Convergence analysis of a variational quasi-reversibility approach for an inverse hyperbolic heat conduction problem

Vo Anh Khoa and Manh-Khang Dao

Abstract. We study a time-reversed hyperbolic heat conduction problem based upon the Maxwell–Cattaneo model of non-Fourier heat law. This heat and mass diffusion problem is a hyperbolic type equation for thermodynamics systems with thermal memory or with finite time-delayed heat flux, where the Fourier or Fick law is proven to be unsuccessful with experimental data. In this work, we show that our recent variational quasi-reversibility method for the classical time-reversed heat conduction problem, which obeys the Fourier or Fick law, can be adapted to cope with this hyperbolic scenario. We establish a generic regularization scheme in the sense that we perturb both spatial operators involved in the PDE. Driven by a Carleman weight function, we exploit the natural energy method to prove the well-posedness of this regularized scheme. Moreover, we prove the Hölder rate of convergence in the mixed \( L^2 - H^1 \) spaces. Under some certain choice of the perturbations and stabilizations, we thereupon obtain the Lipschitz rate in \( L^2 \). We also show that under a weaker conditional estimate, it is sufficient to perturb only the highest order differential operator to gain the Hölder convergence.

Keywords. Backward heat conduction problem, hyperbolic equation, quasi-reversibility method, energy estimates, Carleman weight, Hölder–Lipschitz convergence.

2010 Mathematics Subject Classification. 65L70, 65L09, 65L60.

1 Introduction

1.1 Statement of the inverse problem

In this work, we are interested in the extension of our new quasi-reversibility (QR) method in [20] for terminal boundary value problems. In this regard, we want...
to recover the initial distribution of an evolutionary equation, given the terminal
data. This model is well known to be one of the classical problems in the field
of inverse and ill-posed problems; cf. e.g. [11] for some background of typical
models in this research line. As to the applications of this model, having a reliable
stable approximation of this backward-in-time problem is significantly helpful
in many physical, biological and ecological contexts. Those are concretely involved
in, e.g., the works [4, 9, 21, 24]. In particular, the first contribution of this model be-
ing in mind relies on the heating/cooling transfer problem based upon the fact that
sometimes, we want to measure the initial temperature of a material and our equip-
ment only works at a given later time. Recently, this scenario has been extended
to the case of a two-slab composite system with an ideal transmission condition in
[24]. The second application we would like to address here is recovering blurry
digital images acquired by camera sensors. This practical concern was initiated
in [4] and has been scrutinized in the framework of source localization for brain
tumor in [9]. In mathematical oncology, reconstructing the initial images of the
tumor can be used for analyzing behaviors of cancer cells and then potentially for
predicting the progression of neoplasms of early-stage patients. This initial recon-
struction is also part of the so-called data assimilation procedure that has been of
interest so far in weather forecasting (cf. [1]).

It is worth mentioning that considerations of such parabolic models indicate the
use of the Fourier or Fick law. However, in some contexts of thermodynamics
this typical law is proven to be unsuccessful with experimental data. In fact, any
initial disturbance in a medium is propagated instantly when taking into account
the parabolic case; cf. e.g. [5]. We also refer to the monograph [10] and some
impressive works [19, 26], where some electromechanical models were studied
to unveil this non-standard incompatibility. In order to avoid the phenomenon of
infinite propagation, the Cattaneo–Vernotte law was derived, proposing that the
parabolic case should be upgraded to a hyperbolic form. In terms of PDEs, it
means one should consider

\[ u_{tt} + u_t - \Delta u = 0 \quad \text{in } \Omega \times (0, T) , \tag{1} \]

where \( T > 0 \) is the final time and \( \Omega \subset \mathbb{R}^d \) \((d = 1, 2, 3)\) is a regular bounded do-
main of interest with a sufficiently smooth boundary. In electrodynamics, equation
(1) is the same as the telegrapher’s equation derived from the Maxwell equation.
That is why one usually refers (1) to as the Maxwell–Cattaneo model.

In this work, we investigate a generalized model of (1) due to our mathematical
interest. We assume to look for \( u(x, t) : \Omega \times (0, T) \to \mathbb{R} \) satisfying the following evolution-
ary equation:

\[ u_{tt} + u_t - \Delta u - \Delta u_t = 0 \quad \text{in } \Omega \times (0, T) . \tag{2} \]
In the studies of the motion of viscoelastic materials, this is well-known to be the linear strongly damped wave equation, where the weak and strong damping terms \( (u_t \text{ and } -\Delta u_t) \) are altogether involved in the PDE. Cf. [3] and references cited therein, the solution \( u \) in that setting can be viewed as a displacement, whilst it is a temperature field in the context of thermodynamics we have mentioned above.

Going back to the heat context, we note that the underlying equation (2) is also related to the so-called Gurtin–Pipkin model, which reads as

\[
\theta_t = \int_0^t \kappa(t-s)\theta_{xx}(s)ds. \tag{3}
\]

When the kernel \( \kappa \) is a constant, (3) becomes an integrated wave equation after differentiation in time. If \( \kappa(t) = e^{-t} \), one has the weakly damped wave equation \( u_{tt} + u_t - u_{xx} = 0 \). Furthermore, when \( \kappa(t) = \delta(t) \), we get back to the classical heat equation. Therefore, we can conclude that our mathematical analysis for (2) really works for many distinctive physical applications at the same time.

To complete the time-reversed model, we endow (2) with the following boundary and terminal conditions:

\[
\begin{align*}
\{u(x,t) &= 0 \quad \text{on } \partial\Omega \times (0,T), \\
u(x,T) &= f_0(x), \quad u_t(x,T) = f_1(x) \quad \text{in } \Omega.
\end{align*} \tag{4}
\]

Hence, (2) and (4) form our terminal boundary value problem. As to the ill-posedness of this problem, we refer to [25] for proof of its natural instability using the spectral approach.

### 1.2 Historical remarks and contributions of the paper

In the context of the time-reversed parabolic problem, many regularization schemes were extensively designed in order to circumvent its natural ill-posedness. Inverse problems for parabolic equations with memory effects were investigated in [2, 18]. Since the aim of this work is extending our new QR method in [20] to the hyperbolic heat conduction scenario, we would like to address some existing literature just on the QR topic close to the explicit technique we are developing. Meanwhile, some implicit QR methods for the backward heat conduction problem can be referred to the works [6, 7, 16, 17]. The “implicit” here means that the scheme is designed by perturbing the kernel of the unbounded operator itself. Another QR-based approaches using minimization were studied in e.g. [13, 14].

The very first idea about quasi-reversibility of time-reversed parabolic problems was established by Lattès and Lions in the monograph [15] when they used a fourth-order spatial perturbation to stabilize the Laplace operator involved in the
classical time-reversed parabolic equation. Motivated by this approach, several modifications and variants were constructed and analyzed through five decades, which makes this method considerable in the field of inverse and ill-posed problems. For example, we mention here the pioneering work [22], where a third-order operator in space and time was proposed to obtain a regularization scheme in the form of a pseudoparabolic equation. Recently, Kaltenbacher et al. [12] has used a nonlocal perturbing operator in time with fractional order to regularize the ill-posed problem.

Our newly developed QR method follows the original idea of Lattès and Lions, i.e. we only use the spatial perturbation to stabilize the unbounded spatial operator. The key ingredient of our method lies in the fact that we use the perturbation operator to turn the inverse problem into a forward-like problem involving the stabilized operator. This notion has been studied in a spectral form in our recent work [23]. As a follow-up, we generalize this method in [20] by the establishment of conditional estimates for both the perturbation and stabilized operators. Driven by a Carleman weight function, we further apply the conventional energy method to show both well-posedness of the regularized system and error bounds. This way allows us to derive the scheme in the finite element setting and prove the error estimates in the finite-dimensional space. This will be our next target work in the future.

This work is the first time we extend our new method to the ill-posed problem (2) and (4). Intuitively, we construct in section 2 a generic regularized system in the sense that we perturb all the spatial terms $-\Delta u$ and $-\Delta u_t$. We then use the conditional estimates established in [20] to obtain the Hölder rate of convergence in section 4. Besides, well-posedness of the regularized system is considered in section 3 using a priori estimates and compactness arguments. Section 5 is then devoted to the following improvements. First, under a weaker conditional estimate we show that perturbing the term $-\Delta u_t$ is sufficient to gain the Hölder convergence. Second, we propose a modified generic scheme to obtain a Lipschitz rate of convergence under a special choice of the involved perturbations. All choices of the perturbing and stabilized operators are established with relevance to mathematical analysis of the direct problem. Well-posedness of the forward problem of (2) was already proven in [8].

2 A variational quasi-reversibility framework

To this end, $\langle \cdot, \cdot \rangle$ indicates either the scalar product in $L^2(\Omega)$ or the dual pairing of a continuous linear functional and an element of a function space. Also, $\|\cdot\|$ is the norm in $L^2(\Omega)$. Different inner products and norms should be written as
\langle \cdot, \cdot \rangle_X \text{ and } \| \cdot \|_X, \text{ respectively, where } X \text{ is a certain Banach space. In the sequel, we denote } \varepsilon \in (0, 1) \text{ by the noise level of the terminal data } f_0, f_1 \text{ in (4). Any constant } C > 0 \text{ may vary from line to line. We usually indicate its dependencies if necessary.}

We introduce an auxiliary function \( \gamma := \gamma(\varepsilon) \geq 1 \) satisfying \( \lim_{\varepsilon \to 0} \gamma(\varepsilon) = \infty \).

**Definition 2.1** (perturbing operator). The linear mapping \( Q_\varepsilon : L^2(\Omega) \to L^2(\Omega) \) is said to be a perturbing operator if there exist a function space \( \mathcal{W} \subset L^2(\Omega) \) and a noise-independent constant \( C_0 > 0 \) such that

\[
\| Q_\varepsilon u \| \leq C_0 \| u \|_{\mathcal{W}} / \gamma(\varepsilon) \quad \text{for any } u \in \mathcal{W}.
\]

**Definition 2.2** (stabilized operator). The linear mapping \( P_\varepsilon : L^2(\Omega) \to L^2(\Omega) \) is said to be a stabilized operator if there exists a noise-independent constant \( C_1 > 0 \) such that

\[
\| P_\varepsilon u \| \leq C_1 \log(\gamma(\varepsilon)) \| u \| \quad \text{for any } u \in L^2(\Omega).
\]

In this work, we start off with the generic approach of this modified version by stabilizing both two terms \( -\Delta u \) and \( -\Delta u_t \). In this sense, we add the perturbations \( -Q_1^\varepsilon \) and \( -Q_2^\varepsilon \) for these operators, respectively. In particular, we are led to the following type of stabilizations \( P_1^\varepsilon = 2\Delta + Q_1^\varepsilon \) and \( P_2^\varepsilon = 2\Delta + Q_2^\varepsilon \). Since both \( Q_1^\varepsilon \) and \( Q_2^\varepsilon \) should altogether satisfy the conditional estimate (5), we can take \( Q_1^\varepsilon = Q_2^\varepsilon \) for simplicity. Therefore, we assume the same stabilizations \( P_1^\varepsilon = P_2^\varepsilon \) are applied in the following regularized equation:

\[
u_{tt}^\varepsilon + u_t^\varepsilon + \Delta u^\varepsilon + \Delta u_t^\varepsilon = P_1^\varepsilon u^\varepsilon + P_2^\varepsilon u_t^\varepsilon \quad \text{in } \Omega \times (0, T).
\]

Since in real-world applications the terminal data are usually noisy, we endow (7) with the following boundary and terminal conditions:

\[
\begin{cases}
u^\varepsilon(x,t) = 0 & \text{on } \partial \Omega \times (0, T), \\
u^\varepsilon(x, T) = f_0^\varepsilon(x), u_t^\varepsilon(x, T) = f_1^\varepsilon(x) & \text{in } \Omega.
\end{cases}
\]

In (8), we assume to have a noise level \( \varepsilon \in (0, 1) \) such that

\[
\| u^\varepsilon(\cdot, T) - u(\cdot, T) \|_{H^1(\Omega)} + \| u_t^\varepsilon(\cdot, T) - u_t(\cdot, T) \| \leq \varepsilon.
\]

To validate our mathematical analysis below, we suppose that \( f_0, f_0^\varepsilon \in H^1(\Omega) \) and \( f_1, f_1^\varepsilon \in L^2(\Omega) \).
3 Well-posedness of the regularized system (7)–(8)

Let $v^{\varepsilon}(x, t) := e^{\rho(t-T)}u^{\varepsilon}(x, t)$ where $\rho > 1$ is a constant chosen later, then (7)-(8) become

$$v^{\varepsilon}_{tt} + (1-2\rho)v^{\varepsilon}_t + (\rho^2 - \rho)v^{\varepsilon} + (1-\rho)\Delta v^{\varepsilon} + \Delta v^{\varepsilon}_{\varepsilon} = (1-\rho)P^1_{\varepsilon} v^{\varepsilon} + P^2_{\varepsilon} v^{\varepsilon}_t$$

in $\Omega \times (0, T)$

(10)

and the boundary and terminal conditions:

$$\begin{cases}
v^{\varepsilon}(x, t) = 0 & \text{on } \partial\Omega \times (0, T), \\
v^{\varepsilon}(x, T) = f^{\varepsilon}_0(x), & v^{\varepsilon}_t(x, T) = \rho f^{\varepsilon}_0(x) + f^{\varepsilon}_1(x) \quad \text{in } \Omega.
\end{cases}
$$

(11)

Remark 3.1. The most important difficult need to solve the regularized system (7)–(8) lies in the term $+\Delta u^\varepsilon$, which is bad for our energy estimations for $u^\varepsilon$. More precisely, the sign of this term is technically impeding the energy of the gradient term and eventually, it ruins our mathematical analysis in this section. In order to circumvent this, we consider the system (10)–(11) for $v^\varepsilon$, which is equivalent to the regularized system (7)–(8). Since $\rho > 1$, then $(1-\rho)\Delta u^\varepsilon$ becomes a “good term” and we shall use its effect to obtain the energy estimate for $v^\varepsilon$ in Theorem 3.6. This leads us to the well-posedness of (10)–(11) as well as that of (7)–(8).

Definition 3.2. A function $v \in L^2(0, T; H^1_0(\Omega))$ with $v_t \in L^2(0, T; H^1(\Omega))$ and $v_{tt} \in L^2(0, T; H^{-1}(\Omega))$ is a weak solution of (10)–(11) if for every test function $\varphi \in H^1_0(\Omega)$, it holds that

$$\langle v_{tt}(t), \varphi \rangle_{H^{-1},H^1_0} + (1-2\rho)\langle v_t(t), \varphi \rangle + (\rho^2 - \rho)\langle v(t), \varphi \rangle$$

$$+ (\rho - 1)\langle \nabla v(t), \nabla \varphi \rangle - \langle \Delta v(t), \nabla \varphi \rangle = (1-\rho)\langle P^1_{\varepsilon} v(t), \varphi \rangle + \langle P^2_{\varepsilon} v_t(t), \varphi \rangle$$

for a.e. $t \in (0, T)$, and $v(x, T) = f^\varepsilon_0(x)$, $v_t(x, T) = \rho f^\varepsilon_0(x) + f^\varepsilon_1(x)$ in $\Omega$.

Our proof of well-posedness relies on the conventional Galerkin method. This means that we construct solution of some finite-dimensional approximations to (12). When doing so, we need to recall some auxiliary results.

Remark 3.3. By the standard Fredholm theory, there exist

- a non decreasing sequence of nonegative real numbers $\{\mu_k\}_{k=1}^\infty$ that tends to $+\infty$ as $k \to \infty$,

- a Hilbert basis $\{\phi_k\}_{k=1}^\infty$ of $L^2(\Omega)$ such that $\phi_k \in H^1_0(\Omega)$ such that

$$\int_{\Omega} \nabla \phi_k \cdot \nabla \phi dx = \mu_k \int_{\Omega} \phi_k \phi dx \quad \text{for all } \phi \in H^1_0(\Omega).$$
Lemma 3.4. For any positive $n$, there exist $n$ absolutely continuous functions $y_k^n : [0, T] \to \mathbb{R}$, $k = 1, \ldots, n$ and a function $v_n \in L^2(0, T; H^0_0(\Omega))$, where $\partial_t v_n \in L^2(0, T; H^1(\Omega))$ and $\partial_{tt} v_n \in L^2(0, T; H^{-1}(\Omega))$, of the form

$$v_n(x, t) = \sum_{k=1}^{n} y_k^n(t) \phi_k(x),$$  \hspace{1cm} (13)

such that for $k = 1, \ldots, N$

$$
\begin{align*}
\left\{
\begin{array}{l}
y_k^n(T) = \int_{\Omega} f_0^x(x) \phi_k(x) dx =: g_{0k}(T), \\
\partial_t y_k^n(T) = \int_{\Omega} (\rho f_0^x(x) + f_1^x(x)) \phi_k(x) dx =: g_{1k}(T),
\end{array}
\right.
\end{align*}$$  \hspace{1cm} (14)

and $v_n$ satisfies

$$
\begin{align*}
&\int_{\Omega} \partial_{tt} v_n(t) \phi_k dx + (1 - 2\rho) \int_{\Omega} \partial_t v_n(t) \phi_k dx + (\rho^2 - \rho) \int_{\Omega} v_n(t) \phi_k dx \\
&\quad + (\rho - 1) \int_{\Omega} \nabla v_n(t) \cdot \nabla \phi_k dx - \int_{\Omega} \nabla \partial_t v_n(t) \cdot \nabla \phi_k dx \\
&\quad = (1 - \rho) \int_{\Omega} P_1^x v(t) \phi_k dx + \int_{\Omega} P_2^x \partial_t v(t) \phi_k dx.
\end{align*}$$  \hspace{1cm} (15)

\textbf{Proof.} By the properties of $\{\phi_i\}_{i=1}^{\infty}$ in Remark 3.3, (15) is equivalent to

$$
\begin{align*}
\partial_{tt} y_k^n(t) + (1 - 2\rho - \mu_k) \partial_t y_k^n(t) + (\rho^2 + (\mu_k - 1)\rho - \mu_k) y_k^n(t) \\
= (1 - \rho) \sum_{i=0}^{n} y_i^n(t) \langle P_1^x \phi_i, \phi_k \rangle + \sum_{i=0}^{n} \partial_t y_i^n(t) \langle P_2^x \phi_i, \phi_k \rangle \text{ for a.e. } t \in (0, T).
\end{align*}$$  \hspace{1cm} (16)

Let $z_k^n = \frac{d}{dt} y_k^n$, it follows from (14) and (16) that

$$
\frac{d}{dt} \begin{bmatrix} y_k^n \\ z_k^n \end{bmatrix} + A_k \begin{bmatrix} y_k^n \\ z_k^n \end{bmatrix} = F_k, \hspace{1cm} \begin{bmatrix} y_k^n(T) \\ z_k^n(T) \end{bmatrix} = \begin{bmatrix} g_{0k}(T) \\ g_{1k}(T) \end{bmatrix},
$$

where $F_k = [0, (1 - \rho) \sum_{i=0}^{n} y_i^n \langle P_1^x \phi_i, \phi_k \rangle + \sum_{i=0}^{n} z_k^n \langle P_2^x \phi_i, \phi_k \rangle]^{T}$ and

$$A_k = \begin{bmatrix} 0 & 1 \\
\rho^2 + (\mu_k - 1)\rho - \mu_k & 1 - 2\rho - \mu_k \end{bmatrix}.$$  \hspace{1cm}

Consider $w_k^n := [y_k^n, z_k^n]^{T}$. We thus obtain the following integral equation:

$$w_k^n(t) = w_k^n(T) + A_k \int_{t}^{T} w_k^n(s) ds - \int_{0}^{t} F_k(s) ds. \hspace{1cm} (17)$$
Hereafter, we denote by \( w_n := [w_1^n, \ldots, w_n^n] : [0, T] \to \mathbb{R}^{2n} \). The integral equation (17) can be rewritten as \( w_n = H[w_n] \), where the same notation as \( w_n \) is applied to \( H \) with \( H_k \) being the right-hand side of (17). Define the norm in \( Y = C([0, T]; \mathbb{R}^{2n}) \) as follows:

\[
\|c\|_Y := \sup_{t \in [0, T]} \sum_{j=1}^n |c_j(t)| \quad \text{with} \quad c = [c_j] \in C([0, T]; \mathbb{R}^{2n}).
\]

We claim that there exists \( n_0 \in \mathbb{N}^* \) such that the operator \( H^{n_0} := H[H^{n_0-1}] : Y \to Y \) is a contraction mapping. In other words, we find \( K \in [0, 1) \) such that

\[
\|H^{n_0}[w_n] - H^{n_0}[\tilde{w}_n]\|_Y \leq K \|w_n - \tilde{w}_n\|_Y \quad \text{for any} \quad w_n, \tilde{w}_n \in Y.
\]

This can be done by induction. Indeed, let us observe that

\[
|H_k[w_n](t) - H_k[\tilde{w}_n](t)| \leq \int_t^T |A_k||w_k^n(s) - \tilde{w}_k^n(s)|ds
\]

\[
\quad + \int_t^T \left( C_1 C \log_1(\gamma) \sum_{i=1}^n \left( |1 - \rho||y_i^n(s) - \tilde{y}_i^n(s)| + |z_i^n(s) - \tilde{z}_i^n(s)| \right) \right) ds
\]

\[
\quad \leq \int_t^T (|A_k| + C_1 C \log_1(\gamma)(\rho - 1)) |w_k^n(s) - \tilde{w}_k^n(s)|ds
\]

\[
\quad \leq (|A_k| + C_1 C \log_1(\gamma)(\rho - 1)) (T - t) \|w_n - \tilde{w}_n\|_Y,
\]

aided by the conditional estimate (6). Here, we indicate \( C = \max_i C(\|\phi_i\|_{H^1_0(\Omega)}) > 0 \). Furthermore, for any \( m \in \mathbb{N}^* \)

\[
|H_k^m[w_n](t) - H_k^m[\tilde{w}_n](t)|
\]

\[
\quad \leq \int_t^T (|A_k| + C_1 C \log_1(\gamma)(\rho - 1)) |H^{m-1}[w_n](s) - H^{m-1}[\tilde{w}_n](s)|ds,
\]

and it follows by induction that

\[
|H_k^m[w_n](t) - H_k^m[\tilde{w}_n](t)|
\]

\[
\quad \leq (|A_k| + C_1 C \log_1(\gamma)(\rho - 1))^m \frac{(T - t)^m}{m!} \|w_n - \tilde{w}_n\|_Y.
\]
Therefore, we obtain

\[ |H^m[w_n](t) - H^m[\tilde{w}_n](t)| \leq \frac{||w_n - \tilde{w}_n||_Y}{m!} \sum_{k=1}^{n} (|A_k| + C_1 C \log(\gamma)(\rho - 1))^m. \]

Since the left-hand side tends to 0 as \( m \to \infty \), we can find a sufficiently large \( n_0 \) such that

\[ \frac{1}{n_0!} \sum_{k=1}^{n_0} (|A_k| + C_1 C \log(\gamma)(\rho - 1))^{n_0} < 1. \]

The claim is proved and by the Banach fixed-point argument, there exists a unique solution \( \tilde{w}_n \in Y \) such that \( H^{n_0}[\tilde{w}_n] = \tilde{w}_n \). Since \( H^{n_0}[H[\tilde{w}_n]] = H[H^{n_0}[\tilde{w}_n]] = H[\tilde{w}_n] \), then the integral equation (17) admits a unique solution in \( Y \). Hence, we complete the proof of the lemma.

\[ \square \]

**Remark 3.5.** By Lemma 3.4, it is easy to check that there exists a constant \( C > 0 \) such that

\[ ||\partial_t v_n^\varepsilon(T)||^2, ||v_n^\varepsilon(T)||^2, ||\nabla v_n^\varepsilon(T)||^2 \leq C \quad \text{for all } n \in \mathbb{N}. \]  

(18)

**Theorem 3.6.** Assume (9) holds. For each \( \varepsilon > 0 \), the regularized system (10)–(11) admits a weak solution \( v^\varepsilon \) in the sense of Definition 3.2.

**Proof.** To prove this theorem, we need to derive some energy estimates for approximate solution \( v_n^\varepsilon \). Thanks to Lemma 3.4, we have \( \partial_t v_n^\varepsilon \in C([0, 1]; H^1(\Omega)) \). Multiplying (15) by \( \partial_t y_k^n(t) \), summing for \( k = 1, \ldots, N \) and using the formula (13) for \( v_n^\varepsilon \), we get

\[
\int_{\Omega} \partial_{tt} v_n^\varepsilon(t) \partial_t v_n^\varepsilon(t) dx + (1 - 2\rho) \int_{\Omega} |\partial_t v_n^\varepsilon(t)|^2 dx + (\rho^2 - \rho) \int_{\Omega} v_n^\varepsilon(t) \partial_t v_n^\varepsilon(t) dx \\
+ (\rho - 1) \int_{\Omega} \nabla v_n^\varepsilon(t) \cdot \nabla \partial_t v_n^\varepsilon(t) dx - \int_{\Omega} |\nabla \partial_t v_n^\varepsilon(t)|^2 dx \\
= (1 - \rho) \int_{\Omega} P_1^\varepsilon(v_n^\varepsilon(t)) \partial_t v_n^\varepsilon(t) dx + \int_{\Omega} P_2^\varepsilon(\partial_t v_n^\varepsilon(t)) \partial_t v_n^\varepsilon(t) dx.
\]
This implies

\[
\frac{1}{2} \partial_t \left[ \| \partial_t v_n^\varepsilon(t) \|^2 + (\rho^2 - \rho) \| v_n^\varepsilon(t) \|^2 + (\rho - 1) \| \nabla v_n^\varepsilon(t) \|^2 \right] \\
- (2\rho - 1) \int _\Omega |\partial_t v_n^\varepsilon(t)|^2 dx - \int _\Omega |\nabla \partial_t v_n^\varepsilon(t)|^2 dx \\
= (1 - \rho) \int _\Omega P_\varepsilon^1(\partial_t v_n^\varepsilon(t)) \partial_t v_n^\varepsilon(t) dx + \int _\Omega P_\varepsilon^2(\partial_t v_n^\varepsilon(t)) \partial_t v_n^\varepsilon(t) dx \\
\geq (1 - \rho) C_1 \log(\gamma) \left( \| v_n^\varepsilon(t) \|^2_{H^1(\Omega)} + \| \partial_t v_n^\varepsilon(t) \|^2 \right) - C_1 \log(\gamma) \| \partial_t v_n^\varepsilon(t) \|^2,
\]

(19)

where the last inequality comes from the Hölder inequality and (6).

*Estimate* \( v_n^\varepsilon \) in \( L^\infty(0, T; H^1(\Omega)) \) and \( \partial_t v_n^\varepsilon \) in \( L^\infty(0, T; L^2(\Omega)) \). It follows from (19) that

\[
\partial_t \left( \frac{\| \partial_t v_n^\varepsilon(t) \|^2}{\rho - 1} + \rho \| v_n^\varepsilon(t) \|^2 + \| \nabla v_n^\varepsilon(t) \|^2 \right) \\
\geq 2C_1 \log(\gamma) \rho \left( \frac{\| \partial_t v_n^\varepsilon(t) \|^2}{\rho - 1} + \rho \| v_n^\varepsilon(t) \|^2 + \| \nabla v_n^\varepsilon(t) \|^2 \right),
\]

By Grönwall’s inequality, we get

\[
\frac{\| \partial_t v_n^\varepsilon(t) \|^2}{\rho - 1} + \rho \| v_n^\varepsilon(t) \|^2 + \| \nabla v_n^\varepsilon(t) \|^2 \\
\leq \left( \frac{\| \partial_t v_n^\varepsilon(T) \|^2}{\rho - 1} + \rho \| v_n^\varepsilon(T) \|^2 + \| \nabla v_n^\varepsilon(T) \|^2 \right) \gamma^2 C_1 \rho(T - t).
\]

(20)

From (18), one gets

\[
\begin{aligned}
\text{if } \partial_t v_n^\varepsilon \text{ is uniformly bounded in } L^\infty(0, T; L^2(\Omega)), \\
v_n^\varepsilon \text{ is uniformly bounded in } L^\infty(0, T; H^1(\Omega)).
\end{aligned}
\]

(21)

It follows from the Banach–Alaoglu theorem, and the argument that a weak limit of derivative is the derivative of the weak limit, that we can extract a subsequence of scaled approximate solutions \( v_n^\varepsilon \), which we still denote by \( \{ v_n^\varepsilon \}_{n \in \mathbb{N}} \), such that for each \( \varepsilon > 0 \)

\[
\begin{aligned}
\text{if } \partial_t v_n^\varepsilon \to \partial_t v^\varepsilon \text{ weakly } - * \text{ in } L^\infty(0, T; L^2(\Omega)), \\
v_n^\varepsilon \to v^\varepsilon \text{ weakly } - * \text{ in } L^\infty(0, T; H^1(\Omega)).
\end{aligned}
\]

(22)
Estimate \( \partial_t v_n^\varepsilon \) in \( L^2(0, T; H^1_0(\Omega)) \). Integrating both sides of (19) from 0 to \( T \), we get
\[
(2\rho - 1) \| \partial_t v_n^\varepsilon \|^2_{L^2(0, T; L^2(\Omega))} + \| \nabla v_n^\varepsilon \|^2_{L^2(0, T; L^2(\Omega))} \leq C_1 \log(\gamma) \left( (\rho - 1) \| v_n^\varepsilon \|^2_{L^2(0, T; H^1_0(\Omega))} + \rho \| \partial_t v_n^\varepsilon \|^2_{L^2(0, T; L^2(\Omega))} \right)
\]
\[+ \frac{1}{2} \left[ \| \partial_t v_n^\varepsilon (T) \|^2 + (\rho^2 - \rho) \| v_n^\varepsilon (T) \|^2 + (\rho - 1) \| \nabla v_n^\varepsilon (T) \|^2 \right].
\]
From (18) and (21), it is straightforward to see that
\[
\| \partial_t v_n^\varepsilon \|_{L^2(0, T; H^1_0(\Omega))} \leq \tilde{C} \quad \text{for all } n \in \mathbb{N},
\]
for some constant \( \tilde{C} \).

Estimate \( \partial_t v_n^\varepsilon \) in \( L^2(0, T; H^{-1}(\Omega)) \). Let \( S_n \) be a closed subspace of \( H^1_0(\Omega) \) defined by \( S_n \) = \{ \varphi \in H^1_0(\Omega) : \int_\Omega \varphi \varphi_k \, dx = 0 \text{ for all } k \leq n \}. \) Let \( S_n^{\perp} \) be a closed subspace of \( H^1_0(\Omega) \) such that \( H^1_0(\Omega) = S_n \oplus S_n^{\perp} \). In other words, for all \( \varphi \in H^1_0(\Omega) \), we can write \( \varphi \) of the form \( \varphi = \varphi_n + \varphi_n^{\perp} \) where \( \varphi \in S_n \) and \( \varphi_n^{\perp} \in S_n^{\perp} \). Therefore, for a.e. \( t \in [0, T] \), from (15), one gets
\[
\langle \partial_t v_n^\varepsilon (t), \varphi \rangle = (2\rho - 1) \langle \partial_t v_n^\varepsilon (t), \varphi_n \rangle + (\rho^2 - \rho) \| v_n^\varepsilon (t) \| \| \varphi_n \| + (\rho - 1) \| \nabla v_n^\varepsilon (t) \| \| \nabla \varphi_n \| + \| \nabla \varphi_n \| \| \varphi_n \| + \| \varphi_n^{\perp} \| \| \varphi_n^{\perp} \|.
\]
Since \( \| \varphi_n \|_{H^1_0(\Omega)} \leq \| \varphi_n \|_{H^1_0(\Omega)} + \| \varphi_n^{\perp} \|_{H^1_0(\Omega)} = \| \varphi \|_{H^1_0(\Omega)} \) for all \( n \in \mathbb{N} \), we get
\[
\| \partial_t v_n^\varepsilon (t) \|_{H^{-1}(\Omega)} = \sup_{\varphi \in H^1_0(\Omega) \setminus \{0\}} \frac{\langle \partial_t v_n^\varepsilon (t), \varphi \rangle}{\| \varphi \|_{H^1_0(\Omega)}} \leq (2\rho - 1) \| \partial_t v_n^\varepsilon (t) \| + (\rho^2 - \rho) \| v_n^\varepsilon (t) \| + (\rho - 1) \| \nabla v_n^\varepsilon (t) \|
\]
\[+ \| \partial_t \nabla v_n^\varepsilon (t) \| + C_1 \log(\gamma) \left( (1 - \rho) \| v_n^\varepsilon (t) \|_{H^1_0(\Omega)} + \| \partial_t v_n^\varepsilon (t) \|_{H^1_0(\Omega)} \right),
\]
where the last term in the right-hand side comes from the properties of \( P_\varepsilon^1 \) and \( P_\varepsilon^2 \).

From (21) and (23), there exists a constant \( \tilde{C} > 0 \) such that
\[
\| \partial_t v_n^\varepsilon \|_{L^2(0, T; H^{-1}(\Omega))} \leq \tilde{C} \quad \text{for all } n \in \mathbb{N}.
\]

Quasi-reversibility method for an inverse hyperbolic problem
Henceforth, from the Banach–Alaoglu theorem, there exists a subsequence of \( \{ v^n \} \) (still denoted by \( \{ v^n \} \)) such that
\[
\partial_{tt} v^n \rightharpoonup \partial_{tt} v \quad \text{weakly in} \quad L^2(0, T; H^{-1}(\Omega)).
\] (25)
Combining the above weak-star and weak limits, the function \( v \) satisfies
\[
\begin{cases}
  v \in L^\infty(0, T; H^1_0(\Omega)), \\
  \partial_t v \in L^\infty(0, T; L^2(\Omega)) \cap L^2(0, T; H^1_0(\Omega)), \\
  \partial_{tt} v \in L^2(0, T; H^{-1}(\Omega)).
\end{cases}
\]
Furthermore, since \( H^1_0(\Omega) \) is compactly embedded in \( L^2(\Omega) \) and \( L^2(\Omega) \) is continuously embedded in \( H^{-1}(\Omega) \) (by Rellich–Kondrachov), from Aubin–Lions lemma, we get
\[
\begin{cases}
  v^n \rightharpoonup v \quad \text{strongly in} \quad C([0, T]; H^1_0(\Omega)), \\
  \partial_t v^n \rightharpoonup \partial_t v \quad \text{strongly in} \quad C([0, T]; L^2(\Omega)).
\end{cases}
\] (26)
Fix an integer \( N \) and choose a function \( \bar{v} \in C^1(0, T; H^1_0(\Omega)) \) having the form
\[
\bar{v}(t) = \sum_{k=1}^{N} d_k(t)\phi_k,
\] (27)
where \( d_1, \ldots, d_N \) are given real valued \( C^1 \) functions defined in \([0, T]\). For all \( x \geq N \), multiplying (15), summing for \( k = 1, \ldots, N \) and integrating over \((0, T)\) lead to
\[
\begin{align*}
\int_{\Omega} \partial_{tt} v_n^\varepsilon(t) \bar{v} dx + (1 - 2\rho) \int_{\Omega} \partial_t v_n^\varepsilon(t) \bar{v} dx + (\rho^2 - \rho) \int_{\Omega} v_n^\varepsilon(t) \bar{v} dx \\
+ (\rho - 1) \int_{\Omega} \nabla v_n^\varepsilon(t) \cdot \nabla \bar{v} dx - \int_{\Omega} \nabla \partial_t v_n^\varepsilon(t) \cdot \nabla \bar{v} dx \\
= (1 - \rho) \int_{\Omega} \mathbf{P}^1_\varepsilon v_n^\varepsilon(t) \bar{v} dx + \int_{\Omega} \mathbf{P}^2_\varepsilon \partial_t v_n^\varepsilon(t) \bar{v} dx.
\end{align*}
\]
Letting \( n \to \infty \), we obtain from (26) that
\[
\begin{align*}
\int_{\Omega} \partial_{tt} v^\varepsilon(t) \bar{v} dx + (1 - 2\rho) \int_{\Omega} \partial_t v^\varepsilon(t) \bar{v} dx + (\rho^2 - \rho) \int_{\Omega} v^\varepsilon(t) \bar{v} dx \\
+ (\rho - 1) \int_{\Omega} \nabla v^\varepsilon(t) \cdot \nabla \bar{v} dx - \int_{\Omega} \nabla \partial_t v^\varepsilon(t) \cdot \nabla \bar{v} dx \\
= (1 - \rho) \int_{\Omega} \mathbf{P}^1_\varepsilon v^\varepsilon(t) \bar{v} dx + \int_{\Omega} \mathbf{P}^2_\varepsilon \partial_t v^\varepsilon(t) \bar{v} dx.
\end{align*}
\] (28)
Since the functions of the form (27) are dense in $L^2(0,T;H^1_0(\Omega))$, the equality (28) holds for all test function $\tilde{v} \in L^2(0,T;H^1_0(\Omega))$. We deduce that the function $v^\varepsilon$ obtained from approximate solutions $v_n^\varepsilon$ satisfies the weak formulation in Definition 3.2.

It now remains to verify the initial data for $v^\varepsilon$. Take $\kappa \in C^1([0,T])$ satisfying $\kappa(T) = 1$ and $\kappa(0) = 0$. It follows from (22) that

$$\int_0^T \langle \partial_t v_n^\varepsilon(t), \phi \rangle \kappa(t) dt \to \int_0^T \langle \partial_t v^\varepsilon(t), \phi \rangle \kappa(t) dt$$

for all $\phi \in H^1_0(\Omega)$. Then by integration by parts, one gets

$$\int_0^T \langle v_n^\varepsilon(t), \phi \rangle \partial_t \kappa(t) dt - \langle v_n^\varepsilon(T), \phi \rangle \kappa(T)$$

$$\to \int_0^T \langle v^\varepsilon(t), \phi \rangle \partial_t \kappa(t) dt - \langle v^\varepsilon(T), \phi \rangle \kappa(T)$$

and thereupon, we get $\langle v_n^\varepsilon(T), \phi \rangle \to \langle v^\varepsilon(T), \phi \rangle$ for all $\phi \in H^1_0(\Omega)$ by virtue of (22). From Lemma 3.4, we also have that $v_n^\varepsilon(T) \to f_0^\varepsilon$ in $L^2(\Omega)$ as $n \to \infty$. Thus $\langle v^\varepsilon(T), \phi \rangle = \langle f_0^\varepsilon, \phi \rangle$ for all $\phi \in H^1_0(\Omega)$, which implies that $v^\varepsilon(T) = f_0^\varepsilon$ a.e. in $\Omega$. Similarly, it follows from (25) that

$$\int_0^T \langle \partial_t v_n^\varepsilon(t), \phi \rangle \kappa(t) dt \to \int_0^T \langle \partial_t v^\varepsilon(t), \phi \rangle \kappa(t) dt$$

for all $\phi \in H^1_0(\Omega)$. Then by integration by parts, one gets

$$- \int_0^T \langle \partial_t v_n^\varepsilon(t), \phi \rangle \partial_t \kappa(t) dt + \langle \partial_t v_n^\varepsilon(T), \phi \rangle \kappa(T)$$

$$\to - \int_0^T \langle \partial_t v^\varepsilon(t), \phi \rangle \partial_t \kappa(t) dt + \langle \partial_t v^\varepsilon(T), \phi \rangle \kappa(T)$$

as $n \to \infty$.

Using the similar arguments as in the proof for $v^\varepsilon(T)$, we obtain that $\partial_t v^\varepsilon(T) = \rho f_0^\varepsilon + f_1^\varepsilon$ a.e. in $\Omega$. Hence, we complete the proof of the theorem. □

**Theorem 3.7.** Assume (9) holds. For each $\varepsilon > 0$, the regularized system (10)–(11) admits a unique weak solution $v^\varepsilon$ in the sense of Definition 3.2.

**Proof.** We sketch out some important steps because this proof is standard. Indeed, let $v^\varepsilon$ and $\tilde{v}^\varepsilon$ be two weak solutions of the system (10)–(11) Since the system is linear, it is straightforward to see that the function $k^\varepsilon = v^\varepsilon - \tilde{v}^\varepsilon$ satisfies (10) with zero terminal conditions $k^\varepsilon(T) = \partial_t k^\varepsilon(T) = 0$. Taking $\varphi = \partial_t k^\varepsilon$ as a test
function, we proceed as in the way to get the estimate (20). Hence, \( k^\varepsilon(t) = 0 \) a.e. in \((0, T)\) because of the fact that
\[
\frac{\|\partial_t k^\varepsilon(t)\|^2}{\rho - 1} + \rho \|k^\varepsilon(t)\|^2 + \|\nabla k^\varepsilon(t)\|^2 \leq 0 \quad \text{a.e. in } (0, T).
\]

This completes the proof of the theorem. \(\square\)

4 Convergence analysis

In this part, our focus is on the convergence analysis of the variational QR framework adapted to solve the time-reversed hyperbolic heat conduction problem. The error estimate obtained below can be viewed as a “worst-case” scenario of convergence of this QR scheme in case the stabilized operators \(P^1_\varepsilon\) and \(P^2_\varepsilon\) are bounded logarithmically. Some improvements are discussed in section 5.

It is worth noting that our analysis in section 3 does not care about the dependence of \(C\) (and any type of constants in there) on the noise level \(\varepsilon\), since basically we fix \(\varepsilon\). However, to this end any constant \(C > 0\) used below should be \(\varepsilon\)-independent because we are going to show the error estimates with respect to only \(\varepsilon\).

**Theorem 4.1.** Assume (9) holds. Let \(\varepsilon \in (0, 1)\) be a sufficiently small number such that \(\gamma := \gamma(\varepsilon) \geq e^{2/C_1}\). Suppose the following conditions hold
\[
\begin{cases}
3C_1T < 2, \\
\lim_{\varepsilon \to 0} \gamma^2(\varepsilon) \varepsilon \leq K.
\end{cases}
\]  

Next, assume the original system (2)–(4) admits a unique solution \(u\) such that \(u \in C([0, T]; W_1)\) and \(u_t \in L^2(0, T; W_2)\), where \(W_1, W_2\) are obtained in Definition 5. Let \(M > 0\) be such that
\[
\|u\|^2_{C([0, T]; W_1)} + \|u_t\|^2_{L^2(0, T; W_2)} \leq M.
\]

Let \(u^\varepsilon\) be a unique weak solution of the regularized system (7)–(8) analyzed in Theorems 3.6 and 3.7. Then the following error estimates hold:
\[
\begin{align*}
\|u^\varepsilon(t) - u(t)\|^2 &\leq C \left( \varepsilon + (\log(\gamma))^{-1} \gamma^{3C_1(T-t)-2} \right), \\
\|\nabla u^\varepsilon(t) - \nabla u(t)\|^2 &\leq C \left( \log(\gamma) \varepsilon + \gamma^{3C_1(T-t)-2} \right), \\
\|u^\varepsilon_t(t) - u_t(t)\|^2 &+ \int_t^T \|\nabla u^\varepsilon(s) - \nabla u(s)\|^2 \, ds \\
&\leq C \left( (\log(\gamma))^2 \varepsilon + \log(\gamma) \gamma^{3C_1(T-t)-2} \right).
\end{align*}
\]
where \( C = C (K, M, C_0, C_1) > 0 \) is independent of \( \varepsilon \).

**Proof.** Let \( w^\varepsilon (x, t) = [u^\varepsilon (x, t) - u(x, t)] e^{\rho_\varepsilon (t-\tau)} \) for some \( \rho_\varepsilon > 0 \), viewing as a weighted difference function in our proof of convergence. The notion behind this use of the Carleman weight function is to “maximize” the measured terminal data that we are having and thus, we can take full advantage of the noise level \( \varepsilon \). In principle, the downscaling (with respect to the noise level) used here is helpful in getting rid of the large stability magnitude by a suitable choice of the auxiliary parameter \( \rho_\varepsilon \), which is also relatively large. Now, we compute the equation for \( w^\varepsilon \), calling as the difference equation between the regularized problem (7)–(8) and the original system (2)–(4). In fact, we have

\[
w^\varepsilon_t = [u^\varepsilon_t - u_t] e^{\rho_\varepsilon (t-\tau)} + \rho_\varepsilon [u^\varepsilon - u] e^{\rho_\varepsilon (t-\tau)} \]

\[
= [u^\varepsilon_t - u_t] e^{\rho_\varepsilon (t-\tau)} + \rho_\varepsilon w^\varepsilon, \tag{30}
\]

\[
\Delta w^\varepsilon = [\Delta u^\varepsilon - \Delta u] e^{\rho_\varepsilon (t-\tau)}, \tag{31}
\]

which lead to

\[
w^\varepsilon_{tt} - \rho_\varepsilon w^\varepsilon_t = [u^\varepsilon_{tt} - u_{tt}] e^{\rho_\varepsilon (t-\tau)} + \rho_\varepsilon [u^\varepsilon_t - u_t] e^{\rho_\varepsilon (t-\tau)} \]

\[
= [u^\varepsilon_{tt} - u_{tt}] e^{\rho_\varepsilon (t-\tau)} + \rho_\varepsilon (w^\varepsilon_t - \rho_\varepsilon w^\varepsilon), \tag{32}
\]

\[
\Delta w^\varepsilon_t - \rho_\varepsilon \Delta w^\varepsilon = [\Delta u^\varepsilon_t - \Delta u_t] e^{\rho_\varepsilon (t-\tau)}. \tag{33}
\]

Hereby, we notice that when multiplying both sides of the systems (7)–(8) and (2)–(4) by the weight \( e^{\rho_\varepsilon (t-\tau)} \), it yields

\[
[u^\varepsilon_{tt} - u_{tt}] e^{\rho_\varepsilon (t-\tau)} + [u^\varepsilon_t - u_t] e^{\rho_\varepsilon (t-\tau)} + \Delta (u^\varepsilon - u) e^{\rho_\varepsilon (t-\tau)}
\]

\[
+ \Delta (u^\varepsilon_t - u_t) e^{\rho_\varepsilon (t-\tau)} = P_1^\varepsilon (u^\varepsilon - u) e^{\rho_\varepsilon (t-\tau)} + Q_1^\varepsilon u e^{\rho_\varepsilon (t-\tau)}
\]

\[
+ P_2^\varepsilon (u^\varepsilon_t - u_t) e^{\rho_\varepsilon (t-\tau)} + Q_2^\varepsilon u_t e^{\rho_\varepsilon (t-\tau)}. \tag{34}
\]

Henceforth, we plug the identities (30)–(33) into the equation (34) to get

\[
w^\varepsilon_{tt} + (\rho_\varepsilon^2 - \rho_\varepsilon) w^\varepsilon - (\rho_\varepsilon - 1) \Delta w^\varepsilon + \Delta w^\varepsilon_t
\]

\[
= P_1^\varepsilon w^\varepsilon + Q_1^\varepsilon u e^{\rho_\varepsilon (t-\tau)} + (2\rho_\varepsilon - 1) w^\varepsilon + P_2^\varepsilon (w^\varepsilon_t - \rho_\varepsilon w^\varepsilon) + Q_2^\varepsilon u_t e^{\rho_\varepsilon (t-\tau)}. \tag{35}
\]

which is the PDE for the difference function \( w^\varepsilon \).
Now, we multiply both sides of (35) by \( w^e_t \) and integrate the resulting equation over \( \Omega \). After some manipulations, we arrive at

\[
\frac{1}{2} \frac{d}{dt} \| w^e_t \|^2 + \frac{1}{2} (\rho^2 - \rho) \frac{d}{dt} \| w^e \|^2 + \frac{1}{2} (\rho - 1) \frac{d}{dt} \| \nabla w^e \|^2 - \| \nabla w^e_t \|^2
\]

\[
= (2\rho - 1) \| w^e_t \|^2 + \langle (P^e - \rho P^2) w^e, w^e \rangle + \langle P^2 w^e, w^e \rangle
\]

\[+ e^{\rho (t-T)} \langle Q^1_{\varepsilon} u, w^e \rangle + e^{\rho (t-T)} \langle Q^2_{\varepsilon} u, w^e \rangle. \tag{36} \]

Based upon the conditional estimates (5)–(6) we estimate the right-hand side of (36) as follows:

\[
\langle (P^e - \rho P^2) w^e, w^e \rangle \geq -\frac{1}{2} (\rho - 1) C_1^2 (\log(\gamma))^2 \| w^e \|^2 - \frac{1}{2} (\rho - 1) \| w^e \|^2,
\]

\[
\langle P^2 w^e, w^e \rangle \geq -C_1 \log(\gamma) \| w^e \|^2, \tag{37} \]

\[
e^{\rho (t-T)} \langle Q^1_{\varepsilon} u, w^e \rangle \geq -\frac{1}{2} \left( \frac{1}{4} \| w^e \|^2 + 4e^{2\rho (t-T)} C_0^2 \gamma^{-2} \| u \|^2_{W_1} \right), \tag{38} \]

\[
e^{\rho (t-T)} \langle Q^2_{\varepsilon} u, w^e \rangle \geq -\frac{1}{2} \left( \frac{1}{4} \| w^e \|^2 + 4e^{2\rho (t-T)} C_0^2 \gamma^{-2} \| u \|^2_{W_2} \right). \tag{39} \]

Therefore, by integrating (36) from \( t \) to \( T \) we estimate that

\[
\| w^e_t (t) \|^2 + (\rho^2 - \rho) \| w^e (t) \|^2 + (\rho - 1) \| \nabla w^e (t) \|^2 + 2 \int_t^T \| \nabla w^e (s) \|^2 ds
\]

\[
\leq \| w^e_t (T) \|^2 + (\rho^2 - \rho) \| w^e (T) \|^2 + (\rho - 1) \| \nabla w^e (T) \|^2
\]

\[+ 4C_0^2 \gamma^{-2} \rho e^{-1} \left( 1 - e^{2\rho (t-T)} \right) \| u \|^2_{C([0,T] ; W_1)} + 4C_0^2 \gamma^{-2} \| u \|_{L^2(0,T; W_2)}^2
\]

\[+ C_1^2 \rho e^{-1} (\log(\gamma))^2 \int_t^T \rho e (\rho - 1) \| w^e (s) \|^2 ds
\]

\[+ 2 \left( \frac{1}{2} (\rho - 1) + C_1 \log(\gamma) + \frac{1}{2} - 2 \rho e + 1 \right) \int_t^T \| w^e (s) \|^2 ds. \]

By choosing \( \rho = C_1 \log(\gamma) \geq 2 \) (since \( \gamma \geq e^{2/C_1} \)), the last term in the right-hand side becomes \( (2 - \rho) \int_t^T \| w^e_t (s) \|^2 ds \leq 0 \), we apply the Grönwall inequality to obtain

\[
\| w^e_t (t) \|^2 + (\rho^2 - \rho) \| w^e (t) \|^2 + (\rho - 1) \| \nabla w^e (t) \|^2 + 2 \int_t^T \| \nabla w^e (s) \|^2 ds
\]

\[
\leq \left[ 2 (\rho^2 + 1) \varepsilon^2 + \varepsilon^2 (\rho^2 - 1) + 4C_0^2 \gamma^{-2} M \right] \gamma^{C_1(T-t)}, \tag{40} \]
where we have used the measurement assumption (9) and the fact that
\[ \|w^\varepsilon(T)\|^2 = \|[w^\varepsilon(T) - u_t(T)] + \rho_\varepsilon [w^\varepsilon(T) - u(T)]\|^2 \leq 2(\rho_\varepsilon^2 + 1)\varepsilon^2. \quad (41) \]

Thus, using the back-substitution
\[ w^\varepsilon(x,t) = [w^\varepsilon(x,t) - u(x,t)] e^{\rho_\varepsilon(t-T)} = [w^\varepsilon(x,t) - u(x,t)] \gamma^{C_1(t-T)}, \quad (42) \]
we conclude the convergence in $L^2(\Omega)$ type as follows:
\[
\|u^\varepsilon(t) - u(t)\|^2 \\
\leq \left( \frac{2(\rho_\varepsilon^2 + 1)}{\rho_\varepsilon^2 - \rho_\varepsilon} + \frac{\rho_\varepsilon^2 - 1}{\rho_\varepsilon^2 - \rho_\varepsilon} \right) \varepsilon^2 \gamma^{3C_1(T-t)} + \frac{4}{\rho_\varepsilon^2 - \rho_\varepsilon} C_0^2 M \gamma^{-2} \gamma^{3C_1(T-t)} \\
\leq \frac{\rho_\varepsilon^2 - 1}{\rho_\varepsilon^2 - \rho_\varepsilon} \left( 3\gamma^{3C_1(T-t)} + \gamma^{3C_1(T-t)} \right) \varepsilon^2 + 4C_0^2 M \rho_\varepsilon^{-1} \gamma^{3C_1(T-t)-2} \\
\leq 2 \left( 4\gamma^{3C_1(T-t)} \varepsilon^2 + C_1^{-1} 2C_0^2 M (\log(\gamma))^{-1} \gamma^{3C_1(T-t)-2} \right).
\]

From (29), we get $\gamma^{3C_1(T-t)} \varepsilon^2 \leq K^{\frac{3C_1}{2}} \varepsilon$ and it follows from the previous inequality that
\[ \|u^\varepsilon(t) - u(t)\|^2 \leq C \left( \varepsilon + (\log(\gamma))^{-1} \gamma^{3C_1(T-t)-2} \right), \quad (43) \]
for some constant $C > 0$. In the same manner, we derive from (40) the convergence for the gradient terms:
\[ \|\nabla u^\varepsilon(t) - \nabla u(t)\|^2 \leq 2 \left( 4C_1 \log(\gamma) \gamma^{3C_1(T-t)} \varepsilon^2 + 2C_0 M \gamma^{3C_1(T-t)-2} \right) \\
\leq C \left( \log(\gamma) \varepsilon + \gamma^{3C_1(T-t)-2} \right).
\]

Now using the back-substitution (42), we get
\[ \nabla w^\varepsilon(t) = [\nabla w^\varepsilon(t) - \nabla u_t(t)] \gamma^{C_1(t-T)} + \rho_\varepsilon [\nabla u^\varepsilon(t) - \nabla u(t)] \gamma^{C_1(t-T)}. \]

It yields
\[
2 \int_t^T \|\nabla w^\varepsilon(s)\|^2 ds + 2 \int_t^T \|\nabla u^\varepsilon(s) - \nabla u(s)\|^2 \rho_\varepsilon^2 \gamma^{2C_1(s-T)} ds \\
\geq \int_t^T \|\nabla u^\varepsilon(s) - \nabla u_t(s)\|^2 \gamma^{2C_1(s-T)} ds \\
\geq \gamma^{2C_1(t-T)} \int_t^T \|\nabla u^\varepsilon(s) - \nabla u_t(s)\|^2 ds
\]
Thus it follows from (40) that
\[
\int_t^T \| \nabla u_\varepsilon^\gamma (s) - \nabla u_t (s) \|^2 \, ds
\]
\[
\leq \left( 4\rho_\varepsilon^2 + 4\rho_0^2 \gamma^{-2} M \right) \gamma^{3C_1(T-t)} + 2\rho_\varepsilon \gamma \int_t^T \| \nabla u_\varepsilon^\gamma (s) - \nabla u(s) \|^2 \, ds
\]
\[
\leq \left[ 4C_1^2 (\log(\gamma))^2 \varepsilon^2 \gamma^{3C_1(T-t)} + 4C_0^2 M \gamma^{3C_1(T-t)-2} \right]
\]
\[+ CT \rho_\varepsilon^2 \gamma^{2C_1(t-T)} \left( \varepsilon + (\log(\gamma))^{-1} \gamma^{3C_1(T-t)-2} \right),
\]
where we have used the estimate (43) for the last inequality. This implies
\[
\int_t^T \| \nabla u_\varepsilon^\gamma (s) - \nabla u_t (s) \|^2 \, ds \leq C \left( (\log(\gamma))^2 \varepsilon + \log (\gamma) \gamma^{3C_1(T-t)-2} \right).
\]

Finally, using the back-substitution (42), one has
\[
w_\varepsilon^\gamma (t) = [u_\varepsilon^\gamma (t) - u_t (t)] \gamma^{C_1(t-T)} + \rho_\varepsilon \left[ u_\varepsilon^\gamma (t) - u(t) \right] \gamma^{C_1(t-T)}.
\]

This implies
\[
\| u_\varepsilon^\gamma (t) - u_t (t) \|^2 \gamma^{2C_1(t-T)} \leq 2 \| w_\varepsilon^\gamma (s) \|^2 + 2\rho_\varepsilon^2 \| u_\varepsilon^\gamma (t) - u(t) \|^2.
\]

Applying the estimate of \( \| w_\varepsilon^\gamma \|^2 \) in (40) and \( \| u_\varepsilon^\gamma (t) - u(t) \|^2 \) in (43), we obtain
\[
\| u_\varepsilon^\gamma (t) - u_t (t) \|^2 \leq C \left( (\log(\gamma))^2 \varepsilon + \log (\gamma) \gamma^{3C_1(T-t)-2} \right).
\]

Hence, we complete the proof of the theorem.

\[\square\]

As a by-product of Theorem 4.1, an appropriate choice of \( \gamma \) is taken to state the following convergence result with the Hölder rates.

**Corollary 4.2.** Under the assumptions of Theorem 4.1, if we choose \( \gamma (\varepsilon) = \varepsilon^{-1/2} \), then for any \( \varepsilon \leq e^{-4/C_1} \) the following error estimates hold:
\[
\| u_\varepsilon (t) - u (t) \|^2 \leq C \left( \varepsilon + (\log(\varepsilon^{-1/2}))^{-1} \varepsilon^{1-3C_1(T-t)/2} \right),
\]
\[
\| \nabla u_\varepsilon (t) - \nabla u (t) \|^2 \leq C \left( \log(\varepsilon^{-1/2}) \varepsilon + \varepsilon^{1-3C_1(T-t)/2} \right),
\]
\[
\| u_\varepsilon^\gamma (t) - u_t (t) \|^2 + \int_t^T \| \nabla u_\varepsilon^\gamma (s) - \nabla u_t (s) \|^2 \, ds
\]
\[
\leq C \left( (\log(\varepsilon^{-1/2}))^2 \varepsilon + \log(\varepsilon^{-1/2}) \varepsilon^{1-3C_1(T-t)/2} \right).
\]

where \( C = C (M, C_0, C_1) > 0 \) is independent of \( \varepsilon \).
5 Amendments based upon some particular stabilizations

As we have enjoyed the Hölder rates of convergence in a certain mixed $L^2$–$H^1$ space in section 4, we remark that stabilizing the Laplacian operator $-\Delta$ in (2) can be neglected if we can facilitate the logarithmic bound of the stabilized operator (cf. Definition 2.2). This shows the flexibility of the QR method we are developing. Note again that Definition 2.2 is the “worst” case that the stabilized operator needs to gain the strong convergence of the QR scheme. Based upon a particular choice of the involved operators, we show that stabilizing the highest order differential operator (i.e. $-\Delta u_t$) is sufficient to gain the convergence of $u^\varepsilon$ towards the “ideal” exact solution. This leads to the consideration of the following regularized equation for (2):

$$u^\varepsilon_{tt} + u^\varepsilon_t - \Delta u^\varepsilon + \Delta u^\varepsilon_t = P^2_\varepsilon u^\varepsilon_t \quad \text{in } \Omega \times (0, T),$$

where, similar to (7), $P^2_\varepsilon = 2\Delta + Q^2_\varepsilon$. We recall according to the standard result for the Dirichlet eigenvalue problem (see Remark 3.3) that there exists an orthonormal basis of $L^2(\Omega)$, denoted by $\{\phi_p\}_{p \in \mathbb{N}}$, such that $\phi_p \in H^1_0(\Omega) \cap C^\infty(\overline{\Omega})$ and $-\Delta \phi_p = \mu_p \phi_p$. The Dirichlet eigenvalues $\{\mu_p\}_{p \in \mathbb{N}}$ form an infinite sequence which goes to infinity in the following sense

$$0 \leq \mu_0 < \mu_1 < \mu_2 < \ldots, \lim_{p \to \infty} \mu_p = \infty.$$

Thus, we introduce

$$Q^2_\varepsilon u = 2\gamma^{-1} \sum_{p \in \mathbb{N}} \mu_p^{1/2} \langle u, \phi_p \rangle \phi_p,$$

and it is immediate to see that the conditional one (5) holds for $W_2 = H^1(\Omega)$ and $C_0 = 2$ by using the Parseval identity. Therefore, one has

$$P^2_\varepsilon u_t = 2\Delta u_t + Q^2_\varepsilon u_t = -2 \sum_{p \in \mathbb{N}} \mu_p \langle u_t, \phi_p \rangle \phi_p + 2\gamma^{-1} \sum_{p \in \mathbb{N}} \mu_p^{1/2} \langle u_t, \phi_p \rangle \phi_p$$

$$= 2\gamma^{-2} \sum_{p \in \mathbb{N}} \mu_p^{1/2} (1 - \gamma \mu_p^{1/2}) \langle u_t, \phi_p \rangle \phi_p.$$

Since $|\gamma \mu_p^{1/2} (1 - \gamma \mu_p^{1/2})| \leq 1/4$, we conclude that $\|P^2_\varepsilon u_t\| \leq \frac{1}{2} \gamma^{-2} \|u_t\|$, which avoids the logarithmic boundedness we are supposing. It is also worth mentioning that since $W_2 = H^1(\Omega)$, we mean to assume an usual weak solution of (2) in the forward problem, viz. $u_t \in L^2(0, T; H^1(\Omega))$. Henceforth, we state the following convergence result.
Theorem 5.1. Let $u^\varepsilon$ be a unique solution of the regularized system (44)–(8). Assume that we have
\[
\left\| \mathbf{P}_\varepsilon^2 h \right\| \leq C_1 \| h \|, \quad \text{for } h \in L^2(\Omega),
\]
instead of the generic conditional estimate (6). Under the assumptions of Theorem 4.1, we replace (29) by
\[
\begin{cases}
C_1 T < 1, \\
\lim_{\varepsilon \to 0} \gamma^2(\varepsilon) \varepsilon \leq K.
\end{cases}
\]

Then the following error estimates hold:
\[
\begin{align*}
\| u^\varepsilon(t) - u(t) \|^2 &\leq C \left( \varepsilon + (\log(\gamma))^{-1} \gamma^{2C_1(T-t) - 2} \right), \\
\| \nabla u^\varepsilon(t) - \nabla u(t) \|^2 &\leq C \left( \log(\gamma) \varepsilon + \gamma^{2C_1(T-t) - 2} \right), \\
\| u_1^\varepsilon(t) - u_1(t) \|^2 + \int_t^T \| \nabla u_1^\varepsilon(s) - \nabla u_1(s) \|^2 ds &\leq C \left( (\log(\gamma))^2 \varepsilon + \log(\gamma) \gamma^{2C_1(T-t) - 2} \right),
\end{align*}
\]
where $C = C(K, M, C_0, C_1) > 0$ is independent of $\varepsilon$.

Proof. To prove this theorem, we again take into account the difference equation
\[
w^\varepsilon(x, t) = [u^\varepsilon(x, t) - u(x, t)] e^{\rho_\varepsilon(t-T)},
\]
which obeys the following equation:
\[
w_1^\varepsilon + \left( \rho_\varepsilon^2 - \rho_\varepsilon \right) w^\varepsilon - (\rho_\varepsilon + 1) \Delta w^\varepsilon + \Delta w_1^\varepsilon \\
= (2\rho_\varepsilon - 1) w_1^\varepsilon + \mathbf{P}_\varepsilon^2 (w_1^\varepsilon - \rho_\varepsilon w^\varepsilon) + Q_\varepsilon^2 u_1 e^{\rho_\varepsilon(t-T)}. \tag{47}
\]

Next, we multiply both sides of (47) by $w_1^\varepsilon$ and integrate the resulting equation over $\Omega$ to get the following energy identity:
\[
\begin{align*}
\frac{1}{2} \frac{d}{dt} \| w_1^\varepsilon \|^2 + \frac{1}{2} \left( \rho_\varepsilon^2 - \rho_\varepsilon \right) \frac{d}{dt} \| w^\varepsilon \|^2 + \frac{1}{2} (\rho_\varepsilon + 1) \frac{d}{dt} \| \nabla w^\varepsilon \|^2 &- \| \nabla w_1^\varepsilon \|^2 \\
= (2\rho_\varepsilon - 1) \| w_1^\varepsilon \|^2 - \rho_\varepsilon \langle \mathbf{P}_\varepsilon^2 w^\varepsilon, w_1^\varepsilon \rangle + \langle \mathbf{P}_\varepsilon^2 w_1^\varepsilon, w_1^\varepsilon \rangle + e^{\rho_\varepsilon(t-T)} \langle Q_\varepsilon^2 u_1, w_1^\varepsilon \rangle.
\end{align*}
\]

Thus, by finding the lower bounds of the right-hand side of (48) we can easily obtain the target estimates. More specifically, we use the condition (45) to obtain the following estimates
\[
\begin{align*}
\langle -\rho_\varepsilon \mathbf{P}_\varepsilon^2 w^\varepsilon, w_1^\varepsilon \rangle &\geq -\frac{1}{2} \rho_\varepsilon C_1^2 \| w^\varepsilon \|^2 - \frac{1}{2} \rho_\varepsilon \| w_1^\varepsilon \|^2, \\
\langle \mathbf{P}_\varepsilon^2 w_1^\varepsilon, w_1^\varepsilon \rangle &\geq -C_1 \| w_1^\varepsilon \|^2,
\end{align*}
\]
and keep the estimate (39). Therefore, by integrating (48) from $t$ to $T$ we estimate that
\[ \|w_t^e(t)\|^2 + (\rho_e^2 - \rho_e) \|w^e(t)\|^2 + (\rho_e - 1) \|\nabla w^e(t)\|^2 + 2 \int_t^T \|\nabla w_t^e(s)\|^2 \, ds \]
\[ \leq \|w_t^e(T)\|^2 + (\rho_e^2 - \rho_e) \|w^e(T)\|^2 + (\rho_e - 1) \|\nabla w^e(T)\|^2 \]
\[ + 4C_0^2 \gamma^{-2} \|u_t\|^2_{L^2[0,T;W^1_2]} + C_1^2 (\rho_e - 1)^{-1} \int_t^T \rho_e (\rho_e - 1) \|w^e(s)\|^2 \, ds \]
\[ + 2 \left( \frac{1}{2} \rho_e + C_1 + \frac{1}{4} - 2 \rho_e + 1 \right) \int_t^T \|w_t^e(s)\|^2 \, ds. \]

By choosing $\rho_e = C_1 \log(\gamma) \geq 2$, the sum of two last terms of the right-hand side is less than or equal to $C_1^2 \int_t^T \rho_e (\rho_e - 1) \|w^e(s)\|^2 \, ds$. Thus, this implies from the Grönwall inequality that
\[ \|w_t^e(t)\|^2 + (\rho_e^2 - \rho_e) \|w^e(t)\|^2 + (\rho_e - 1) \|\nabla w^e(t)\|^2 + 2 \int_t^T \|\nabla w_t^e(s)\|^2 \, ds \]
\[ \leq \left[ 2 (\rho_e^2 + 1) \varepsilon^2 + \varepsilon^2 (\rho_e^2 - 1) + 4C_0^2 \gamma^{-2} M \right] e^{C_1^2 (T-t)}. \] (49)

The rest of the proof can be proceeded as in the proof of Theorem 4.1.

**Remark 5.2.** In Theorem 5.1, thanks to the assumption (45), we have obtained the term $e^{C_1^2 (T-t)}$ in (49) instead of the term $\gamma^{C_1 (T-t)} = e^{\rho_e (T-t)}$ in (40). This leads to the term $\gamma^{2C_1 (T-t)-2}$ in the error estimates. Thus, condition (46) certainly follows. Note that if we replace (45) by (6), then the term $\gamma^{3C_1 (T-t)-2}$ will appear in the error estimates of Theorem 5.1, which brings us back to condition (29).

The above-mentioned perturbing and stabilized operators result in another amendment. Now, we consider a way to get rid of the term $\langle (P_\varepsilon^1 - \rho_\varepsilon P_\varepsilon^2) w^e, w_t^e \rangle$ in the identity (36), which is the main factor of slowing down our rate of convergence. To improve this Hölder rate, we introduce another regularized equation for (2):
\[ u_{tt}^e + u_t^e + (\rho_e - 1) \Delta u^e + \Delta u_t^e = P_\varepsilon^3 u^e + P_\varepsilon^4 u_t^e \quad \text{in} \quad \Omega \times (0,T), \] (50)

where, in principle, $P_\varepsilon^3 = \rho_\varepsilon \Delta + Q_\varepsilon^3$ and $P_\varepsilon^4 = 2 \Delta + Q_\varepsilon^4$. Keep in mind that $\rho_\varepsilon = C_1 \log(\gamma) \geq 2$ as we have chosen in the proof of Theorem 4.1. Here, we propose a particular choice of the operator $Q_\varepsilon^4$:
\[ Q_\varepsilon^4 h = \gamma^{-1} \rho_\varepsilon^{-1} \sum_{p \in \mathbb{N}} \mu_p^{1/2} \langle h, \phi_p \rangle \phi_p + \gamma^{-1} \rho_\varepsilon^{-1} \sum_{\mu_p \geq \gamma^{-2}} \mu_p^{1/2} \langle h, \phi_p \rangle \phi_p, \] (51)
and the operator $P_4^\varepsilon$:

$$
P_4^\varepsilon h = 2\Delta h + Q_4^\varepsilon h
= 2 \sum_{\mu_p \geq \gamma^{-2}} \mu_p^{1/2} \left( \gamma^{-1} \rho_p^{-1} - \mu_p^{1/2} \right) \langle h, \phi_p \rangle \phi_p
+ \gamma^{-1} \rho_p^{-1} \sum_{\mu_p < \gamma^{-2}} \mu_p^{1/2} \langle h, \phi_p \rangle \phi_p - 2 \sum_{\mu_p < \gamma^{-2}} \mu_p \langle h, \phi_p \rangle \phi_p.
$$

Observe that the perturbation $Q_4^\varepsilon$ satisfies (5) with $W = H_0^1(\Omega)$ and $C_0 = 1$ since

$$
\|Q_4^\varepsilon h\|^2 \leq 4\gamma^{-2} \rho_p^{-2} \sum_{p \in \mathbb{N}} \mu_p |\langle h, \phi_p \rangle|^2 \leq \frac{\|\nabla h\|^2}{\gamma^2},
$$

where the last inequality comes from the fact that $\rho_p \geq 2$. On the other hand, the stabilized $P_4^\varepsilon$ satisfies $\|P_4^\varepsilon h\| \leq 3\gamma^{-2} \|h\|$ for $h \in L^2(\Omega)$. Next, we choose $P_3^\varepsilon = \rho_\varepsilon \Delta + Q_3^\varepsilon$, where

$$
Q_3^\varepsilon h = \rho_\varepsilon \Delta h + \rho_\varepsilon Q_4^\varepsilon h
= \rho_\varepsilon^{-1} \sum_{\mu_p \geq \gamma^{-2}} \rho_\varepsilon \mu_p^{1/2} \left( \gamma^{-1} - \rho_\varepsilon \mu_p^{1/2} \right) \langle h, \phi_p \rangle \phi_p
+ \gamma^{-1} \sum_{\mu_p < \gamma^{-2}} \mu_p^{1/2} \langle h, \phi_p \rangle \phi_p - \rho_\varepsilon \sum_{\mu_p < \gamma^{-2}} \mu_p \langle h, \phi_p \rangle \phi_p.
$$

Here, one can check that the perturbation $Q_3^\varepsilon$ satisfies (5) with $W = L^2(\Omega)$ and $C_0 = 1/8$ since

$$
\|Q_3^\varepsilon h\|^2 = \rho_\varepsilon^{-2} \sum_{\mu_p \geq \gamma^{-2}} \rho_\varepsilon^2 \mu_p \left( \gamma^{-1} - \rho_\varepsilon \mu_p^{1/2} \right)^2 |\langle h, \phi_p \rangle|^2
+ \sum_{\mu_p < \gamma^{-2}} \left( \gamma^{-1} \mu_p^{1/2} - \rho_\varepsilon \mu_p \right)^2 |\langle h, \phi_p \rangle|^2
\leq \frac{\gamma^{-4}}{16\rho_\varepsilon^2} \sum_{p \in \mathbb{N}} |\langle h, \phi_p \rangle|^2 \leq \frac{1}{64} \frac{\|h\|^2}{\gamma^2}.
$$

Moreover, this way we have $P_3^\varepsilon = 2\rho_\varepsilon \Delta + \rho_\varepsilon Q_4^\varepsilon = \rho_\varepsilon P_4^\varepsilon$. In the following, we show that the solution $u_\varepsilon$ of (50) converges to $u$ of (2) with the Lipschitz rate.
Theorem 5.3. Let \( u^\varepsilon \) be a unique solution of the regularized system (50)–(8). Let \( \varepsilon \in (0, 1) \) be a sufficiently small number such that \( \gamma := \gamma (\varepsilon) \geq \varepsilon^2/C_1 \). Suppose that we can choose \( P^3_\varepsilon = C_1 \log (\gamma) P^4_\varepsilon \). Under the assumptions of Theorem 4.1 without (29), we obtain the following error estimates:

\[
\|u^\varepsilon (t) - u(t)\|^2 \leq C \left( \varepsilon^2 + \gamma^{-2} \right),
\]

\[
\|\nabla u^\varepsilon (t) - \nabla u(t)\|^2 + \|u^\varepsilon (t) - u(t)\|^2 + \int_t^T \|\nabla u^\varepsilon (s) - \nabla u_t (s)\|^2 ds \leq C \left( (\log(\gamma))^2 \varepsilon^2 + \gamma^{-2} \right),
\]

where \( C = C (T, M, C_0, C_1) > 0 \) is independent of \( \varepsilon \).

Proof. As in the proof of Theorem 5.1, we only address some important parts that enable us to prove the aimed rate of convergence. The weighted difference function \( w^\varepsilon (x, t) = [u^\varepsilon (x, t) - u (x, t)] e^{\rho \varepsilon (t-T)} \) satisfies

\[
w^\varepsilon_t + \left( \rho^2 - \rho_\varepsilon \right) w^\varepsilon - \Delta w^\varepsilon + \Delta w^\varepsilon_i = P^3_\varepsilon w^\varepsilon + Q^3_\varepsilon u e^{\rho \varepsilon (t-T)} + (2\rho_\varepsilon - 1) w^\varepsilon_i + P^4_\varepsilon (w^\varepsilon_i - \rho \varepsilon w^\varepsilon) + Q^4_\varepsilon u_t e^{\rho \varepsilon (t-T)}.
\]

(54)

where we have relied on the fact that \( P^3_\varepsilon = \rho_\varepsilon \Delta + Q^3_\varepsilon \) and \( P^4_\varepsilon = 2\Delta + Q^4_\varepsilon \). Therefore, by the choice \( P^3_\varepsilon = C_1 \log (\gamma) P^4_\varepsilon \) and by taking \( \rho_\varepsilon = C_1 \log (\gamma) \geq 2 \), we obtain the resulting energy identity after testing (54) with \( w^\varepsilon_i \):

\[
\frac{1}{2} \frac{d}{dt} \|w^\varepsilon_i\|^2 + \frac{1}{2} \left( \rho^2 - \rho_\varepsilon \right) \frac{d}{dt} \|w^\varepsilon\|^2 + \frac{1}{2} \frac{d}{dt} \|\nabla w^\varepsilon\|^2 - \|\nabla w^\varepsilon_i\|^2 = (2\rho_\varepsilon - 1) \|w^\varepsilon_i\|^2 + \langle P^4_\varepsilon w^\varepsilon_i, w^\varepsilon \rangle + e^{\rho \varepsilon (t-T)} \langle Q^3_\varepsilon u, w^\varepsilon_i \rangle + e^{\rho \varepsilon (t-T)} \langle Q^4_\varepsilon u_t, w^\varepsilon_i \rangle.
\]

(55)

By the same arguments as in (37)–(39) applied to the last three terms of (55), we integrate (55) from \( t \) to \( T \) and estimate the resulting equality as follows:

\[
\|w_t (t)\|^2 + \left( \rho^2 - \rho_\varepsilon \right) \|w^\varepsilon (t)\|^2 + \|\nabla w^\varepsilon (t)\|^2 + 2 \int_t^T \|\nabla w^\varepsilon_i (s)\|^2 ds \leq \|w_t (T)\|^2 + \left( \rho^2 - \rho_\varepsilon \right) \|w^\varepsilon (T)\|^2 + \|\nabla w^\varepsilon (T)\|^2 + 4C_0^2 \gamma^{-2} \rho^{-1} \left( 1 - e^{2\rho_\varepsilon (t-T)} \right) \|u\|^2_{C([0,T]; \overline{\mathbb{W}})} + 4C_0^2 \gamma^{-2} \|u_t\|^2_{L^2(0,T; \overline{\mathbb{W}})} + 2 [-C_1 \log (\gamma) + 2] \int_t^T \|u^\varepsilon_i (s)\|^2 ds.
\]
Using (41) and applying the Grönwall inequality, we arrive at
\[
\|w_\varepsilon(t)\|^2 + (\rho_\varepsilon^2 - \rho_\varepsilon) \|w_\varepsilon(t)\|^2 + \|\nabla w_\varepsilon(t)\|^2 + 2 \int_t^T \|\nabla w_\delta^\varepsilon(s)\|^2 \, ds \\
\leq [2 (\rho_\varepsilon^2 + 1) \varepsilon^2 + \varepsilon^2 (\rho_\varepsilon^2 - \rho_\varepsilon + 1) + 4C_0^2 \gamma^{-2} M] \gamma^{2C_1(T-T)} e^{2T(C_0+1)},
\]
which yields the target error estimates. Hence, we complete the proof of the theorem.

\[\Box\]

**Corollary 5.4.** Under the assumptions of Theorem 5.3, if we choose \(\gamma(\varepsilon) = \varepsilon^{-1}\), then for any \(\varepsilon \leq e^{-2/C_1}\) the following error estimates hold:
\[
\|u_\varepsilon(t) - u(t)\|^2 \leq C \varepsilon^2,
\]
\[
\|\nabla u_\varepsilon(t) - \nabla u(t)\|^2 + \|u_\varepsilon(t) - u_t(t)\|^2 + \int_t^T \|\nabla u_\varepsilon^\varepsilon(s) - \nabla u_t(s)\|^2 \, ds \\
\leq C (\log(\varepsilon^{-1}))^2 \varepsilon^2,
\]
where \(C = C (T, M, C_0, C_1) > 0\) is independent of \(\varepsilon\).

In Corollary 5.4, we see the Lipschitz rate of convergence in \(C([0,T]; L^2(\Omega))\) when a special regularized equation (50) is investigated. We have also found that under the choice of \(P_3^\varepsilon\) and \(P_4^\varepsilon\) that we have proposed in (51)–(53), our convergence works for any finite time \(T > 0\), compared to the ones assumed in (29) and (46). Furthermore, the source condition in this scenario is merely \(u \in C([0,T]; L^2(\Omega))\) and \(u_t \in L^2(0,T; H^1(\Omega))\), which is a very least one for weak solutions of the forward problem. Last but not least, we remark that if the measurement assumption (9) is only given by \(\|u_\varepsilon(\cdot,T) - u(\cdot,T)\|_{L^2(\Omega)} \leq \varepsilon\), we obtain the logarithmic rate of convergence in the following sense:
\[
\|u_\varepsilon(t) - u(t)\|^2 \leq C \rho_\varepsilon^{-2} \leq C / (\log(\gamma))^2,
\]
by virtue of (56).

**Acknowledgments.** V.A.K would like to thank Prof. Michael Victor Klibanov (Charlotte, USA) for his wholehearted guidance during the fellowship at UNCC and for giving a chance to delve into the Carleman estimates and convexification.

**Bibliography**

[1] M. Asch, M. Bocquet and M. Nodet, *Data Assimilation: Methods, Algorithms, and Applications*, Society for Industrial and Applied Mathematics, 2016.
[2] S. A. Avdonin, S. A. Ivanov and J. M. Wang, Inverse problems for the heat equation with memory, *Inverse Problems & Imaging* **13** (2019), 31–38.

[3] E. Bonetti, E. Rocca, R. Scala and G. Schimperna, On the strongly damped wave equation with constraint, *Communications in Partial Differential Equations* **42** (2017), 1042–1064.

[4] A. S. Carasso, J. G. Sanderson and J. M. Hyman, Digital removal of random media image degradations by solving the diffusion equation backwards in time, *SIAM Journal on Numerical Analysis* **15** (1978), 344–367.

[5] C. I. Christov and P. M. Jordan, Heat conduction paradox involving second-sound propagation in moving media, *Physical Review Letters* **94** (2005).

[6] V. N. Doan, H. T. Nguyen, V. A. Khoa and V. A. Vo, A note on the derivation of filter regularization operators for nonlinear evolution equations, *Applicable Analysis* **97** (2017), 3–12.

[7] R. E. Ewing, The approximation of certain parabolic equations backward in time by Sobolev equations, *SIAM Journal on Mathematical Analysis* **6** (1975), 283–294.

[8] F. Gazzola and M. Squassina, Global solutions and finite time blow up for damped semilinear wave equations, *Annales de l'Institut Henri Poincare (C) Non Linear Analysis* **23** (2006), 185–207.

[9] R. Jaroudi, G. Baravdish, F. Åström and B. T. Johansson, *Source Localization of Reaction-Diffusion Models for Brain Tumors*, Lecture Notes in Computer Science, Springer International Publishing, 2016, pp. 414–425.

[10] D. Jou, J. Casas-Vázquez and G. Lebon, *Extended Irreversible Thermodynamics*, Springer Berlin Heidelberg, 2001.

[11] S. I. Kabanikhin, Definitions and examples of inverse and ill-posed problems, *Journal of Inverse and Ill-posed Problems* **16** (2008), 317–357.

[12] B. Kaltenbacher and W. Rundell, Regularization of a backwards parabolic equation by fractional operators, *Inverse Problems & Imaging* **13** (2019), 401–430.

[13] M. V. Klibanov, Carleman estimates for the regularization of ill-posed Cauchy problems, *Applied Numerical Mathematics* **94** (2015), 46–74.

[14] M. V. Klibanov and A. G. Yagola, Convergent numerical methods for parabolic equations with reversed time via a new Carleman estimate, *Inverse Problems* **35** (2019), 115012.

[15] R. Lattès and J. L. Lions, *Méthode de Quasi-réversibilité et Applications*, Paris, Dunod, 1967.

[16] N. T. Long and A. P. N. Dinh, Approximation of a parabolic non-linear evolution equation backwards in time, *Inverse Problems* **10** (1994), 905–914.

[17] N. T. Long and A. P. N. Dinh, Note on a regularization of a parabolic nonlinear evolution equation backwards in time, *Inverse Problems* **12** (1996), 455–462.
[18] T. N. Luan and T. Q. Khanh, On the backward problem for parabolic equations with memory, *Applicable Analysis* (2019), 1–18.

[19] V. Méndez and J. Camacho, Dynamics and thermodynamics of delayed population growth, *Physical Review E* 55 (1997), 6476–6482.

[20] H. T. Nguyen, V. A. Khoa and V. A. Vo, Analysis of a quasi-reversibility method for a terminal value quasi-linear parabolic problem with measurements, *SIAM Journal on Mathematical Analysis* 51 (2019), 60–85.

[21] S. B. M. Sambatti, H. F. de Campos Velho and L. D. Chiwiacowsky, *Epidemic Genetic Algorithm for Solving Inverse Problems: Parallel Algorithms*, Integral Methods in Science and Engineering, Springer International Publishing, 2019, pp. 381–394.

[22] R. E. Showalter and T. W. Ting, Pseudoparabolic partial differential equations, *SIAM Journal on Mathematical Analysis* 1 (1970), 1–26.

[23] N. H. Tuan, V. V. Au, V. A. Khoa and D. Lesnic, Identification of the population density of a species model with nonlocal diffusion and nonlinear reaction, *Inverse Problems* 33 (2017), 055019.

[24] N. H. Tuan, V. A. Khoa, M. T. N. Truong, T. T. Hung and M. N. Minh, Application of the cut-off projection to solve a backward heat conduction problem in a two-slab composite system, *Inverse Problems in Science and Engineering* 27 (2018), 460–483.

[25] N. H. Tuan, D. V. Nguyen, V. V. Au and D. Lesnic, Recovering the initial distribution for strongly damped wave equation, *Applied Mathematics Letters* 73 (2017), 69–77.

[26] J. J. Vadasz, S. Govender and P. Vadasz, Heat transfer enhancement in nano-fluids suspensions: Possible mechanisms and explanations, *International Journal of Heat and Mass Transfer* 48 (2005), 2673–2683.

**Author information**

Vo Anh Khoa, Department of Mathematics and Statistics, University of North Carolina at Charlotte, Charlotte, North Carolina 28223, USA.  
E-mail: vakhoa.hcmus@gmail.com

Manh-Khang Dao, Department of Mathematics, KTH Royal Institute of Technology, 100 44 Stockholm, Sweden.  
E-mail: mkdao@kth.se