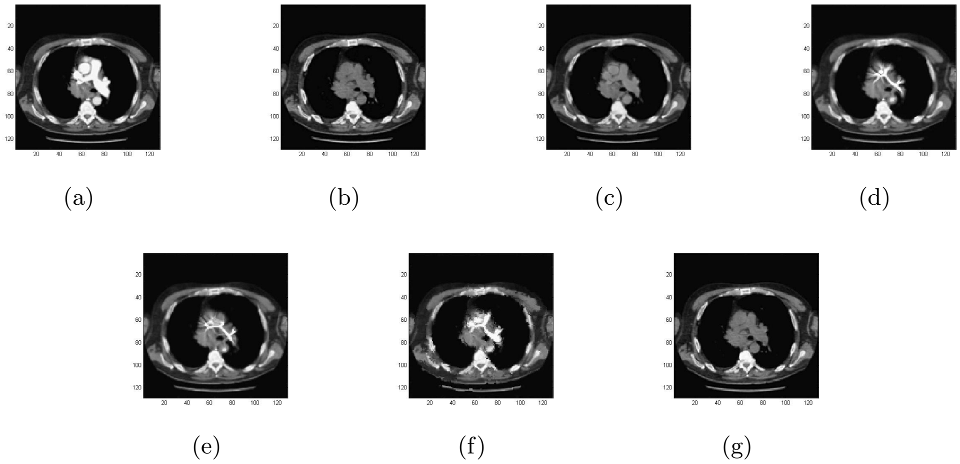


Figure 10. *Comparison on VII: (a) floating image $T(\cdot)$; (b) target image $D(\cdot)$; (c) $T \circ \tilde{\boldsymbol{\varphi}}_N(\cdot)$ in Algorithm 4.2,* Re_SSD=9.76%*; (d) 1DFDIM,* Re_SSD=25.72%*; (e) DFIRA,* Re_SSD=22.8%*; (f) LDDMM,* Re_SSD=50.05%*; (g) FBNE,* Re_SSD=12.57%.

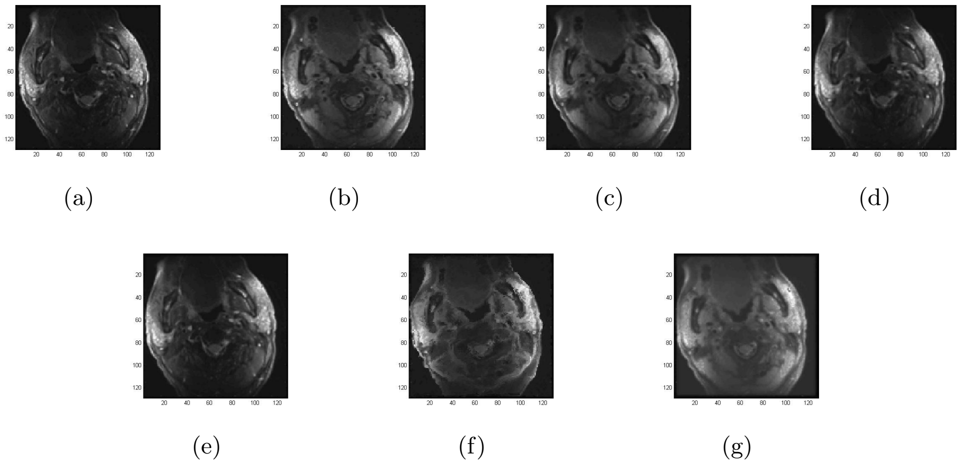
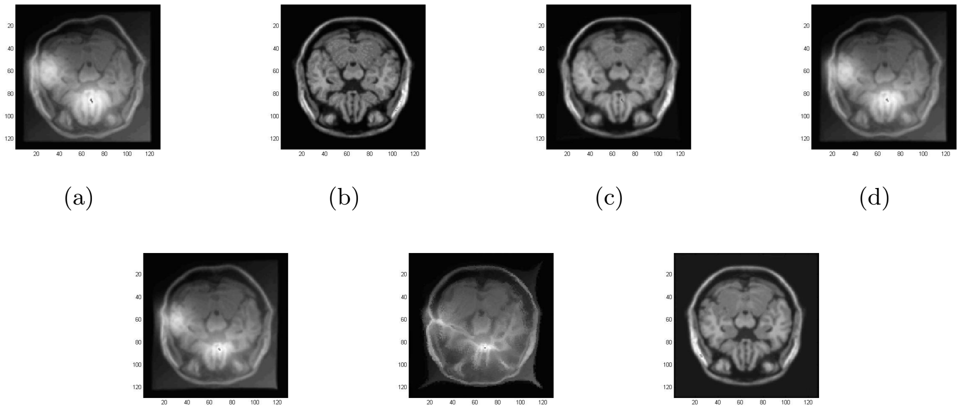


Figure 11. *Comparison on VIII: (a) floating image $T(\cdot)$; (b) target image $D(\cdot)$; (c) $T \circ \tilde{\boldsymbol{\varphi}}_N(\cdot)$ in Algorithm 4.1,* Re_SSD=5.28%*; (d) 1DFDIM,* Re_SSD=76.09%*; (e) DFIRA,* Re_SSD=66.54%*; (f) LDDMM,* Re_SSD=51.23%*; (g) FBNE,*Re_SSD=24.41%.

6. Conclusion. In this paper, we propose a variational model for joint diffeomorphic image registration and intensity correction. Based on the joint model, some related greedy matching problem (2.17) is proposed. For solving the greedy matching problem (2.17), the multiscale approach is introduced which addresses the instability by directly solving the greedy matching problem (2.17). This provides a theoretical support for this kind of research. For the numerical computation of the multiscale approach, an ADM method is proposed and the convergence of this process is proved. In addition, a coarse-to-fine strategy is introduced to accelerate the registration algorithm and the convergence of the coarse-to-fine strategy is proved. Finally, three different kinds of numerical tests are performed to validate the theoretical results in this



(a) (b) (c) (d) (e) (f) (g)

Figure 12. *Comparison on IX: (a) floating image $T(\cdot)$; (b) target image $D(\cdot)$; (c) $T \circ \tilde{\varphi}_N(\cdot)$ in Algorithm 4.1,* Re_SSD=3.05%*; (d) 1DFDIM,*Re_SSD=83.2%*; (e) DFIRA,* Re_SSD=68.96%*; (f) LDDMM,* Re_SSD=36.18%*; (g) FBNE,* Re_SSD=15.44%*.*

Table 3
Quantitative comparison between five different image registration algorithms

Data	Algorithm	Re-SSD(%)	MFN	CPU/s
data VII	**Algorithm 4.1**	**9.76**	**0**	**40.3**
	1DFDIM [17]	25.72	0	62.7
	DFIRA [18]	22.80	0	372.6
	LDDMM [24]	50.05	0	21.2
	FBNE [31]	12.57	0	118.6
data VIII	**Algorithm 4.1**	**5.28**	**0**	**42.2**
	1DFDIM [17]	76.09	0	86.3
	DFIRA [18]	66.54	0	456.7
	LDDMM [24]	51.23	0	30.2
	FBNE [31]	24.41	0	101.6
data IX	**Algorithm 4.1**	**3.05**	**0**	**50.4**
	1DFDIM [17]	83.2	0	90.3
	DFIRA [18]	68.96	0	748.4
	LDDMM [24]	36.18	0	23.7
	FBNE [31]	15.44	0	117.3

paper. For the future research, we may extend this work to the field of image registration joint segmentation.

Acknowledgments. The authors of this paper would like to thank Hang Li in Hospital of Stomatology, Wuhan University for providing the data for Test II.

REFERENCES

[1] M. Burger, J. Modersitzki, and L. Ruthotto, *A hyperelastic regularization energy for image registration*, SIAM J. Sci. Comput., 35(1) (2013), pp. 132–148, https://doi.org/10.1137/110835955.

[2] A. Chambolle, *An algorithm for total variation minimization and applications*, J. Math. Imaging Vis., 20 (2004), pp. 89–97, https://doi.org/10.1023/B:JMIV.0000011325.36760.1e.

[3] C. Chen and O. Oktem, *Indirect image registration with large diffeomorphic deformations*, SIAM J. Imaging Sci., 11 (2018), pp. 575–617, https://doi.org/10.1137/17M1134627.

[4] Y. Chen, J. Shi, M. Rao, and J. Lee, *Deformable multi-modal image registration by maximizing rnyi's statistical dependence measure*, Inverse Probl. Imaging, 9(1) (2015), pp. 79–103, https://doi.org/10.3934/ipi.2015.9.79.

[5] G. Christensen, R. Rabbitt, and M. Miller, *Deformable templates using large deformation kinematics*, IEEE Trans. Image Process., 5(10) (1996), pp. 1435–1447, https://doi.org/10.1109/83.536892.

[6] N. Chumchob, *Vectorial total variation-based regularization for variational image registration*, IEEE Trans. Image Process., 22(11) (2013), pp. 4551–4559, https://doi.org/10.1109/TIP.2013.2274749.

[7] N. Chumchob, K. Chen, and C. Loeza, *A fourth-order variational image registration model and its fast multi grid algorithm*, Multiscale Model. Simul., 9(1) (2011), pp. 89–128, https://doi.org/10.1137/100788239.

[8] L. K. Chun and L. M. Lui, *Landmark and intensity based registration with large deformations via quasi-conformal maps*, SIAM J. Imaging Sci., 7(4) (2013), pp. 2364–2392, https://doi.org/10.1137/130943406.

[9] F. Demengel, G. Demengel, and R. Erne, *Functional spaces for the theory of elliptic partial differential equations*, Springer London, 2012.

[10] M. Droske and M. Rumpf, *A variational approach to nonrigid morphological image registration*, SIAM J. Appl. Math, 64(2) (2004), pp. 668–687, https://doi.org/10.2307/4096004.

[11] M. Ebrahimi and S. Kulaseharan, *Deformable image registration and intensity correction of cardiac perfusion mri*, Statistical Atlases and Computational Models of the Heart - Imaging and Modelling Challenges, 8896 (2015), pp. 13–20, https://doi.org/10.1007/978-3-319-14678-2_2.

[12] M. Ebrahimi, A. Lausch, and A. L. Martel, *A gauss-newton approach to joint image registration and intensity correction*, Computer Methods and Programs in Biomedicine, 112(3) (2013), pp. 398–406, https://doi.org/10.1016/j.cmpb.2013.07.026.

[13] L. Evans, *Partial differential equations*, American Mathematical Society, 2010.

[14] D. Ferreira, E. Ribeiro, and C. Barcelos, *A variational approach to non-rigid image registration with bregman divergences and multiple features*, Pattern Recognition, 77 (2018), pp. 237–247, https://doi.org/10.1016/j.patcog.2017.12.015.

[15] E. Haber and J. Modersitzki, *A multilevel method for image registration*, SIAM J. Sci. Comput., 27(5) (2006), pp. 1594–1607, https://doi.org/10.1137/040608106.

[16] H. Han, *A variational model with fractional-order regularization term arising in registration of diffusion tensor image*, Inverse Probl. Imaging, 12(6) (2018), pp. 1263–1291, https://doi.org/aimsciences.org//article/doi/10.3934/ipi.2018053.

[17] H. Han and A. Wang, *A fast multi grid algorithm for 2D diffeomorphic image registration model*, J. Comput. Appl. Math., 394 (2021), p. 113567, https://doi.org/10.1016/j.cam.2021.113576.

[18] H. Han and Z. Wang, *A diffeomorphic image registration model with fractional-order regularization and Cauchy-Riemann constraint*, SIAM J. Imaging Sci., 13 (2020), pp. 1240–1271, https://doi.org/10.1137/19M1260621.

[19] H. Han and Z. Wang, *3d diffeomorphic image registration with cauchy-riemann constraint and lower bounded deformation divergence*, ESAIM Math. Model. Num., accpeted (2023), https://doi.org/10.1051/m2an/2022080.

[20] H. Han, Z. Wang, and Y. Zhang, *Multiscale approach for two-dimensional diffeomorphic image registration*, Multiscale Model. Simul., 19(4) (2021), pp. 1538–1572, https://doi.org/10.1137/20M1383987.

[21] H. Han, Z. Wang, and Y. Zhang, *Multiscale approach for three-dimensional conformal image registration*, SIAM J. Imaging Sci., 15(3) (2022), pp. 1431–1468, https://doi.org/10.1137/21M1455929.

[22] C. Kanzow and Y. Shehu, *Strong convergence of a double projection-type method for monotone variational inequalities in hilbert spaces*, J. Fixed Point Theory Appl., 20 (1) (2018), pp. 51–72, https://doi.org/doi.org/10.1007/s11784-018-0531-8.

[23] H. Li, W. Guo, J. Liu, L. Cui, and X. D., *Image segmentation with adaptive spatial priors from joint*

registration, SIAM J. Imaging Sci., 15 (2022), pp. 1314–1344, https://doi.org/10.1137/21M1444874.

[24] J. Li, Y. Shi, G. Tran, I. Dinov, D. Wang, and A. Toga, *Fast local trust region for diffusion tensor registration using exact reorientation and regularization*, IEEE Tans Med Imaging, 33(5) (2014), pp. 1–43, https://doi.org/10.1109/TMI.2013.2274051.

[25] L. M. Lui and T. C. Ng, *A splitting method for diffeomorphism optimization problem using beltrami coefficients*, J. Sci. Comput., 63 (2015), pp. 573–611, https://doi.org/10.1007/s10915-014-9903-4.

[26] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, *Diffeomorphic demons: Efficient non-parametric image registration*, IEEE Trans. Med. Imaging., 16(2) (1997), pp. 187–198, https://doi.org/10.1109/42.563664.

[27] K. Maleelai, C. Watchararuangwit, and N. Chumchob, *An improved numerical method for variational image registration model with intensity correction*, Journal of Interdisciplinary Mathematics, 25(4) (2022), pp. 1005–1022, https://doi.org/10.1080/09720502.2021.1889781.

[28] D. Qiu, K. C. Lam, and L. M. Lui, *Computing quasiconformal folds*, SIAM J. Imaging Sci., 12(3) (2019), pp. 1392–1424, https://doi.org/10.1137/18M1220042.

[29] D. Qiu and L. M. Lui, *Inconsistent surface registration via optimization of mapping distortions*, J. Sci. Comput., 83 (2020), pp. 1–31, https://doi.org/10.1007/s10915-020-01246-5.

[30] E. Tadmor and L. V. S. Nezzar, *A multiscale image representation using hierarchical (bv, l^2) decompositions*, Multiscale Model. Simul., 2 (2004), pp. 554–579, https://doi.org/10.1137/030600448.

[31] A. Theljani and K. Chen, *A nash game based variational model for joint image intensity correction and registration to deal with varying illumination*, Inverse Probl., 36 (2020), p. 034002, https://doi.org/10.1088/1361-6420/ab2934.

[32] T. Vercauteren, X. Pennecb, A. Perchanta, and N. Ayacheb, *Diffeomorphic demons:efficient non-parametric image registration*, Neuroimage, 45(1) (2009), pp. 61–72, https://doi.org/10.1109/TMI.2016.2610583.

[33] R. Xu, P. Athavale, A. Nachman, and G. A. Wright, *Multiscale registration of real-time and prior mri data for image-guided cardiac interventions*, IEEE Trans. Bio-Med. Eng., 61(10) (2014), pp. 2621–2632, https://doi.org/10.1109/TBME.2014.2324998.

[34] C. P. Yung, G. Choi, K. Chen, and L. M. Lui, *Efficient feature-based image registration by mapping sparsified surfaces*, J. Vis. Commun. Image R., 55 (2018), pp. 561–571, https://doi.org/10.1016/j.jvcir.2018.07.005.

[35] D. Zhang and K. Chen, *A novel diffeomorphic model for image registration and its algorithm*, J. Math. Imaging Vision, 2018 (2018), pp. 1–30, https://doi.org/10.1007/s10851-018-0811-3.

[36] D. Zhang and K. Chen, *3D orientation-preserving variational models for accurate image registration*, SIAM J. Imaging Sci., 13(3) (2020), pp. 1653–1691, https://doi.org/10.1137/20M1320006.

[37] D. Zhang, G. P. T. Choi, J. Zhang, and L. M. Lui, *A unifying framework for n-dimensional quasiconformal mappings*, SIAM J. Imaging Sci., 15(2) (2022), pp. 960–988, https://doi.org/10.1137/21M1457497.

[38] J. Zhang and K. Chen, *Variational image registration by a total fractional-order variation model*, J. Comput. Phys., 293 (2015), pp. 442–461, https://doi.org/10.1016/j.jcp.2015.02.021.

[39] J. Zhang, K. Chen, and B. Yu, *A novel high-order functional based image registration model with inequality constraint*, Comput. Math. Appl., 72(12) (2014), pp. 2887–2899, https://doi.org/10.1016/j.camwa.2016.10.018.

Appendix A. Estimation on L^∞ norm for the Hessian matrix of $\ln(T\circ\tilde{\boldsymbol{\varphi}}_{n-1}(\mathbf{x}+\mathbf{v}_n^k)-s_{n-1}-\delta s_n^k)$ with respect to $\mathbf{v}_n^k$.

The Hessian matrix of $\ln(T\circ\tilde{\boldsymbol{\varphi}}_{n-1}(\mathbf{x}+\mathbf{v}_n^k)-s_{n-1}-\delta s_n^k)$ is formulated as

$$H=\frac{1}{(T\circ\tilde{\boldsymbol{\varphi}}_{n-1}(\mathbf{x}+\mathbf{v}_n^k)-s_{n-1}-\delta s_n^k)^2}\begin{pmatrix} H_1 & H_2\\ H_2 & H_3\end{pmatrix},$$

where $H_1=T\circ\tilde{\boldsymbol{\varphi}}_{n-1}(\mathbf{x}+\mathbf{v}_n^k)\frac{\partial^2}{\partial x_1^2}\left(T\circ\tilde{\boldsymbol{\varphi}}_{n-1}(\mathbf{x}+\mathbf{v}_n^k)\right)-\left(\frac{\partial}{\partial x_1}T\circ\tilde{\boldsymbol{\varphi}}_{n-1}(\mathbf{x}+\mathbf{v}_n^k)\right)^2$, $H_2=T\circ\tilde{\boldsymbol{\varphi}}_{n-1}(\mathbf{x}+\mathbf{v}_n^k)\frac{\partial^2}{\partial x_1\partial x_2}\left(T\circ\tilde{\boldsymbol{\varphi}}_{n-1}(\mathbf{x}+\mathbf{v}_n^k)\right)-\left(\frac{\partial}{\partial x_1}T\circ\tilde{\boldsymbol{\varphi}}_{n-1}(\mathbf{x}+\mathbf{v}_n^k)\right)\left(\frac{\partial}{\partial x_2}T\circ\tilde{\boldsymbol{\varphi}}_{n-1}(\mathbf{x}+\mathbf{v}_n^k)\right)$ and

$H_3 = T\circ\tilde{\boldsymbol{\varphi}}_{n-1}(\mathbf{x}+\mathbf{v}_n^k)\frac{\partial^2}{\partial x_2^2}\left(T\circ\tilde{\boldsymbol{\varphi}}_{n-1}(\mathbf{x}+\mathbf{v}_n^k)\right) - \left(\frac{\partial}{\partial x_2}T\circ\tilde{\boldsymbol{\varphi}}_{n-1}(\mathbf{x}+\mathbf{v}_n^k)\right)^2$. By (1.3) and (3.5), we know $(T\circ\tilde{\boldsymbol{\varphi}}_{n-1}(\mathbf{x}+\mathbf{v}_n^k) - s_{n-1} - \delta s_n^k)^2 \geq (\kappa-\kappa_0)^2$.

Now we give an estimate on the L^∞ norm of H_1. Since $\alpha > 3.5$, by the Sobolev embddding Theorem [9] ($H_0^\alpha(\Omega)\hookrightarrow C^2(\Omega)$), we obtain

$$\begin{aligned}\|\nabla^2\boldsymbol{\varphi}\|^2_{C(\Omega)} = \|\nabla^2\mathbf{u}\|^2_{C(\Omega)} \leq& CR_1(\mathbf{u}) \leq C\lambda_n\int_\Omega (m_{n-1}+\ln D - \ln(T\circ\tilde{\boldsymbol{\varphi}}_{n-1}-s_{n-1}))^2 d\mathbf{x}\\ \leq& 2C\lambda_n|\Omega|(K^2+\ln^2(\overline{M}/\kappa)^2) \leq \lambda_n\widetilde{M},\end{aligned}$$

where $\widetilde{M} \triangleq \widetilde{M}(\Omega,\alpha) = 2C|\Omega|(K^2+\ln^2(\overline{M}/\kappa)^2) > 0$ and $C = C(\Omega,\alpha)$ is a positive constant (see Lemma 3.2 and Lemma 3.3 in [16] for details). Similarly, there holds

$$\|\mathbf{u}_n\|^2_{C^1(\Omega)} \leq \lambda_n\widetilde{M}.$$

Note that

$$\nabla_{\mathbf{x}}\tilde{\boldsymbol{\varphi}}_n(\mathbf{x}) = \nabla_{\mathbf{g}_1}\boldsymbol{\varphi}_0\cdot\nabla_{\mathbf{g}_2}\boldsymbol{\varphi}_1\cdots\nabla_{\mathbf{g}_{n-1}}\boldsymbol{\varphi}_{n-2}\cdot\nabla_{\mathbf{g}_n}\boldsymbol{\varphi}_{n-1}\cdot\nabla_{\mathbf{x}}\boldsymbol{\varphi}_n(\mathbf{x}),$$

where $\mathbf{g}_k = \boldsymbol{\varphi}_k\circ\boldsymbol{\varphi}_2\cdots\boldsymbol{\varphi}_n$ for $k=1,2,\cdots,n$. Since $\mathbf{g}_k$ are mappings from Ω to Ω, by (A), we obtain

$$\|\nabla_{\mathbf{x}}\tilde{\boldsymbol{\varphi}}_n(\mathbf{x})\|^2_{C(\Omega)} \leq \lambda_0\lambda_1\cdots\lambda_n\widetilde{M}^n \leq (\lambda_n\widetilde{M})^n.$$

Then by the chain rule, we have

$$\frac{\partial T\circ\tilde{\boldsymbol{\varphi}}_{n-1}}{\partial x_1} = \frac{\partial T\circ\tilde{\boldsymbol{\varphi}}_{n-1}}{\partial\tilde{\varphi}^1_{n-1}}\frac{\partial\tilde{\varphi}^1_{n-1}}{\partial x_1} + \frac{\partial T\circ\tilde{\boldsymbol{\varphi}}_{n-1}}{\partial\tilde{\varphi}^2_{n-1}}\frac{\partial\tilde{\varphi}^2_{n-1}}{\partial x_1}$$

and

$$\begin{aligned}\frac{\partial^2 T\circ\tilde{\boldsymbol{\varphi}}_{n-1}}{\partial x_1^2} =& \frac{\partial^2 T\circ\tilde{\boldsymbol{\varphi}}_{n-1}}{\partial(\varphi^1_{n-1})^2}\left(\frac{\partial\varphi^1_{n-1}}{\partial x_1}\right)^2 + 2\frac{\partial^2 T\circ\tilde{\boldsymbol{\varphi}}_{n-1}}{\partial\varphi^1_{n-1}\partial\varphi^2_{n-1}}\frac{\partial\tilde{\varphi}^1_{n-1}}{\partial x_1}\frac{\partial\tilde{\varphi}^2_{n-1}}{\partial x_1},\\ &+\frac{\partial T\circ\tilde{\boldsymbol{\varphi}}_{n-1}}{\partial\tilde{\varphi}^2_{n-1}}\frac{\partial^2\tilde{\varphi}^2_{n-1}}{\partial x_1^2} + \frac{\partial^2 T\circ\tilde{\boldsymbol{\varphi}}_{n-1}}{\partial(\varphi^2_{n-1})^2}\left(\frac{\partial\varphi^2_{n-1}}{\partial x_1}\right)^2\\ &+\frac{\partial T\circ\tilde{\boldsymbol{\varphi}}_{n-1}}{\partial\tilde{\varphi}^1_{n-1}}\frac{\partial^2\tilde{\varphi}^1_{n-1}}{\partial x_1^2} + \frac{\partial T\circ\tilde{\boldsymbol{\varphi}}_{n-1}}{\partial\tilde{\varphi}^2_{n-1}}\frac{\partial^2\tilde{\varphi}^2_{n-1}}{\partial x_1^2}.\end{aligned}$$

This concludes

$$\left\|\frac{\partial T\circ\tilde{\boldsymbol{\varphi}}_{n-1}}{\partial x_1}\right\|^2_{C(\Omega)} \leq 4\overline{M}(\lambda_n\widetilde{M})^n$$

and

$$\left\|\frac{\partial^2 T\circ\tilde{\boldsymbol{\varphi}}_{n-1}}{\partial x_1^2}\right\|^2_{C(\Omega)} \leq 6\overline{M}(\lambda_n\widetilde{M})^n.$$

Based on the above mentioned discussion, we can get

$$\|H_1\|_{C(\Omega)} \leq 10\overline{M}^2(\lambda_n\widetilde{M})^n.$$

In addition, we can also obtain the similar estimate for H_2, H_3. Therefore, we conclude

$$\|H(\sigma)\|_{C(\Omega)} \leq \frac{10\overline{M}^2(\lambda_n\widetilde{M})^n}{(\kappa-\kappa_0)^2}.$$

Appendix B. Multigrid method for PDE (3.27). By adopting Grunwald approximation [38], $\frac{\partial^\alpha f(\mathbf{x})}{\partial x_i^\alpha}$, $\frac{\partial^{\alpha*} f(\mathbf{x})}{\partial x_i^{\alpha*}}(i=1,2)$ are discretized as follows:

$$\text{(B.1)} \qquad \frac{\partial^\alpha f(\mathbf{x}_{p,q})}{\partial x_i^\alpha} = \delta_{i-}^\alpha f(\mathbf{x}_{p,q}) + O(h), \quad \frac{\partial^{\alpha*} f(\mathbf{x}_{p,q})}{\partial x_i^{\alpha*}} = \delta_{i+}^\alpha f(\mathbf{x}_{p,q}) + O(h),$$

where $\delta_{1-}^\alpha f_{p,q} = \frac{1}{h^\alpha}\sum_{l=0}^{p+1}\rho_l^{(\alpha)} f_{p-l+1,q}$, $\delta_{1+}^\alpha f_{p,q} = \frac{1}{h^\alpha}\sum_{l=0}^{N-p+2}\rho_l^{(\alpha)} f_{p+l-1,q}$, $\delta_{2-}^\alpha f_{p,q} = \frac{1}{h^\alpha}\sum_{m=0}^{q+1}\rho_m^{(\alpha)} f_{p,q-m+1}$, $\delta_{2+}^\alpha f_{p,q,r} = \frac{1}{h^\alpha}\sum_{m=0}^{N-q+2}\rho_m^{(\alpha)} f_{p,q+m-1}$, and $\rho_l^{(\alpha)}$ is computed by the formula $\rho_0^{(\alpha)} = 1$, $\rho_l^{(\alpha)} = (1-\frac{1+\alpha}{l})\rho_{l-1}^{(\alpha)}$.

Let $U_q = (f_{1,q}, f_{2,q}, \cdots, f_{N,q})^T$, then it follows from (B.1) that

$$\frac{\partial^\alpha U_q}{\partial x_1^\alpha} \approx B_{N,\alpha}U_q, \quad \frac{\partial^{\alpha*} U_q}{\partial x_1^{\alpha*}} \approx B_{N,\alpha}^T U_q,$$

where $\frac{\partial^\alpha U_q}{\partial x_1^\alpha} = \left(\frac{\partial^\alpha f_{1,q}}{\partial x_1^\alpha}, \frac{\partial^\alpha f_{2,q}}{\partial x_1^\alpha}, \cdots, \frac{\partial^\alpha f_{N,q}}{\partial x_1^\alpha}\right)^T$, $\frac{\partial^{\alpha*} U_q}{\partial x_1^{\alpha*}} = \left(\frac{\partial^{\alpha*} f_{1,q}}{\partial x_1^{\alpha*}}, \frac{\partial^{\alpha*} f_{2,q}}{\partial x_1^{\alpha*}}, \cdots, \frac{\partial^{\alpha*} f_{N,q}}{\partial x_1^{\alpha*}}\right)^T$ and

$$B_{N,\alpha} = \frac{1}{h^\alpha}\begin{pmatrix} \rho_1^{(\alpha)} & \rho_0^{(\alpha)} & 0 & \cdots & 0 & 0 \\ \rho_2^{(\alpha)} & \rho_1^{(\alpha)} & \rho_0^{(\alpha)} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \rho_{N-1}^{(\alpha)} & \rho_{N-2}^{(\alpha)} & \rho_{N-3}^{(\alpha)} & \cdots & \rho_1^{(\alpha)} & \rho_0^{(\alpha)} \\ \rho_N^{(\alpha)} & \rho_{N-1}^{(\alpha)} & \rho_{N-2}^{(\alpha)} & \cdots & \rho_2^{(\alpha)} & \rho_1^{(\alpha)} \end{pmatrix}.$$

Hence, we obtain

$$\frac{\partial^{\alpha*}}{\partial x_1^{\alpha*}}\left(\frac{\partial^\alpha U_q}{\partial x_1^\alpha}\right) = B_{N,\alpha}^T B_{N,\alpha}U_q \triangleq A_{N,\alpha}U_q.$$

In a similar way, we obtain the following two approximations for $\frac{\partial^{\alpha*}}{\partial x_2^{\alpha*}}\left(\frac{\partial^\alpha f(\mathbf{x})}{\partial x_2^\alpha}\right)$:

$$\frac{\partial^{\alpha*}}{\partial x_2^{\alpha*}}\left(\frac{\partial^\alpha V_p}{\partial x_2^\alpha}\right) = A_{N,\alpha}V_p,$$

where $V_p = (f_{p,1}, f_{p,2}, \cdots, f_{p,N})^T$. By adopting the Grunwald approximation, $\mathrm{div}^{\alpha *}(\nabla^{\alpha}\mathbf{u}_n^{k+1})$ and $\Delta\mathbf{u}_n^{k+1}$ are approximated by the following two formulas:

$$\left(\mathrm{div}^{\alpha *}(\nabla^{\alpha} u_{n,\beta}^{k+1})\right)_{p,q} \approx \sum_{l=0}^{N}\left(a_{N,\alpha}(p,l)(u_{n,\beta}^{k+1})_{l,q} + a_{N,\alpha}(q,l)(u_{n,\beta}^{k+1})_{p,l}\right) \tag{B.2}$$

and

$$\left(\Delta u_{n,\beta}^{k+1}\right)_{p,q} \approx \frac{1}{h^2}((u_{n,\beta}^{k+1})_{p+1,q} + (u_{n,\beta}^{k+1})_{p-1,q} + (u_{n,\beta}^{k+1})_{p,q+1} + (u_{n,\beta}^{k+1})_{p,q-1} - 4(u_{n,\beta}^{k+1})_{p,q}), \tag{B.3}$$

where $\beta = 1, 2$. Based on (B.2) and (B.3), (3.27) is discretized as follows:

$$\begin{aligned} &(1 + 4\gamma_n + 2\theta_n\mu(a_{N,\alpha}(p,p) + a_{N,\alpha}(q,q)))(u_{n,\beta}^{k+1})_{p,q} \\ &+2\theta_n\mu \sum_{l=1, l\neq p,q}^{N_S}\left(a_{N,\alpha}(p,l)(u_{n,\beta}^{k+1})_{l,q} + a_{N,\alpha}(q,l)(u_{n,\beta}^{k+1})_{p,l}\right) \\ &-\gamma_n\left((u_{n,\beta}^{k+1})_{p+1,q} + (u_{n,\beta}^{k+1})_{p-1,q} + (u_{n,\beta}^{k+1})_{p,q+1} + (u_{n,\beta}^{k+1})_{p,q-1}\right) = (v_{n,\beta}^{k+1})_{p,q}, \end{aligned} \tag{B.4}$$

where $\gamma_n = \frac{2\theta_n\Theta}{h^2}$. Then (B.4) induces the following solver for (3.27):

$$\begin{aligned} (u_{n,\beta}^{k+1})_{p,q}^{(t+1)} =& \frac{1}{\Upsilon_n}\Bigg((v_{n,\beta}^{k+1})_{p,q} - 2\theta_n\mu \sum_{l=1, l\neq p,q}^{N_S}\left(a_{N,\alpha}(p,l)(u_{n,\beta}^{k+1})_{l,q}^{(t)} + a_{N,\alpha}(q,l)(u_{n,\beta}^{k+1})_{p,l}^{(t)}\right) \\ &+ \gamma_n\left((u_{n,\beta}^{k+1})_{p+1,q}^{(t)} + (u_{n,\beta}^{k+1})_{p-1,q}^{(t)} + (u_{n,\beta}^{k+1})_{p,q+1}^{(t)} + (u_{n,\beta}^{k+1})_{p,q-1}^{(t)}\right)\Bigg), \end{aligned} \tag{B.5}$$

where $\Upsilon_n = 1 + 4\gamma_n + 2\theta_n\mu(a_{N,\alpha}(p,p) + a_{N,\alpha}(q,q))$ and $t = 0, 1, 2, \cdots$. Hence, based on (B.5), the multigrid algorithm for (3.27) can be summarized in Algorithm B.1.

Algorithm B.1 2D multigrid algorithm for **u**-problem

Initialization: $\mathbf{u}_n^{k+1,h} = \mathbf{u}_n^{k,h}$, $\mathbf{u}_{n,0}^{k+1,h} = \mathbf{u}_n^{k,h} + \mathbf{\Pi}$, $\mu > 0$, $\bar{k} = 0$ and maximum iteration times K.

while $\|\mathbf{u}_n^{k+1,h} - \mathbf{u}_{n,0}^{k+1,h}\| \geq \|\mathbf{\Pi}\|$ and $\bar{k} \leq K$

$\mathbf{u}_{n,0}^{k+1,h} = \mathbf{u}_n^{k+1,h}$.

Step 1. relax (B.5) with initial guess $\mathbf{u}_n^{k+1,h}$; compute residual error $\mathbf{r}_n^{k+1,h}$ on Ω^h;
Set $level = L$;

Step 2. restrict the residual error to Ω^H by using $\mathbf{r}_n^{k+1,H} = R_h^H \mathbf{r}_n^{k+1,h}$.
Set $level = level - 1$, $H = 2h$, and relax (B.5) by replacing $\mathbf{v}_n^{k+1}$ with $\mathbf{r}_n^{k+1,H}$, and with initial guess $\mathbf{u}_n^{k+1,H} = \mathbf{0}$ to obtain approximations $\bar{\mathbf{u}}_n^{k+1,H}$; update residual error $\mathbf{r}_n^{k+1,H}$.

Step 3. **If** $level = 1$,

do: accurately solve the system (B.5) by replacing $\mathbf{v}_n^{k+1}$ with $\mathbf{r}_n^{k+1,H}$ to obtain the solution $\mathbf{u}_n^{k+1,H}$;

else

do: repeat **Step 2** until $level = 1$.

endif.

Step 4. **If** $level = L$,

do: relax (B.5) to obtain the final solution $\mathbf{u}_n^{k+1,h}$ for this round and let $\bar{k} = \bar{k} + 1$;

else

do(repeat): interpolate the correction to next fine grid by letting $\mathbf{u}_{n,t}^{k+1,h} = I_H^h \mathbf{u}_n^{k+1,H}$; update current grid approximations using correction $\hat{\mathbf{u}}_n^{k+1,h} = \mathbf{u}_{n,t}^{k+1,h} + \bar{\mathbf{u}}_n^{k+1,h}$; relax (B.5) with initial guess $\hat{\mathbf{u}}_n^{k+1,h}$ on fine grid to obtain approximations $\mathbf{u}_n^{k+1,h}$ and let $level = level + 1$. Repeat this process until $level = L$.

endif.

Set $\bar{k} = \bar{k} + 1$;

endwhile

Output: $\mathbf{u}_n^{k+1} = \mathbf{u}_n^{k+1,h}$.