ERROR ESTIMATES FOR FINITE ELEMENT APPROXIMATIONS OF PARABOLIC EQUATIONS WITH MEASURE DATA

WEI GONG

ABSTRACT. In this paper we study the a priori error estimates for the finite element approximations of parabolic equations with measure data, especially we consider problems with separate measure data in time and space, respectively. The solutions of these kinds of problems exhibit low regularities due to the existence of measure data, this introduces some difficulties in both theoretical and numerical analysis. For both cases we use standard piecewise linear and continuous finite elements for the space discretization and derive the a priori error estimates for the semi-discretization problems, while the backward Euler method is then used for time discretization and a priori error estimates for the fully discrete problems are also derived. Numerical results are provided at the end of the paper to confirm our theoretical findings.

1. Introduction

The aim of this paper is to analyze the finite element approximations of parabolic equations with measure data. Let \( \Omega_T = \Omega \times (0, T) \), \( \Gamma_T = \partial \Omega \times (0, T) \), \( \Omega \) is an open bounded domain in \( \mathbb{R}^d \) (\( d = 2 \) or 3) with boundary \( \Gamma = \partial \Omega \). We consider the following parabolic problems:

\[
\begin{align*}
\partial_t y + A y &= \mu \quad \text{in } \Omega_T, \\
y &= 0 \quad \text{on } \Gamma_T, \\
y(\cdot, 0) &= y_0 \quad \text{in } \Omega,
\end{align*}
\]

(1.1)

where \( \partial_t y = \frac{\partial y}{\partial t} \), the operator \( A \) is assumed to be a second order elliptic partial differential operator, \( y_0 \in L^2(\Omega) \) and \( T > 0 \) are fixed.

Here we consider two kinds of problems with measure data. At first, we consider problem (1.1) with measure data in time, i.e., \( \mu = g \omega \), \( g \) and \( \omega \) are given functions such that \( g \in C([0, T]; L^2(\Omega)) \) and \( \omega \in \mathcal{M}[0, T] \), where \( \mathcal{M}[0, T] \) is the space of the real and regular Borel measures in \([0, T]\), which can be defined as the dual space of \( C[0, T] \) with its natural norm

\[
\|\mu\|_{\mathcal{M}[0,T]} = \sup \left\{ \int_0^T v d\mu : v \in C[0, T] \text{ and } \|v\|_{C[0,T]} \leq 1 \right\}.
\]

Received by the editor February 5, 2011 and, in revised form, August 8, 2011 and September 13, 2011.

2010 Mathematics Subject Classification. Primary 49J20, 49K20, 65N15, 65N30.

Key words and phrases. Finite element method, parabolic equation, measure data, semi-discrete error estimates, fully discrete error estimates.

This work was partially supported by the National Natural Science Foundation of China under grant 11171337 and the National Basic Research Program of China under grant 2012CB821204.

©2012 American Mathematical Society
Reverts to public domain 28 years from publication
One of the most important applications for parabolic equations with measure data in time is related to the first order optimality conditions of parabolic optimal control problems with pointwise state constraints. Optimal control problems governed by parabolic PDE reads as:

\[
\begin{align*}
\min \quad & J(y, u) = \frac{1}{2} \|y - y_d\|_{L^2(\Omega_T)}^2 + \frac{\alpha}{2} \|u\|_{L^2(\Omega_T)}^2 \\
\text{subject to} \quad & \begin{cases}
\partial_t y + Ay = u \quad \text{in } \Omega_T, \\
y = 0 \quad \text{on } \Gamma_T, \\
y(\cdot, 0) = y_0 \quad \text{in } \Omega,
\end{cases}
\end{align*}
\]

where \(y\) denotes the state variable and \(u\) denotes the control (see, e.g., [18]). If we impose state constraints pointwise in time, for example (see, e.g., Example 2.3 in [6] and [22]),

\[
\int_{\Omega} g(x, t, y(x, t)) dx \leq b(t) \quad \forall \ t \in [0, T]
\]

with given function \(b(t)\), then the adjoint state \(p\) associated to the first order optimality conditions satisfies

\[
\begin{align*}
\begin{cases}
-\partial_t p + A^* p = y - y_d + \mu_{\Omega_T} \quad \text{in } \Omega_T, \\
p = \mu_{\Sigma_T} \quad \text{on } \Sigma_T, \\
p(\cdot, T) = \mu_T \quad \text{in } \Omega,
\end{cases}
\end{align*}
\]

in the sense of distributions (see, e.g., [6], [10] and [22]). In general, the Lagrange multiplier \(\mu\) associated to the state constraints for parabolic optimal control problems with pointwise state constraints belongs to \(\mathcal{M}(\Omega_T)\), where \(\mathcal{M}(\Omega_T)\) is the space of regular Borel measures on \(\Omega_T\), \(\mu_{\Omega_T} := \mu|_{\Omega_T}\), \(\mu_{\Gamma_T} := \mu|_{\Gamma_T}\) and \(\mu_T := \mu|_{\partial \Omega \times \{T\}}\). However, in this case where only pointwise in time state constraints are imposed, the Lagrange multiplier \(\mu\) associated to the state constraints (1.4) appears to be a measure only in time, and can be decomposed as \(\mu = g\omega\), \(g\) and \(\omega\) are given functions such that \(g \in C([0, T]; L^2(\Omega))\) and \(\omega \in \mathcal{M}[0, T]\). Thus the associated (to the state) adjoint equation exhibits the similar structure of (1.1).

Then, we consider problems with measure data in space, i.e., \(\mu = g\omega\), \(g\) and \(\omega\) are given functions such that \(g \in L^2(0, T; C(\Omega))\) and \(\omega \in \mathcal{M}(\Omega)\). Here \(\mathcal{M}(\Omega)\) is the space of the real and regular Borel measures on \(\Omega\), and can be defined as the dual space of \(C(\Omega)\) with its natural norm

\[
\|\mu\|_{\mathcal{M}(\Omega)} = \sup \left\{ \int_{\Omega} v \mu : \ v \in C(\Omega) \text{ and } \|v\|_{C(\Omega)} \leq 1 \right\}.
\]

The problems of form (1.1) with measure data in space can be used to model the potential of an electric field with an electric charge distribution \(\mu\) (see [5]). These kinds of problems also arise in other types of applications, for instance, modeling of acoustic monopoles, transport equations for effluent discharge in aquatic media (see [1]). In the design and management of waste-water treatment systems, mainly the disposal of sea outfalls discharging polluting effluent from a sewerage system (see [21] for details), the problem can be formulated as an optimal control problem with
pointwise state and control constraints and pointwise control, while the governing state equation of which is of the form (1.1) with measure data in space

\[
\begin{aligned}
\begin{cases}
y_t + Ay = \sum_{i=1}^{m} u_i(t) \delta_{X_i} & \text{in } \Omega_T, \\
y(x, t) = 0 & \text{on } \Gamma_T, \\
y(x, 0) = y_0(x) & \text{in } \Omega,
\end{cases}
\end{aligned}
\]

(1.6)

where \( \delta_{X_i} \) represents the Dirac measure concentrated at \( X_i, i = 1, 2, \ldots, m \). Some other kind of optimal control problems with state equation of form (1.6), for example, problems with pointwise control, can be found in, e.g., [11], [14] and [25].

The existence of solutions for quasi-linear elliptic and parabolic equations involving measure data has been studied by Boccardo and Gallouët in [4], Casas studied linear parabolic problems and improved the results of [4] by exploiting the linearity of the equation in [6]. The finite element method for elliptic equation with measure data has been extensively studied (see, e.g., [2], [5], [28] and [29]). Casas gave an optimal error estimate of order \( O(h^{2-d}) \) in [5], where \( h \) is the mesh size of space triangulation and \( d \) is the dimension of \( \Omega \). Araya et al. obtained a posteriori error estimates for elliptic problems with Dirac delta source terms in [1]. However, there seems to be no such kind of contributions to finite element approximations of parabolic equations with measure data. To the best of our knowledge this paper is among the few contributions on this topic.

In this paper we study the finite element approximations of parabolic equations with measure data, especially we consider problems with separate measure data in time and space, respectively. We use standard piecewise linear and continuous finite elements for the space discretization and derive the a priori error estimates for the semidiscretization problems, while the backward Euler method is then used for time discretization and a priori error estimates for the fully discrete problems are also derived.

We denote by \( k \) the step size in the temporal discretization and by \( h \) the maximum element size of the spatial mesh. Then the main results of this paper are as follows. For parabolic equations with measure data in time, we obtain the following estimates of the error between the solution \( y \) of the continuous problem and the solution \( y_h \) of the semidiscretization one:

\[
\| y - y_h \|_{L^2(0,T;L^2(\Omega))} \leq Ch \left( \| g \|_{L^\infty(0,T;L^2(\Omega))} \| \omega \|_{\mathcal{M}[0,T]} + \| y_0 \|_{0,\Omega} \right)
\]

and the estimates of the error between the solution \( y \) of the continuous problem and the solution \( Y_h \) of the fully discrete one:

\[
\| y - Y_h \|_{L^2(0,T;L^2(\Omega))} \leq C(h + k^{\frac{1}{2}}) \left( \| g \|_{L^\infty(0,T;L^2(\Omega))} \| \omega \|_{\mathcal{M}[0,T]} + \| y_0 \|_{0,\Omega} \right).
\]

For parabolic equations with measure data in space, we obtain the following estimates of the error between the solution \( y \) of the continuous problem and the solution \( y_h \) of the semidiscretization one:

\[
\| y - y_h \|_{L^2(0,T;L^2(\Omega))} \leq Ch^{2-d} \left( \| g \|_{L^2(0,T;L^\infty(\Omega))} \| \omega \|_{\mathcal{M}(\Omega)} + \| y_0 \|_{0,\Omega} \right)
\]
and the estimates of the error between the solution $y$ of the continuous problem and the solution $Y_h$ of the fully discrete one:
\[
\|y - Y_h\|_{L^2(0,T;L^2(\Omega))} \leq C(h^{2-\frac{d}{2}} + k^2) \left( \|g\|_{H^{\frac{1}{2}}(0,T;L^\infty(\Omega))} \|\omega\|_{\mathcal{M}(\Omega)} + \|y_0\|_{0,\Omega} \right).
\]

Numerical results are provided at the end of the paper to confirm our theoretical findings.

The rest of this paper is organized as follows. In Section 2 we give some notations and present the parabolic equations with measure data in time and space, and analyze the existence results for the unique solution. In Section 3 we establish the continuous time semi-discrete finite element approximation schemes for the above two kinds of problems and derive a priori error estimates. Then the fully discrete finite element approximation based on the backward Euler method is introduced and an a priori estimate for the discretization error is derived in Section 4. We also carry out some numerical experiments in Section 5 to confirm our theoretical findings. At the end of the paper we give a conclusion and discuss the future work.

2. PARABOLIC EQUATIONS WITH MEASURE DATA

2.1. Notation. Assume that $\Omega \subset \mathbb{R}^d$, $d = 2$ or $3$ is a convex polygonal or polyhedral domain, or domain with a $C^{1,1}$ boundary. We denote by $W^{m,p}(\Omega)$ the usual Sobolev space of order $m \geq 0$, $1 \leq p < \infty$ with norm $\| \cdot \|_{m,p,\Omega}$ and seminorm $| \cdot |_{m,p,\Omega}$, and the standard modification for $p = \infty$. For $p = 2$ we denote $W^{m,p}(\Omega)$ by $H^m(\Omega)$ and $\| \cdot \|_{m,\Omega} = \| \cdot \|_{m,2,\Omega}$, which is a Hilbert space. Note that $H^0(\Omega) = L^2(\Omega)$ and $H^1_0(\Omega) = \{ v \in H^1(\Omega) : v = 0 \text{ on } \partial \Omega \}$.

For $p \in [1, \infty)$, the interval $[0, T] \subset \mathbb{R}$ and the Banach space $A$ with norm $\| \cdot \|_A$, we denote by $L^p(0,T;A)$ the set of measurable functions $y : [0, T] \rightarrow A$ such that $\int_0^T \|y(t)\|^p_A dt \leq \infty$. The norm on $L^p(0,T;A)$ is defined by
\[
\|y(t)\|_{L^p(0,T;A)} = \left\{ \begin{array}{ll}
\left( \int_0^T \|y(t)\|^p_A dt \right)^{\frac{1}{p}} & 1 \leq p < \infty, \\
\text{ess sup}_{t \in [0,T]} \|y(t)\|_A & p = \infty.
\end{array} \right.
\]

We denote $\mathcal{D}(\Omega_T)$ the set of $C^\infty(\Omega_T)$ functions with compact support in $\Omega_T$. Let $H^{s,r}(\Omega_T) = L^2(0,T;H^s(\Omega)) \cap H^r(0,T;L^2(\Omega))$ be equipped with the norm
\[
\|w\|_{s,r} = \left( \int_0^T \|w(\cdot,t)\|^2_{L^2} dt + \int_{\Omega} \|w(x,\cdot)\|^2_{L^2(\Omega)} dx \right)^{\frac{1}{2}},
\]
where $\| \cdot \|_{r,[0,T]}$ denotes the norm on $H^r([0,T])$. We set
\[
W(0,T) := L^2(0,T;H^1_0(\Omega)) \cap H^1(0,T;H^{-1}(\Omega)),
\]
it is straightforward that $W(0,T) \hookrightarrow C([0,T];L^2(\Omega))$ (see [19]). We also set
\[
X(0,T) := L^2(0,T;H^2(\Omega) \cap H^1_0(\Omega)) \cap H^1(0,T;L^2(\Omega)) \hookrightarrow C([0,T];H^1_0(\Omega)).
\]

We denote the $L^2$-inner products on $L^2(\Omega)$ and $L^2(\Omega_T)$ by
\[
(v,w) = \int_{\Omega} vw dx \quad \forall v, w \in L^2(\Omega)
\]
and
\[
(v,w)_{\Omega_T} = \int_{\Omega_T} vw dt \quad \forall v, w \in L^2(\Omega_T),
\]
respectively. The operator $A$ is assumed to be a second order elliptic partial differential operator of the form

$$Ay = -\sum_{i,j=1}^{d} \partial_{x_j}(a_{ij}(a_{i}y)) + a_0y,$$

where $a_0 \in L^\infty(\Omega)$, $a_0(x,t) \geq 0$ for all $(x,t) \in \Omega_T$, $a_{ij}$ $(1 \leq i, j \leq d)$ is Lipschitz continuous on $\Omega_T$ and satisfies the following uniform ellipticity condition:

$$\sum_{i,j=1}^{d} a_{ij} \xi_i \xi_j \geq c|\xi|^2, \quad c > 0 \quad \forall \xi \in \mathbb{R}^d, \quad x \in \Omega.$$

Moreover, $\partial_{n}A = \sum_{i,j=1}^{d} a_{ij} \partial_{x_j} n_i$ and $n$ is the unit outer normal to $\partial\Omega$. We will denote by $A^*$ the adjoint operator of $A$:

$$A^*y = -\sum_{i,j=1}^{d} \partial_{x_j}(a_{ji}(a_{i}y)) + a_0y.$$

In addition, $c$ and $C$ denote generic positive constants.

We introduce the following bilinear forms associated with $A$ on $\Omega$ and $\Omega_T$:

$$a(v, w) = \sum_{i,j=1}^{d} \int_{\Omega} (a_{ij} \partial_{x_i} v \partial_{x_j} w + a_0 vw) dx \quad \forall \ v, w \in H^1(\Omega)$$

and

$$a(v, w)_{\Omega_T} = \sum_{i,j=1}^{d} \int_{\Omega_T} (a_{ij} \partial_{x_i} v \partial_{x_j} w + a_0 vw) dx dt \quad \forall \ v, w \in L^2(0, T; H^1(\Omega)).$$

For $f \in L^2(\Omega_T)$, we assume that $\phi$ and $\psi$ are the solutions of following forward and backward in time parabolic problems:

\[ \begin{align*}
\partial_t \phi + A \phi &= f \quad \text{in } \Omega_T, \\
\phi &= 0 \quad \text{on } \Gamma_T, \\
\phi(0) &= 0 \quad \text{in } \Omega
\end{align*} \tag{2.1} \]

and

\[ \begin{align*}
-\partial_t \psi + A^* \psi &= f \quad \text{in } \Omega_T, \\
\psi &= 0 \quad \text{on } \Gamma_T, \\
\psi(T) &= 0 \quad \text{in } \Omega.
\end{align*} \tag{2.2} \]

Then the following standard stability estimates can be found in, e.g., [19].

**Lemma 2.1.** Let $\phi$ and $\psi$ denote the solutions of problem (2.1) and (2.2), respectively. Then for $v = \phi$ or $v = \psi$ there holds $v \in L^2(0, T; H^2(\Omega) \cap H^1_0(\Omega)) \cap H^1(0, T; L^2(\Omega)) \rightarrow C([0, T]; H^1(\Omega))$ and satisfies

\[ \|v\|_{L^2(0, T; H^2(\Omega))} + \|v_t\|_{L^2(0, T; L^2(\Omega))} \leq C\|f\|_{0,0} \tag{2.3} \]

and

\[ \|\phi(T)\|_{1,\Omega} \leq C\|f\|_{L^2(0, T; L^2(\Omega))}, \quad \|\psi(0)\|_{1,\Omega} \leq C\|f\|_{L^2(0, T; L^2(\Omega))}. \tag{2.4} \]
Parabolic equations with measure data in time. At first, we consider the following parabolic equations with measure data in time:

\[
\begin{align*}
\partial_t y + Ay &= g\omega \quad \text{in } \Omega_T, \\
y &= 0 \quad \text{on } \Gamma_T, \\
y(\cdot, 0) &= y_0 \quad \text{in } \Omega,
\end{align*}
\]

(2.5)

where \( g \in C([0, T]; L^2(\Omega)) \) and \( \omega \in \mathcal{M}[0, T] \). The weak solution of problems (2.5) can be defined by transposition techniques (see Lions and Magenes [19]). In the following theorem we will give the results on the existence and uniqueness as well as regularity of the solution to problem (2.5).

**Theorem 2.2.** With the assumption that \( \mu = g\omega, g \) and \( \omega \) are given functions such that \( g \in C([0, T]; L^2(\Omega)) \) and \( \omega \in \mathcal{M}[0, T] \), problem (2.5) admits a unique solution \( y \in L^2(0, T; H^1(\Omega)) \cap L^\infty(0, T; L^2(\Omega)) \) such that

\[(2.6) - (y, \partial_t v)_{H^1, H^{-1}} + a(y, v)_{\Omega_T} = \langle \mu, v \rangle_{\Omega_T} + (y_0, v(\cdot, 0)) \quad \forall \, v \in W(0, T) \]

with \( v(\cdot, T) = 0 \) and

\[
\|y\|_{L^2(0, T; H^1(\Omega))} + \|y\|_{L^\infty(0, T; L^2(\Omega))} \\
\leq C \left( \|g\|_{L^\infty(0, T; L^2(\Omega))} \|\omega\|_{\mathcal{M}[0, T]} + \|y_0\|_{0, \Omega} \right).
\]

Here

\[
\langle \mu, v \rangle_{\Omega_T} = \int_{\Omega_T} vd\mu = \int_0^T \left( \int_{\Omega} g(x, t)v(x, t)dx \right) d\omega(t), \quad \forall \, v \in C([0, T]; L^2(\Omega)).
\]

**Proof.** The proof follows the idea of [6], here we sketch the proof for completeness. Since the problem is linear, it suffices to consider either \( y_0 \equiv 0 \) or \( \mu \equiv 0 \).

If \( \mu \equiv 0 \), \( y_0 \in L^2(\Omega) \), it is obvious that problem (2.5) admits a unique solution \( y \in L^2(0, T; H^1(\Omega)) \cap L^\infty(0, T; L^2(\Omega)) \) satisfying (see Lions and Magenes [19])

\[
\|y\|_{L^2(0, T; H^1(\Omega))} + \|y\|_{L^\infty(0, T; L^2(\Omega))} \leq C\|y_0\|_{0, \Omega}.
\]

Now we suppose \( y_0 \equiv 0 \), let \( \{\omega_n\}_n \subset C[0, T] \) be the sequence converging weakly-* to \( \omega \) in \( \mathcal{M}[0, T] \) and satisfy

\[
\|\omega_n\|_{L^1[0, T]} \leq \|\omega\|_{\mathcal{M}[0, T]}.
\]

Let \( y_n \) be the solutions of

\[
\begin{align*}
\partial_t y_n + Ay_n &= g\omega_n \quad \text{in } \Omega_T, \\
y_n &= 0 \quad \text{on } \Gamma_T, \\
y_n(\cdot, 0) &= 0 \quad \text{in } \Omega,
\end{align*}
\]

(2.8)

then we have \( y_n \in L^2(0, T; H^2(\Omega) \cap H^1(\Omega)) \cap H^1(0, T; L^2(\Omega)) \). For \( f \in \mathcal{D}(\Omega_T) \), let \( \psi \) be the solution of problem

\[
\begin{align*}
-\partial_t \psi + A^* \psi &= f \quad \text{in } \Omega_T, \\
\psi &= 0 \quad \text{on } \Gamma_T, \\
\psi(\cdot, T) &= 0 \quad \text{in } \Omega,
\end{align*}
\]

(2.9)
thus we have \( \psi \in C(\overline{\Omega_T}) \) from the regularity theory of parabolic equation. Then from \((2.8)\) we have
\[
\int_{\Omega_T} f y_n dx dt = \int_{\Omega_T} (-\partial_t \psi + A^* \psi) y_n dx dt
= \int_{\Omega_T} g \omega_n \psi dx dt
\leq C \|g\|_{L^\infty(0, T; L^2(\Omega))} \|\omega_n\|_{L^1[0, T]} \|\psi\|_{C([0, T]; L^2(\Omega))}
\leq C \|g\|_{L^\infty(0, T; L^2(\Omega))} \|\omega\|_{\mathcal{M}(0, T)} \|\psi\|_{C([0, T]; L^2(\Omega))},
\]
Standard estimates give (see, e.g., \[19\])
\[
\|\psi\|_{C([0, T]; L^2(\Omega))} \leq C \|f\|_{L^1([0, T]; L^2(\Omega))}
\]
and
\[
\|\psi\|_{C([0, T]; L^2(\Omega))} \leq C \|f\|_{L^2(0, T; H^{-1}(\Omega))}.
\]
We can conclude from \((2.10)\) that the solution sequence \( \{y_n\}_n \) is bounded in the space \( L^\infty(0, T; L^2(\Omega)) \), while \( \{y_n\}_n \) is also bounded in the space \( L^2(0, T; H^1_0(\Omega)) \) from \((2.11)\). Thus we can take a subsequence such that \( y_n \rightarrow y \) weakly in \( L^2(0, T; H^1_0(\Omega)) \cap L^\infty(0, T; L^2(\Omega)) \) and \((2.7)\) is satisfied.

Then we prove \((2.6)\). Let \( \psi \in W(0, T) \) and \( \psi(\cdot, T) = 0 \), multiplying \((2.8)\) by \( \psi \) and integrating by parts we have
\[
\int_0^T \left( \int_\Omega g(x, t) \psi(x, t) dx \right) \omega_n(t) dt
= -\int_{\Omega_T} y_n \partial_t \psi dx dt + \int_{\Omega_T} \left( \sum_{i,j=1}^d a_{ij} \partial_x y_n \partial_x \psi + a_0 y_n \psi \right) dx dt,
\]
passing to the limit in \((2.12)\) we get \((2.6)\).

Finally, we note that uniqueness holds since the only solution for zero data of \((2.8)\) is \( y = 0 \).

\[\Box\]

2.3. Parabolic equations with measure data in space. Now we turn to the following parabolic equations with measure data in space:
\[
\begin{cases}
\partial_t y + Ay = \mu = g \omega & \text{in } \Omega_T, \\
y = 0 & \text{on } \Gamma_T, \\
y(0) = y_0 & \text{in } \Omega,
\end{cases}
\]
where \( g \in L^2(0, T; C(\overline{\Omega})) \) and \( \omega \in \mathcal{M}(\Omega) \). Similarly, the weak solution of problem \((2.13)\) can be defined by transposition techniques. The following theorem gives the results concerning the existence, uniqueness and regularity of the solution to problem \((2.13)\).

**Theorem 2.3.** With the assumption that \( y_0 \in L^2(\Omega), \mu = g \omega, g \) and \( \omega \) are given functions such that \( g \in L^2(0, T; C(\overline{\Omega})) \) and \( \omega \in \mathcal{M}(\Omega) \), problem \((2.13)\) admits a unique solution \( y \in L^2(0, T; L^2(\Omega)) \) in the sense that
\[
-(y, \partial_t v)_{\Omega_T} + (y, A^* v)_{\Omega_T} = \langle \mu, v \rangle_{\Omega_T} + (y_0, v(\cdot, 0)) \quad \forall \ v \in X(0, T)
\]
with \( v(\cdot, T) = 0 \), here
\[
\langle \mu, v \rangle_{T} = \int_{\Omega_{T}} vd\mu = \int_{\Omega} \left( \int_{0}^{T} g(x, t)v(x, t)dt \right) d\omega(x) \quad \forall \ v \in L^{2}(0, T; C(\Omega)).
\]

Then, there exist a constant \( C \) only depending on data, such that
\[
\|y\|_{L^{2}(0, T; L^{2}(\Omega))} \leq C \left( \|g\|_{L^{2}(0, T; L^{\infty}(\Omega))} \|\omega\|_{M(\Omega)} + \|y_{0}\|_{0, \Omega} \right).
\]
Moreover, we have
\[
y \in L^{1}(0, T; W^{1, p}(\Omega)) \cap C([0, T]; W^{1, 2, q}(\Omega)) \quad \text{and} \quad \partial_{t}y \in L^{1}(0, T; W^{1, q}(\Omega)'),
\]
and
\[
\|y\|_{L^{1}(0, T; W^{1, p}(\Omega))} \leq C \left( \|g\|_{L^{2}(0, T; L^{\infty}(\Omega))} \|\omega\|_{M(\Omega)} + \|y_{0}\|_{0, \Omega} \right),
\]
where \( p \in \left[1, \frac{d}{d-1}\right) \) and \( q \) is the conjugate number of \( p \) such that \( \frac{1}{p} + \frac{1}{q} = 1 \).

Proof. Similarly, as in the proof of Theorem 2.2 we assume that \( y_{0} \in L^{2}(\Omega) \) and \( g \in L^{2}(0, T; C(\Omega)) \). For \( \mu \equiv 0 \) the assertion is obvious. We set \( y_{0} \equiv 0 \) and let \( \{\omega_{n}\}_{n} \subset C(\Omega) \) be the sequence converging weakly-* to \( \omega \) in \( M(\Omega) \) and satisfy
\[
\|\omega_{n}\|_{L^{1}(\Omega)} \leq \|\omega\|_{M(\Omega)}.
\]
Let \( y_{n} \) be the solutions of (2.8) with right-hand side \( g\omega_{n} \), then we have \( y_{n} \in X(0, T) \).

For \( f \in D(\Omega_{T}) \), let \( \psi \) be the solution of problem (2.9), thus we have \( \psi \in C(\Omega_{T}) \) from the regularity theory of parabolic equation. We deduce from (2.8) that
\[
\int_{\Omega_{T}} fy_{n}dxdt = \int_{\Omega_{T}} (-\partial_{t}\psi + A^{\ast}\psi) y_{n}dxdt
\]
\[
= \int_{\Omega_{T}} g\omega_{n}\psi dxdt
\]
\[
\leq C \|g\|_{L^{2}(0, T; C(\Omega))} \|\omega_{n}\|_{L^{1}(\Omega)} \|\psi\|_{L^{2}(0, T; C(\Omega))}
\]
\[
\leq C \|g\|_{L^{2}(0, T; C(\Omega))} \|\omega\|_{M(\Omega)} \|\psi\|_{L^{2}(0, T; C(\Omega))}.
\]

From embedding theorem we have \( L^{2}(0, T; H^{2}(\Omega)) \hookrightarrow L^{2}(0, T; C(\Omega)) \). Standard estimates give (see, e.g., [19])
\[
\|\psi\|_{L^{2}(0, T; H^{2}(\Omega))} \leq C \|f\|_{L^{2}(0, T; L^{2}(\Omega))}.
\]
We can conclude from (2.14) that the solution sequence \( \{y_{n}\}_{n} \) is bounded in the space \( L^{2}(0, T; L^{2}(\Omega)) \). Thus we can take a subsequence such that \( y_{n} \rightarrow y \) weakly in \( L^{2}(0, T; L^{2}(\Omega)) \) and (2.15) is satisfied. The rest of the proof is standard.

Furthermore, the second part of this theorem has been proved in [6]; see also [21]. Actually, Theorem 6.3 in [6] implies the existence of a unique solution \( y \in L^{1}(0, T; W^{1, p}(\Omega)) \) for all \( p \in \left[1, \frac{d}{d-1}\right) \) and \( \partial_{t}y \in L^{1}(0, T; W^{2, q}(\Omega)') \) in the sense of (2.14), such that (2.16) is satisfied, hence we have \( y \in C([0, T]; W^{1, q}(\Omega)') \) after a modification on a set of zero measure. \( \square \)

Remark 2.4. If \( d = 1 \), the function \( \mu = g(x, t)\omega(x) \) belongs to \( L^{2}(0, T; H^{-1}(\Omega)) \), this property implies in turn that (see [14])
\[
y \in L^{2}(0, T; H^{1}(\Omega)) \cap C([0, T]; L^{2}(\Omega)), \quad \frac{\partial y}{\partial t} \in L^{2}(0, T; H^{-1}(\Omega)).
\]
However, we do not proceed with this issue and consider only cases when \( d=2 \) or 3.
3. Error estimates for the continuous time Galerkin approximations

Let us consider the continuous time finite element approximations of the problems (2.6) and (2.14). To this aim, we consider a family of triangulation $\mathcal{T}^h$ of $\overline{\Omega}$, such that $\overline{\Omega} = \bigcup_{\tau \in \mathcal{T}^h} \overline{\tau}$. We suppose that $\overline{\Omega}$ is the union of the elements of $\mathcal{T}^h$ so that element edges lying on the boundary may be curved. This triangulation is supposed to be regular in the usual sense. For each element $\tau \in \mathcal{T}^h$ we associate two parameters $\rho(\tau)$ and $\sigma(\tau)$, where $\rho(\tau)$ denotes the diameter of the element $\tau$ and $\sigma(\tau)$ is the supremum of the diameters of all circles contained in $\tau$. Define the size of the mesh by $h = \max_{\tau \in \mathcal{T}^h} \rho(\tau)$. We suppose that the following regularity assumptions are satisfied: There exists a positive constant $C$ such that

$$\frac{\rho(\tau)}{\sigma(\tau)} \leq C, \quad \frac{h}{\rho(\tau)} \leq C$$

hold for all $\tau \in \mathcal{T}^h$ and all $h > 0$.

Here we consider only $n$-simplex elements, as they are among the most widely used ones. Associated with $\mathcal{T}^h$ is a finite dimensional subspace $V^h$ of $C(\Omega)$, such that $|\chi|_s$ are polynomials of order $m$ ($m \geq 1$) for $\forall \chi \in V^h$ and $\tau \in \mathcal{T}^h$. Here we only consider piecewise linear elements, i.e., $m = 1$. We also set $V^h_0 = V^h \cap H^1_0(\Omega)$.

Note that the regular assumption (3.1) guarantees the following inverse properties for $v_h \in V^h$:

$$\|v_h\|_{s,\Omega} \leq Ch^{1-s}\|v_h\|_{l,\Omega}, \quad 0 \leq l \leq s \leq 1$$

and

$$\|v_h\|_{0,\infty,\Omega} \leq Ch^{-\frac{d}{2}}\|v_h\|_{0,\Omega},$$

$$\|v_h\|_{0,\infty,\Omega} \leq C\rho(d,h)\|v_h\|_{1,\Omega},$$

where

$$\rho(d,h) = \begin{cases} \sqrt{\log h}, & d = 2; \\ h^{-\frac{d}{2}}, & d = 3. \end{cases}$$

Let $\Pi_h : C(\overline{\Omega}) \to V^h$ denote the standard Lagrange interpolation operator, then interpolation error estimate implies that for $y \in H^2(\Omega)$ (see, e.g., [8])

$$\|y - \Pi_h y\|_{0,\Omega} + h\|y - \Pi_h y\|_{1,\Omega} \leq C h^2\|y\|_{2,\Omega}$$

and

$$\|y - \Pi_h y\|_{0,\infty,\Omega} \leq C h^{2-\frac{d}{2}}\|y\|_{2,\Omega}.$$ 

Let $\mathcal{P}_h$ be the $L^2(\Omega)$-projection operator defined from $L^2(\Omega)$ to $V^h$:

$$\langle \mathcal{P}_h y, v_h \rangle = \langle y, v_h \rangle \quad \forall \, v_h \in V^h$$

and $\mathcal{R}_h : H^1_0(\Omega) \to V^h_0$ denote the Ritz projection operator defined as

$$\mathcal{R}_h y = a(y, v_h) \quad \forall \, v_h \in V^h_0.$$ 

Then we have the following error estimates (see, e.g., [8] and [26])

**Lemma 3.1.** Let $\mathcal{P}_h$ and $\mathcal{R}_h$ be the $L^2$-projection operator and Ritz projection operator defined above. Then there holds:

$$\|y - \mathcal{P}_h y\|_{-1,\Omega} + h\|y - \mathcal{P}_h y\|_{0,\Omega} \leq C h^2\|y\|_{1,\Omega},$$

$$\|y - \mathcal{R}_h y\|_{0,\Omega} + h\|y - \mathcal{R}_h y\|_{1,\Omega} \leq C h^2\|y\|_{2,\Omega}.$$
Moreover, we have

\[(3.12)\quad \|y - \mathcal{R}_h y\|_{0, \infty, \Omega} \leq C h^{2-\frac{d}{2}} \|y\|_{2, \Omega}.
\]

**Proof.** Here we only prove \(\|y - \mathcal{P}_h y\|_{-1, \Omega}\). From the definition of \(L^2\)-projection we have

\[
\|y - \mathcal{P}_h y\|_{-1, \Omega} = \sup_{v \in H^1(\Omega)} \frac{(y - \mathcal{P}_h y, v)}{\|v\|_{1, \Omega}}
= \sup_{v \in H^1(\Omega)} \frac{(y - \mathcal{P}_h y, v - \mathcal{P}_h v)}{\|v\|_{1, \Omega}}
\leq \sup_{v \in H^1(\Omega)} Ch^2 \|y\|_{1, \Omega} \|v\|_{1, \Omega}
\leq Ch^2 \|y\|_{1, \Omega}.
\]

Since the solutions of problems (2.6) and (2.14) have low regularities, it seems to be natural to estimate the error between the solutions of continuous problem and semidiscretization problem under the norm \(L^2(0, T; L^2(\Omega))\). To achieve this we need to use the duality argument. Thus, we introduce the semi-discrete finite element approximation of the backward parabolic problem (2.2):

\[(3.13)\quad \begin{cases}
-(\partial_t \psi_h, v_h)_{\Omega_T} + a(\psi_h, v_h)_{\Omega_T} = (f, v_h)_{\Omega_T} & \forall v_h \in V^h_0, \\
(\psi_h(T), w_h) = 0 & \forall w_h \in V^h_0,
\end{cases}
\]

where \(\psi_h(t) \in H^1(0, T; V^h_0)\).

At first, we need to derive the error estimates for the solutions of the backward parabolic problem (2.2) and its semidiscretization approximation (3.13) under the norms \(L^2(0, T; L^\infty(\Omega))\) and \(L^\infty(0, T; L^2(\Omega))\), which will play a crucial role in the derivation of our main results.

**Lemma 3.2.** Let \(\psi \in X(0, T) \rightarrow C([0, T]; H^1(\Omega))\) and \(\psi_h \in H^1(0, T; V^h_0)\) be the solutions of problem (2.2) and (3.13), respectively. Then we have the following uniformly in time and space error estimate:

\[(3.14)\quad \|\psi - \psi_h\|_{L^\infty(0, T; L^2(\Omega))} \leq Ch \left(\|\psi\|_{L^2(0, T; H^2(\Omega))} + \|\psi_t\|_{L^2(0, T; L^2(\Omega))}\right)
\]

and

\[(3.15)\quad \|\psi - \psi_h\|_{L^2(0, T; L^\infty(\Omega))} \leq Ch^{2-\frac{d}{2}} \left(\|\psi\|_{L^2(0, T; H^2(\Omega))} + \|\psi_t\|_{L^2(0, T; L^2(\Omega))}\right).
\]

**Proof.** Actually, we need to prove the error estimates under the regularity conditions stated in this lemma, and the techniques are different from the standard proof for the semidiscrete error estimates which requires higher regularity; see, e.g., [30]. Using similar arguments as in [7] it is not difficult to prove the following a priori error estimates for the backward parabolic equations

\[(3.16)\quad \|\psi(t) - \psi_h(t)\|_{0, \Omega} + \|\psi - \psi_h\|_{L^2(0, T; H^1(\Omega))} \leq Ch \left(\|\psi\|_{L^2(0, T; H^2(\Omega))} + \|\psi_t\|_{L^2(0, T; L^2(\Omega))}\right)
\]

and

\[(3.17)\quad \|\psi - \psi_h\|_{L^2(0, T; L^2(\Omega))} \leq C h^2 \left(\|\psi\|_{L^2(0, T; H^2(\Omega))} + \|\psi_t\|_{L^2(0, T; L^2(\Omega))}\right).
\]
The key ingredient of the proof is the introduction of the $L^2$-projection instead of the Ritz-projection used in the other literatures.

Error estimate (3.14) is a direct consequence of (3.16). To prove (3.15), let $\Pi_h \psi$ be the piecewise linear interpolation of $\psi$ defined above, then from (3.17) we have

\[
\begin{align*}
\| \psi - \psi_h \|_{L^2(0,T;L^\infty(\Omega))} &\leq \| \psi - \Pi_h \psi \|_{L^2(0,T;L^\infty(\Omega))} + \| \Pi_h \psi - \psi_h \|_{L^2(0,T;L^\infty(\Omega))} \\
&\leq C h^{-\frac{d}{2}} \| \psi \|_{L^2(0,T;H^2(\Omega))} + C h^{-\frac{d}{2}} \| \Pi_h \psi - \psi_h \|_{L^2(0,T;L^2(\Omega))} \\
&\leq C h^{-\frac{d}{2}} \| \psi \|_{L^2(0,T;H^2(\Omega))} + C h^{-\frac{d}{2}} \| \psi - \psi_h \|_{L^2(0,T;L^2(\Omega))},
\end{align*}
\]

where we have used the standard interpolation error estimate and inverse estimate (3.3). This together with (3.17) implies (3.15). 

3.1. Finite element approximations to parabolic equations with measure data in time. We now turn to defining the continuous time finite element approximation scheme for problems (2.6). Based on the weak form stated in Theorem 2.2, we can define the following semi-discrete finite element approximation of (2.6):

\[
- \langle y_h, \partial_t v_h \rangle_{\Omega_T} + a(y_h, v_h)_{\Omega_T} = \langle \mu, v_h \rangle_{\Omega_T} + \langle y_0^h, v_h(\cdot, 0) \rangle \quad \forall v_h \in H^1(0,T;V_0^h)
\]

where $y_0^h \in V_0^h$ is an approximation of $y_0$. We set $y_0^h = P_h y_0$ to be the $L^2$-projection of $y_0$. Here

\[
\langle \mu, v_h \rangle_{\Omega_T} = \int_{\Omega_T} v_h d\mu = \int_0^T \left( \int_{\Omega} g(x,t)v_h(x)dx \right) d\omega(t) \quad \forall v_h \in V^h.
\]

Now we are in a position to state our main result of this subsection, i.e., the estimates of the error between the solution $y$ of the continuous problem (2.6) and the solution $y_h$ of the semidiscretization one (3.19).

**Theorem 3.3.** Assume that $\mu = g \omega$, $g$ and $\omega$ are given functions such that $g \in C([0,T];L^2(\Omega))$ and $\omega \in M([0,T])$. Let $y \in L^2(0,T;H^1_0(\Omega)) \cap L^\infty(0,T;L^2(\Omega))$ and $y_h \in L^2(0,T;V_0^h)$ be the solutions of problem (2.6) and (3.19), respectively. Then we have the following error estimate:

\[
\| y - y_h \|_{L^2(0,T;L^2(\Omega))} \leq C h \left( \| g \|_{L^\infty(0,T;L^2(\Omega))} \| \omega \|_{M([0,T])} + \| y_0\|_{0,\Omega} \right).
\]

**Proof.** Let $\psi$ be the solution of problem (2.2) with $f \in L^2(0,T;L^2(\Omega))$. Then from (2.6), (3.19) and orthogonality property we have

\[
\begin{align*}
\int_{\Omega_T} (y - y_h) f dx dt &= \int_0^T \int_{\Omega} (y - y_h)(-\partial_t \psi + A^* \psi) dx dt \\
&= \int_0^T \left( \int_{\Omega} (y - y_h)(-\partial_t \psi + A^* \psi) dx \right) dt \\
&= \int_0^T \left( \int_{\Omega} \left( -\partial_t \psi + A^* \psi \right) dx \right) dt \\
&= \int_0^T \left( \int_{\Omega} \left( -\partial_t \psi + A^* \psi \right) dx \right) dt \\
&= \int_0^T \left( \int_{\Omega} \left( -\partial_t \psi + A^* \psi \right) dx \right) dt \\
&\leq C \left( \| g \|_{L^\infty(0,T;L^2(\Omega))} \| \omega \|_{M([0,T])} + \| y_0 \|_{0,\Omega} \right) \| \psi - \psi_h \|_{L^\infty(0,T;L^2(\Omega))}.
\end{align*}
\]

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use
Lemma 2.1 and Lemma 3.2 yield
\[ \int_{\Omega_T} (y - y_h) f dx dt \leq C \left( \| g \|_{L^0(0,T;L^2(\Omega))} \| \omega \|_{\mathcal{M}[0,T]} + \| y_0 \|_{0,\Omega} \right) \| \psi - \psi_h \|_{L^0(0,T;L^2(\Omega))} \]
\[ \leq Ch \left( \| g \|_{L^0(0,T;L^2(\Omega))} \| \omega \|_{\mathcal{M}[0,T]} + \| y_0 \|_{0,\Omega} \right) \| \psi \|_{L^2(0,T;H^2(\Omega))} \]
\[ + \| \partial_t \psi \|_{L^2(0,T;L^2(\Omega))} \]
\[ \leq Ch \left( \| g \|_{L^0(0,T;L^2(\Omega))} \| \omega \|_{\mathcal{M}[0,T]} + \| y_0 \|_{0,\Omega} \right) \| f \|_{L^2(0,T;L^2(\Omega))}, \]
then from the definition of $L^2(\Omega_T)$ norm we have
\[ \| y - y_h \|_{L^2(0,T;L^2(\Omega))} = \sup_{f \in L^2(0,T;L^2(\Omega)), f \neq 0} \frac{(f, y - y_h)_{\Omega_T}}{\| f \|_{L^2(0,T;L^2(\Omega))}} \]
\[ \leq Ch \left( \| y_0 \|_{0,\Omega} + \| g \|_{L^0(0,T;L^2(\Omega))} \| \omega \|_{\mathcal{M}[0,T]} \right), \]
which completes the proof. \(\square\)

3.2. Finite element approximations to parabolic equations with measure data in space. Similarly, based on the results of Theorem 2.3, we can define the semi-discrete finite element approximation of (2.14) as follows:

\[ -(y_h, \partial_t v_h)_{\Omega_T} + a(y_h, v_h)_{\Omega_T} = \langle \mu, v_h \rangle_{\Omega_T} + (y_0^h, v_h(\cdot, 0)) \quad \forall \ v_h \in H^1(0, T; V_0^h) \]

with $v_h(\cdot, T) = 0$, where $y_0(t) \in L^2(0, T; V_0^h)$, and $y_0^h \in V_0^h$ is an approximation of $y_0$. We set $y_0^h = P_h y_0$ to be the $L^2$-projection of $y_0$. Here
\[ \langle \mu, v_h \rangle_{\Omega_T} = \int_{\Omega_T} v_h d\mu = \int_0^T \left( \int_{\Omega} g(x, t) v_h(x) d\omega(x) \right) dt \quad \forall \ v_h \in V^h. \]

With the above preparations we now can state our main result in the following theorem.

Theorem 3.4. Assume that $g$ and $\omega$ are given functions such that $g \in L^2(0, T; C(\overline{\Omega}))$ and $\omega \in \mathcal{M}(\Omega)$, $y_0 \in L^2(\Omega)$. Let $y \in L^2(0, T; L^2(\Omega))$ and $y_h \in L^2(0, T; V_0^h)$ be the solutions of problems (2.13) and (3.21), respectively. Then we have the following error estimate:

\[ \| y - y_h \|_{L^2(0,T;L^2(\Omega))} \leq Ch^{2/4} \left( \| g \|_{L^2(0,T;L^\infty(\Omega))} \| \omega \|_{\mathcal{M}(\Omega)} + \| y_0 \|_{0,\Omega} \right). \]

Proof. Let $\psi$ be the solution of problem (2.22) with $f \in L^2(0, T; L^2(\Omega))$. Then from (2.14) and (3.21), orthogonality property and proceeding as in the proof of Theorem 3.3 we have
\[ \int_{\Omega_T} (y - y_h) f dx dt = \int_0^T \int_{\Omega} (y - y_h)(-\partial_t \psi + A^* \psi) dx dt \]
\[ = \int_{\Omega_T} (\psi - \psi_h) d\mu + (y_0, \psi(\cdot, 0) - \psi_h(\cdot, 0)) \]
\[ = \int_0^T \left( \int_{\Omega} g(x, t)(\psi - \psi_h) d\omega(x) \right) dt + (y_0, \psi(\cdot, 0) - \psi_h(\cdot, 0)) \]
\[ \leq C \left( \| g \|_{L^2(0,T;L^\infty(\Omega))} \| \omega \|_{\mathcal{M}(\Omega)} \| \psi - \psi_h \|_{L^2(0,T;L^\infty(\Omega))} \right. \]
\[ + \left. \| y_0 \|_{0,\Omega} \| \psi - \psi_h \|_{C([0,T];L^2(\Omega))} \right). \]
It now follows from Lemma 2.1 and Lemma 3.2 that
\[
\int_{\Omega_T} (y - y_h) f dx dt \leq C \left( h^{2 - \frac{d}{2}} \| g \|_{L^2(0,T;L^\infty(\Omega))} \| \omega \|_{M(\Omega)} + Ch \| y_0 \|_{\Omega} \right)
\]
\[
\left( \| \psi \|_{L^2(0,T;H^1(\Omega))} + \| \partial_t \psi \|_{L^2(0,T;L^2(\Omega))} \right)
\]
\[
\leq C h^{2 - \frac{d}{2}} \left( \| g \|_{L^2(0,T;L^\infty(\Omega))} \| \omega \|_{M(\Omega)} + \| y_0 \|_{\Omega} \right) \| f \|_{L^2(0,T;L^2(\Omega))},
\]
which proves (3.22) from the definition of \(L^2(\Omega_T)\) norm.

Remark 3.5. The a priori error estimate we obtained in Theorem 3.4, which is of order \(O(h^{2 - \frac{d}{2}})\), seems to be optimal compared with the results presented in [5], where finite element approximation for elliptic equations with measure data is studied and an a priori error estimate of order \(O(h^{2 - \frac{d}{2}})\) is derived.

4. Error estimates for fully discrete finite element approximations

We next consider the fully discrete approximations for above semidiscrete problems by using the backward Euler scheme in time. We consider a partitioning of the time interval \(T = [0, T]\) as
\[
\bar{T} = \{0\} \cup I_1 \cup I_2 \cup \cdots \cup I_N
\]
with subintervals \(I_i = (t_{i-1}, t_i]\) of size \(k_i\) and time points
\[
0 = t_0 < t_1 < \cdots < t_{N-1} < t_N = T.
\]
We define the discretization parameter \(k\) as a piecewise constant function by setting \(k|_{I_i} = k_i\) for \(i = 1, 2, \ldots, N\). For \(i = 1, 2, \ldots, N\), construct the finite element spaces \(V_i^h \in H^1(\Omega)\) (similar to \(V^h\)) with the mesh \(\mathcal{T}_h^i\). For simplicity we consider equal partition in time, i.e., \(k_i \equiv k\), where \(k\) denotes the time step size, and the same finite element space on each time step. For our error analysis, in the following we set \(k = O(h^d)\) throughout the paper.

4.1. Fully discrete approximations of parabolic equations with measure data in time. Now we are in a position to define the fully discrete approximations to parabolic equations with measure data in time. The fully discrete approximation scheme of (3.1) is to find \(Y_i^h \in V_0^h, i = 1, 2, \ldots, N\), such that
\[
\frac{Y_i^h - Y_{i-1}^h}{k}, w_h + a(Y_i^h, w_h) = \langle \mu, w_h \rangle_{L_i}, \forall w_h \in V_0^h, i = 1, \ldots, N,
\]
\[
Y_0^h(x) = y_0^h(x), \quad x \in \Omega.
\]
Here
\[
\langle \mu, v_h \rangle_{L_i} = \frac{1}{k} \int_{\Omega \times (t_{i-1}, t_i]} v_h d\mu = \frac{1}{k} \int_{t_{i-1}}^{t_i} \left( \int_{\Omega} g(x, t)v_h(x)dx \right) d\omega(t), \quad \forall v_h \in V^h.
\]
In the following we denote \(Y_h\) the fully discrete finite element approximation of \(y\); it is obvious that \(Y_h\) is piecewise constant in time and piecewise linear in space on each time interval.

Then we can derive the following stability estimate for numerical scheme (4.1) of problem (2.5).
Lemma 4.1. Let \( Y^i_h \in V^h_i, \; i = 1, 2, \cdots, N \) be the solutions of fully discrete scheme (4.1), \( y^h_0 = \mathcal{P}_h y_0 \) and assume that \( k \leq Ch^d \), then there exists a constant \( C \) independent of \( h, k \) and the data \((g, \omega, y_0)\) such that

\[
\sum_{i=1}^{N} \|Y^i_h - Y^{i-1}_h\|_{0, \Omega}^2 + k\|Y^N_h\|_{1, \Omega}^2 \leq C (\|y_0\|^2_{0, \Omega} + \|g\|_{L^\infty(0,T;L^2(\Omega))}^2 \|\omega\|^2_{\mathcal{M}[0,T]})
\]

and

\[
\|Y^N_h\|_{0, \Omega}^2 + \sum_{i=1}^{N} k\|Y^i_h\|_{1, \Omega}^2 \leq C (\|y_0\|^2_{0, \Omega} + k^{-1}\|g\|_{L^\infty(0,T;L^2(\Omega))}^2 \|\omega\|^2_{\mathcal{M}[0,T]}).
\]

Proof. Let \( w_h = k(Y^i_h - Y^{i-1}_h) \) in (4.1) we get

\[
(Y^i_h - Y^{i-1}_h, Y^i_h - Y^{i-1}_h) + ka(Y^i_h, Y^i_h - Y^{i-1}_h) = k\langle \mu, Y^i_h - Y^{i-1}_h \rangle_{t_i},
\]

thus we have

\[
\|Y^i_h - Y^{i-1}_h\|_{0, \Omega}^2 + k\|Y^i_h\|_{1, \Omega}^2 \\
\leq ka(Y^i_h, Y^{i-1}_h) + \int_{t_{i-1}}^{t_i} (g(t), Y^i_h - Y^{i-1}_h)d\omega(t) \\
\leq \frac{1}{2} k\|Y^i_h\|_{1, \Omega}^2 + \frac{1}{2} k\|Y^{i-1}_h\|_{1, \Omega}^2 + \|Y^i_h - Y^{i-1}_h\|_{0, \Omega} \int_{t_{i-1}}^{t_i} \|g(t)\|_{0, \Omega}d\omega(t) \\
\leq \frac{1}{2} k\|Y^i_h\|_{1, \Omega}^2 + \frac{1}{2} k\|Y^{i-1}_h\|_{1, \Omega}^2 + C \left( \int_{t_{i-1}}^{t_i} \|g(t)\|_{0, \Omega}d\omega(t) \right)^2 + \frac{1}{2} \|Y^i_h - Y^{i-1}_h\|_{0, \Omega}^2.
\]

Summing the above equations over \( i \) from 1 to \( N \) we obtain

\[
\sum_{i=1}^{N} \|Y^i_h - Y^{i-1}_h\|_{0, \Omega}^2 + k\|Y^N_h\|_{1, \Omega}^2 \leq k\|\mathcal{P}_h y_0\|_{1, \Omega}^2 \\
+ C \sum_{i=1}^{N} \left( \int_{t_{i-1}}^{t_i} \|g(t)\|_{0, \Omega}d\omega(t) \right)^2 \\
\leq k\|\mathcal{P}_h y_0\|_{1, \Omega}^2 + C \left( \int_{0}^{T} \|g(t)\|_{0, \Omega}d\omega(t) \right)^2 \\
\leq C\|y_0\|^2_{0, \Omega} + C\|g\|_{L^\infty(0,T;L^2(\Omega))}^2 \|\omega\|^2_{\mathcal{M}[0,T]}.
\]
where we have used the inverse estimate $k\|P_h y_0\|_{0,\Omega}^2 \leq kh^{-2}\|P_h y_0\|_{0,\Omega}^2 \leq C\|y_0\|_{0,\Omega}^2$. This proves (4.2). Similarly, by setting $v_h = kY_h^i$ in (4.1) we have
\[(Y_h^i - Y_{h-1}^i, Y_h^i) + k\alpha(Y_h^i, Y_h^i) = k\langle \mu, Y_h^i \rangle_{L^2(\Omega)},\]
thus we have
\[
\|Y_h^i\|_{0,\Omega}^2 + k\|Y_h^i\|_{0,\Omega}^2 = (Y_h^i, Y_{h-1}^i) + \int_{t_{i-1}}^{t_i} (g(t), Y_h^i) d\omega(t)
\]
\[
\leq \frac{1}{2}\|Y_h^i\|_{0,\Omega}^2 + \frac{1}{2}\|Y_{h-1}^i\|_{0,\Omega}^2 + \|Y_h^i\|_{1,\Omega}^2 \int_{t_{i-1}}^{t_i} \|g(t)\|_{0,\Omega} d\omega(t)
\]
\[
\leq \frac{1}{2}\|Y_h^i\|_{0,\Omega}^2 + \frac{1}{2}\|Y_{h-1}^i\|_{0,\Omega}^2 + Ck^{-1} \left( \int_{t_{i-1}}^{t_i} \|g(t)\|_{0,\Omega} d\omega(t) \right)^2
\]
\[+ \frac{1}{2} k\|Y_h^i\|_{0,\Omega}^2.
\]
Summing the above equations over $i$ from 1 to $N$ we obtain
\[
\|Y_h^N\|_{0,\Omega}^2 + \sum_{i=1}^N k\|Y_h^i\|_{0,\Omega}^2
\]
\[
\leq ||P_h y_0||_{0,\Omega}^2 + C \sum_{i=1}^N k^{-1} \left( \int_{t_{i-1}}^{t_i} \|g(t)\|_{0,\Omega} d\omega(t) \right)^2
\]
\[
\leq C\|y_0\|_{0,\Omega}^2 + Ck^{-1} \|g\|_{L^\infty(0,T;L^2(\Omega))}^2 \|\omega\|_{L^2(0,T)}^2.
\]
This proves (4.3). \hfill \Box

With the above preparations we are ready to estimate the error between the solution $y$ of continuous problem (2.6) and the solution $Y_h$ of the fully discrete problem (4.1), which is the main result of this paper. Instead of the standard approaches based on Ritz-projection (see, e.g., [30]), we use the duality argument to carry out the error analysis, and the stability results stated in Lemma 4.1 for numerical scheme (4.1) plays an important role.

**Theorem 4.2.** Assume that $\mu = g\omega$, $g$ and $\omega$ are given functions such that $g \in C([0,T];L^2(\Omega))$ and $\omega \in \mathcal{M}(0,T)$. Let $y \in L^2(0,T;H_0^1(\Omega)) \cap L^\infty(0,T;L^2(\Omega))$ be the solution of problem (2.6), and $Y_h$ be the solution of problem (4.1), then we have
\[
\|y - Y_h\|_{L^2(0,T;L^2(\Omega))}
\]
\[
\leq C(h + k^{1/2}) \left( \|g\|_{L^\infty(0,T;L^2(\Omega))} \|\omega\|_{\mathcal{M}(0,T)} + \|y_0\|_{0,\Omega} \right).
\]
Proof. Let $\psi$ be the solution of problem (2.2) with $f \in L^2(0, T; L^2(\Omega))$. It follows from (2.6) that
\[
\int_{\Omega_T} (y - Y_h) f \, dx \, dt = \int_{0}^{T} \int_{\Omega} (y - Y_h) (-\partial_t \psi + A^* \psi) \, dx \, dt
\]
\[
= - (y, \partial_t \psi)_{\Omega_T} + (a(y, \psi)_{\Omega_T} + \sum_{n=1}^{N} \int_{I_n} ((Y^n_h, \partial_t \psi) - a(Y^n_h, \psi)) \, dt
\]
\[
= \langle \mu, \psi \rangle_{\Omega_T} + (y_0, \psi(0)) + \sum_{n=1}^{N} \int_{I_n} (k^{-1}(Y^n_h, \psi^{n-1}) - a(Y^n_h, \psi)) \, dt
\]
\[
= \langle \mu, \psi \rangle_{\Omega_T} + (y_0, \psi(0)) - \sum_{n=1}^{N} \int_{I_n} (k^{-1}(Y^n_h - Y^{n-1}_h, \psi^{n-1}) + a(Y^n_h, \psi)) \, dt
\]
\[
+ (Y^n_h, \psi^N) - (Y^n_h, \psi(0))
\]
with $\psi^n := \psi(\cdot, t_n)$. Note that from (4.1) we have
\[
\sum_{n=1}^{N} (k^{-1}(Y^n_h - Y^{n-1}_h, \overline{R_h} \psi) + a(Y^n_h, \overline{R_h} \psi)) = \sum_{n=1}^{N} \langle \mu, \overline{R_h} \psi \rangle_{I_n},
\]
where $\overline{R_h} \psi \in V^h$ is defined on $I_n$ as
\[
\overline{R_h} \psi = \overline{R_h} \psi^n = \frac{1}{k} \int_{I_n} R_h \psi(\cdot, t) \, dt, \quad n > 0,
\]
and $\overline{R_h} \psi^N = R_h \psi(T)$. Here and in what follows we denote $\overline{\psi}$ the average of $\psi$ in $I_n$ as defined in (4.9) for all $\psi \in L^1(I_n)$. It is straightforward to see that
\[
\int_{I_n} (\psi - \overline{\psi}) \, dt = 0.
\]
Therefore we have
\[
\int_{\Omega_T} (y - Y_h) f \, dx \, dt = \langle \mu, \psi \rangle_{\Omega_T} - \sum_{n=1}^{N} \int_{I_n} \langle \mu, \overline{R_h} \psi \rangle_{I_n}
\]
\[
- \sum_{n=1}^{N} \int_{I_n} (k^{-1}(Y^n_h - Y^{n-1}_h, \psi^{n-1}) + a(Y^n_h, \overline{\psi})) \, dt
\]
\[
+ (y_0 - Y^0_h, \psi(0)) + \sum_{n=1}^{N} \int_{I_n} (k^{-1}(Y^n_h - Y^{n-1}_h, \overline{R_h} \psi) + a(Y^n_h, \overline{R_h} \psi)) \, dt
\]
\[
= - \sum_{n=1}^{N} \int_{I_n} (k^{-1}(Y^n_h - Y^{n-1}_h, \psi^{n-1} - \overline{R_h} \psi) + a(Y^n_h, \overline{\psi} - \overline{R_h} \psi)) \, dt
\]
\[
+ \langle \mu, \psi \rangle_{\Omega_T} - \sum_{n=1}^{N} \int_{I_n} \langle \mu, \overline{R_h} \psi \rangle_{I_n} + (y_0 - Y^0_h, \psi(0))
\]
\[
= E_1 + E_2 + E_3.
\]
Note that we have
\[
|E_2| = \left| \langle \mu, \psi \rangle_{\Omega_T} - \sum_{n=1}^{N} \int_{I_n} \langle \mu, \overline{R_h \psi} \rangle_{I_n} \right|
\]
(4.12)
\[
= \left| \sum_{n=1}^{N} \int_{I_n}^{I_{n-1}} \left( \int_{\Omega} g(x, t)(\psi - \overline{R_h \psi})(x) \, dx \right) \, dw(t) \right|
\]
\[
\leq C \|g\|_{L^\infty(0,T;L^2(\Omega))} \|\omega\|_{M[0,T]} \|\psi - \overline{R_h \psi}\|_{L^\infty(0,T;L^2(\Omega))}.
\]
Note that we have
\[
\|\psi - \overline{R_h \psi}\|_{L^\infty(0,T;L^2(\Omega))}
\]
(4.13)
\[
\leq \|\psi - \overline{\psi}\|_{L^\infty(0,T;L^2(\Omega))} + \|\overline{R_h \psi}\|_{L^\infty(0,T;L^2(\Omega))}
\]
\[
\leq C k^{1/2} \|\psi\|_{H^1(0,T;L^2(\Omega))} + C h \|\overline{\psi}\|_{L^\infty(0,T;H^1(\Omega))}
\]
\[
\leq C (k^{1/2} + h) \|\psi\|_{2,1},
\]
where standard error estimates were used in the above equation. Thus
\[
|E_2| \leq C (k^{1/2} + h) \|\psi\|_{2,1} \|g\|_{L^\infty(0,T;L^2(\Omega))} \|\omega\|_{M[0,T]} \]
(4.14)
\[
\leq C (k^{1/2} + h) \|g\|_{L^\infty(0,T;L^2(\Omega))} \|\omega\|_{M[0,T]} \|f\|_{L^2(0,T;L^2(\Omega))}.
\]
We also have
\[
|E_3| = \left| (y_0 - Y_0^n, \psi(0)) \right| \leq \|y_0 - \mathcal{P}_h y_0\|_{-1,\Omega} \|\psi(0)\|_{1,\Omega}
\]
(4.15)
\[
\leq C h \|y_0\|_{0,\Omega} \|f\|_{L^2(0,T;L^2(\Omega))}.
\]
Then it remains to estimate $E_1$. Since $Y_h^n \in V_0^h$, from the definition of Ritz-projection we have
\[
\int_{I_n} a(Y_h^n, \overline{\psi} - \overline{R_h \psi}) \, dt = 0.
\]
(4.16)
Application of the Cauchy-Schwarz inequality gives
\[
|E_1| = \left| - \sum_{n=1}^{N} \int_{I_n} (k^{-1} (Y_h^n - Y_h^{n-1}, \psi^{n-1} - \overline{R_h \psi}) + a(Y_h^n, \overline{\psi} - \overline{R_h \psi})) \, dt \right|
\]
(4.17)
\[
= \left| - \sum_{n=1}^{N} \int_{I_n} k^{-1} (Y_h^n - Y_h^{n-1}, \psi^{n-1} - \overline{R_h \psi}) \, dt \right|
\]
\[
\leq F_1 \cdot F_2,
\]
where
\[
F_1 = \left( \sum_{n=1}^{N} \|Y_h^n - Y_h^{n-1}\|_{0,\Omega}^2 \right)^{1/2}
\]
and
\[
F_2 = \left( \sum_{n=1}^{N} \|\psi^{n-1} - \overline{R_h \psi}\|_{0,\Omega}^2 \right)^{1/2}.
\]
Lemma 4.1 gives
\[
F_1 \leq C \left( \|y_0\|_{0,\Omega} + \|g\|_{L^\infty(0,T;L^2(\Omega))} \|\omega\|_{M[0,T]} \right).
\]
(4.18)
Standard error estimates yield
\[
\|\psi^{n-1} - \Psi_h^t \psi\|_{0,\Omega} \leq \|\psi^{n-1} - \bar{\psi}\|_{0,\Omega} + \|\bar{\psi} - \Psi_h^t \psi\|_{0,\Omega}
\]
(4.19)
\[
\leq \|\psi^{n-1} - \bar{\psi}\|_{0,\Omega} + C h^2 \|\bar{\psi}\|_{2,\Omega}
\]
and
\[
\|\psi^{n-1} - \bar{\psi}\|_{0,\Omega} \leq k^{\frac{1}{2}} \|\psi\|_{L^2(t_{n-1}, t_n; L^2(\Omega))}.
\]
(4.20)

It is straightforward to show that
\[
\|\bar{\psi}\|_{2,\Omega} \leq k^{-\frac{1}{2}} \|\psi\|_{L^2(t_{n-1}, t_n; H^2(\Omega))}.
\]
(4.21)

Then from (4.19)-(4.21) we can conclude that
\[
|F_2| \leq C \left( \sum_{n=1}^{N} (h^4 \|\bar{\psi}\|_{2,\Omega}^2 + k \|\psi\|_{L^2(t_{n-1}, t_n; L^2(\Omega))}^2) \right)^{\frac{1}{2}}
\]
(4.22)
\[
\leq C \left( \sum_{n=1}^{N} (h^4 k^{n-1} \|\bar{\psi}\|_{L^2(t_{n-1}, t_n; H^2(\Omega))}^2 + k \|\psi\|_{L^2(t_{n-1}, t_n; L^2(\Omega))}^2) \right)^{\frac{1}{2}}
\]
\[
\leq C (h + k^{\frac{1}{2}}) \|\psi\|_{2,1},
\]

thus
\[
|E_1| \leq C (h + k^{\frac{1}{2}}) \|\psi\|_{2,1} \left( \|y_0\|_{0,\Omega} + \|g\|_{L^\infty(0,T; L^2(\Omega))} \|\omega\|_{\cal M}[0,T] \right)
\]
(4.23)
\[
\leq C (h + k^{\frac{1}{2}}) \|f\|_{L^2(0,T; L^2(\Omega))} \left( \|y_0\|_{0,\Omega} + \|g\|_{L^\infty(0,T; L^2(\Omega))} \|\omega\|_{\cal M}[0,T] \right).
\]

It follows from (4.11), (4.14), (4.15) and (4.23) that
\[
\|y - Y_h\|_{L^2(0,T; L^2(\Omega))} = \sup_{f \in L^2(0,T; L^2(\Omega)), f \neq 0} \frac{(f, y - Y_h)_{\Omega_T}}{\|f\|_{L^2(0,T; L^2(\Omega))}}
\]
\[
\leq C (h + k^{\frac{1}{2}}) \left( \|y_0\|_{0,\Omega} + \|g\|_{L^\infty(0,T; L^2(\Omega))} \|\omega\|_{\cal M}[0,T] \right),
\]

which completes the proof. \(\square\)

**Remark 4.3.** In Theorem 3.3 we obtain a priori error estimates of order \(O(h)\) for the semidiscretization finite element approximation of parabolic equations with measure data in time, and the same result with respect to space discretization is derived for fully discrete approximation in Theorem 4.2. The order \(O(h)\) seems to be optimal in view of the regularity of solution \(y\), which belongs to \(L^2(0,T; H^1(\Omega))\) as presented in Theorem 2.2. The convergence order with respect to time discretization is \(O(k^{\frac{1}{2}})\), which should also be optimal.

### 4.2. Fully discrete approximations of parabolic equations with measure data in space.

This subsection is devoted to the fully discrete approximations of parabolic equations with measure data in space. The fully discrete approximation scheme of (4.20) is to find \(Y^i_h \in V^h_0, \ i = 1, 2, \ldots, N,\) such that
\[
\left\{ \begin{array}{l}
\left( \frac{Y^i_h - Y^{i-1}_h}{k}, w_h \right) + a(Y^i_h, w_h) = \langle \mu, w_h \rangle_{I_i}, \quad \forall \ w_h \in V^h_0, \quad i = 1, \ldots, N, \\
Y^0_h(x) = y^0_h(x) \quad x \in \Omega.
\end{array} \right.
\]
(4.24)
Here
\[ \langle \mu, v_h \rangle_{I_i} = \frac{1}{k} \int_{\Omega} v_h d\mu = \frac{1}{k} \int_{t_{i-1}}^{t_i} \left( \int_{\Omega} g(x, t) v_h(x) d\omega(x) \right) dt \quad \forall \, v_h \in V^h. \]

Also we denote \( Y_h \) the fully discrete finite element approximation of \( y \).

Then we can derive the following stability estimate for the numerical scheme (4.24) of problem (2.13).

**Lemma 4.4.** Assume that \( g \) and \( \omega \) are given functions such that \( g \in L^2(0, T; C(\Omega)) \) and \( \omega \in M(\Omega) \), \( y_0 \in L^2(\Omega) \). Let \( Y^i_h \in V^h_0 \), \( i = 1, 2, \ldots, N \) be the solutions of fully discrete scheme (4.24), \( y^0_h = P_h y_0 \) and assume that \( k \leq C h^d \), then there exists a constant \( C \) independent of \( h, k \) and the data \( g, \omega, y_0 \) such that

\[ \sum_{i=1}^{N} \| Y^i_h - Y^{i-1}_h \|_{0, \Omega}^2 + k \| Y^N_h \|_{1, \Omega}^2 \leq C \left( \| y_0 \|_{0, \Omega}^2 + \| g(t) \|_{L^2(0, T; L^\infty(\Omega))}^2 \| \omega \|_{M(\Omega)}^2 \right) \]

and

\[ \| Y^N_h \|_{0, \Omega}^2 + \sum_{i=1}^{N} k \| Y^i_h \|_{1, \Omega}^2 \leq C \left( \| y_0 \|_{0, \Omega}^2 + \rho^2(d, h) \| g(t) \|_{L^2(0, T; L^\infty(\Omega))} \| \omega \|_{M(\Omega)}^2 \right). \]

**Proof.** The proof is similar to Lemma 4.1. Let \( w_h = k (Y^i_h - Y^{i-1}_h) \) in (4.24), we get

\[ (Y^i_h - Y^{i-1}_h, Y^i_h - Y^{i-1}_h) + k a(Y^i_h, Y^i_h - Y^{i-1}_h) = k \langle \mu, Y^i_h - Y^{i-1}_h \rangle_{I_i}, \]

thus we have

\[ \| Y^i_h - Y^{i-1}_h \|_{0, \Omega}^2 + k \| Y^i_h \|_{1, \Omega}^2 \]

\[ \leq k a(Y^i_h, Y^i_h - Y^{i-1}_h) + \int_{t_{i-1}}^{t_i} \left( \int_{\Omega} g(x, t) (Y^i_h - Y^{i-1}_h) d\omega(x) \right) dt \]

\[ \leq \frac{1}{2} k \| Y^i_h \|_{1, \Omega}^2 + \frac{1}{2} k \| Y^{i-1}_h \|_{1, \Omega}^2 + \int_{t_{i-1}}^{t_i} \| Y^i_h - Y^{i-1}_h \|_{0, \infty, \Omega} \| g(t) \|_{0, \infty, \Omega} \| \omega \|_{M(\Omega)} dt \]

\[ \leq \frac{1}{2} k \| Y^i_h \|_{1, \Omega}^2 + \frac{1}{2} k \| Y^{i-1}_h \|_{1, \Omega}^2 + C \| g(t) \|_{L^2(t_{i-1}, t_i; L^\infty(\Omega))} \| \omega \|_{M(\Omega)}^2 \]

\[ + \frac{1}{2} \| Y^i_h - Y^{i-1}_h \|_{0, \Omega}^2, \]

where we have used the following inverse estimate:

\[ \sqrt{k} \| Y^i_h - Y^{i-1}_h \|_{0, \infty, \Omega} \leq C \sqrt{k} h^{-d/2} \| Y^i_h - Y^{i-1}_h \|_{0, \Omega} \]

\[ \leq C \| Y^i_h - Y^{i-1}_h \|_{0, \Omega}. \]
Summing the above equations over $i$ from 1 to $N$ and using the inverse estimate we get

\[
\sum_{i=1}^{N} \|Y_i^h - Y_i^{i-1}\|_{0,\Omega}^2 + k\|Y_i^h\|_{1,\Omega}^2 \\
\leq k\|Y_0^h\|_{1,\Omega}^2 + C\sum_{i=1}^{N} \|g(t)\|_{L^2(t_i-1,t_i;L^\infty(\Omega))}^2 \|\omega\|_{\mathcal{M}(\Omega)}^2 \\
\leq k\|P_hy_0\|_{1,\Omega}^2 + C\|g(t)\|_{L^2(0,T;L^\infty(\Omega))}^2 \|\omega\|_{\mathcal{M}(\Omega)}^2 \\
\leq C\|y_0\|_{0,\Omega}^2 + C\|g(t)\|_{L^2(0,T;L^\infty(\Omega))}^2 \|\omega\|_{\mathcal{M}(\Omega)}^2,
\]

which proves (4.26).

Similarly, by setting $v_h = kY_i^h$ in (4.24) we have

\[
(Y_i^h - Y_i^{i-1}, Y_i^h) + ka(Y_i^h, Y_i^h) = k\langle \mu, Y_i^h \rangle I_i,
\]

which gives

\[
(Y_i^h, Y_i^h) + ka(Y_i^h, Y_i^h) \\
= (Y_i^{i-1}, Y_i^h) + \int_{t_{i-1}}^{t_i} (\int_{\Omega} g(x,t)Y_i^h d\omega(x)) dt \\
\leq \frac{1}{2} \|Y_i^h\|_{0,\Omega} + \frac{1}{2} \|Y_i^{i-1}\|_{0,\Omega} + \int_{t_{i-1}}^{t_i} \|Y_i^h\|_{0,\infty,\Omega} \|g(t)\|_{0,\infty,\Omega} \|\omega\|_{\mathcal{M}(\Omega)} dt \\
\leq \frac{1}{2} \|Y_i^h\|_{0,\Omega} + \frac{1}{2} \|Y_i^{i-1}\|_{0,\Omega} + \frac{1}{2} k\|Y_i^h\|_{1,\Omega}^2 \\
+ C\rho^2(d,h) \|g(t)\|_{L^2(t_{i-1},t_i;L^\infty(\Omega))}^2 \|\omega\|_{\mathcal{M}(\Omega)}^2.
\]

Here we used Young’s inequality and the inverse estimate (3.4). Summation from $i = 1$ to $N$ leads to

\[
\|Y_i^h\|_{0,\Omega}^2 + k \sum_{i=1}^{N} a(Y_i^h, Y_i^h) \leq C\|Y_0^h\|_{0,\Omega}^2 + C\rho^2(d,h) \|g(t)\|_{L^2(0,T;L^\infty(\Omega))}^2 \|\omega\|_{\mathcal{M}(\Omega)}^2 \\
\leq C\|y_0\|_{0,\Omega}^2 + C\rho^2(d,h) \|g(t)\|_{L^2(0,T;L^\infty(\Omega))}^2 \|\omega\|_{\mathcal{M}(\Omega)}^2.
\]

This proves (4.26). \hfill \Box

Now we are in a position to estimate the error between the solutions of problems (2.14) and (4.24), which is one of the main results of this paper.

**Theorem 4.5.** Assume that $g$ and $\omega$ are given functions such that

\[
g \in L^2(0,T;C(\overline{\Omega})) \cap H^\frac{1}{2}(0,T;L^\infty(\Omega)) \quad \text{and} \quad \omega \in \mathcal{M}(\Omega), y_0 \in L^2(\Omega).
\]

Let $y \in L^2(0,T;L^2(\Omega))$ be the solution of problem (2.14), and let $Y_h$ be the solution of problem (4.24), then we have

\[
\|y - Y_h\|_{L^2(0,T;L^2(\Omega))} \leq C(h^{2-\frac{d}{2}} + k^{\frac{1}{2}}) \left( \|g\|_{H^\frac{1}{2}(0,T;L^\infty(\Omega))} \|\omega\|_{\mathcal{M}(\Omega)} + \|y_0\|_{0,\Omega} \right).
\]

\[
(4.29)
\]
Proof. Again, we use duality argument to prove this theorem. Let \( \psi \) be the solution of problem (2.2) with \( f \in L^2(0, T; L^2(\Omega)) \). Note that \( \psi = 0 \) on \( \partial \Omega \), \( \psi^N = \psi(T) = 0 \), it follows from (2.14) that

\[
\int_{\Omega_T} (y - Y_h) f dx dt = \int_0^T \int_{\Omega} (y - Y_h)(-\partial_t \psi + A^* \psi) dx dt
\]

\[
= -(y, \partial_t \psi)_{\Omega_T} + (y, A^* \psi)_{\Omega_T} + \sum_{n=1}^N \int_{I_n} ((Y_h^n, \partial_t \psi) - a(Y_h^n, \psi)) dt
\]

\[
= (\mu, \psi)_{\Omega_T} + (y_0, \psi(\cdot, 0)) + \sum_{n=1}^N \int_{I_n} (k^{-1}(Y_h^n - Y_h^{n-1}) - a(Y_h^n, \psi)) dt
\]

\[
= (\mu, \psi)_{\Omega_T} + (y_0, \psi(\cdot, 0)) - \sum_{n=1}^N \int_{I_n} (k^{-1}(Y_h^n - Y_h^{n-1}) + a(Y_h^n, \psi)) dt
\]

\[
+ (Y_h^N, \psi^N) - (Y_h^0, \psi(\cdot, 0))
\]

\[
= - \sum_{n=1}^N \int_{I_n} (k^{-1}(Y_h^n - Y_h^{n-1}) + a(Y_h^n, \psi)) dt
\]

\[
+ (\mu, \psi)_{\Omega_T} + (y_0 - Y_h^0, \psi(\cdot, 0)).
\]

Note that from (4.24) we have

\[
\sum_{n=1}^N (k^{-1}(Y_h^n - Y_h^{n-1}, \overline{R_h} \psi) + a(Y_h^n, \overline{R_h} \psi)) = \sum_{n=1}^N (\mu, \overline{R_h} \psi)_{I_n},
\]

where \( \overline{R_h} \psi \in V_0^h \) is defined in (4.9). Thus

\[
\int_{\Omega_T} (y - Y_h) f dx dt = (\mu, \psi)_{\Omega_T} - \sum_{n=1}^N \int_{I_n} (\mu, \overline{R_h} \psi)_{I_n}
\]

\[
+ \sum_{n=1}^N \int_{I_n} (k^{-1}(Y_h^n - Y_h^{n-1}, \overline{R_h} \psi) + a(Y_h^n, \overline{R_h} \psi)) dt
\]

\[
- \sum_{n=1}^N \int_{I_n} (k^{-1}(Y_h^n - Y_h^{n-1}, \psi^{n-1}) + a(Y_h^n, \psi)) dt + (y_0 - Y_h^0, \psi(0))
\]

\[
= - \sum_{n=1}^N \int_{I_n} (k^{-1}(Y_h^n - Y_h^{n-1}, \psi^{n-1} - \overline{R_h} \psi) + a(Y_h^n, \overline{\psi} - \overline{R_h} \psi)) dt
\]

\[
+ (\mu, \psi)_{\Omega_T} - \sum_{n=1}^N \int_{I_n} (\mu, \overline{R_h} \psi)_{I_n} + (y_0 - Y_h^0, \psi(0))
\]

\[
= E_1 + E_2 + E_3.
\]
From (4.10) we deduce

\[ |\tilde{E}_2| = \left| \langle \mu, \psi \rangle_{\Omega_T} - \sum_{n=1}^{N} \int_{I_n} \langle \mu, R_h \psi \rangle I_n \right| \]

\[ = \left| \sum_{n=1}^{N} \int_{\Omega} \left( \int_{t_{n-1}}^{t_n} \left( g(x, t) \psi(x, t) - \tilde{g}(x, t) R_h \psi(x, t) \right) dt \right) d\omega(x) \right| \]

\[ \leq \left| \sum_{n=1}^{N} \int_{\Omega} \left( \int_{t_{n-1}}^{t_n} \psi(x, t) \left( g(x, t) - \tilde{g}(x, t) \right) dt \right) d\omega(x) \right| \]

\[ + \left| \sum_{n=1}^{N} \int_{\Omega} \left( \int_{t_{n-1}}^{t_n} \tilde{g}(x, t) \left( \psi - R_h \psi \right)(x, t) dt \right) d\omega(x) \right| \]

\[ \leq C \| g - \tilde{g} \|_{L^2(0,T;L^\infty(\Omega))} \| \omega \|_{M(\Omega)} \| \psi \|_{L^2(0,T;L^\infty(\Omega))} \]

\[ + C \| \tilde{g} \|_{L^2(0,T;L^\infty(\Omega))} \| \omega \|_{M(\Omega)} \| \psi - R_h \psi \|_{L^2(0,T;L^\infty(\Omega))}. \]

Standard error estimates yield

\[ \| g - \tilde{g} \|_{L^2(0,T;L^\infty(\Omega))} \leq C h^{\frac{1}{2}} \| g \|_{H^\frac{1}{2}(0,T;L^\infty(\Omega))} \]

and

\[ \| \psi - R_h \psi \|_{L^2(0,T;L^\infty(\Omega))} \leq C h^{2 - \frac{d}{2}} \| \psi \|_{L^2(0,T;H^2(\Omega))}. \]

Thus we have

\[ |\tilde{E}_2| \leq C \left( h^{\frac{1}{2}} + h^{2 - \frac{d}{2}} \right) \| \omega \|_{M(\Omega)} \| g \|_{H^\frac{1}{2}(0,T;L^\infty(\Omega))} \| f \|_{L^2(0,T;L^2(\Omega))}. \]

Similar to (4.15) we also have

\[ |\tilde{E}_3| \leq C h \| y_0 \|_{0,\Omega} \| f \|_{L^2(0,T;L^2(\Omega))}. \]

Then it remains to estimate \( \tilde{E}_1 \). (4.16) and Cauchy-Schwarz inequality give

\[ |\tilde{E}_1| \leq \tilde{F}_1 \cdot \tilde{F}_2, \]

where

\[ \tilde{F}_1 = \left( \sum_{n=1}^{N} \| Y_h^n - Y_h^{n-1} \|_{0,\Omega}^2 \right)^{\frac{1}{2}} \]

and

\[ \tilde{F}_2 = \left( \sum_{n=1}^{N} \| \psi^{n-1} - R_h \psi \|_{0,\Omega}^2 \right)^{\frac{1}{2}}. \]

From (4.25) of Lemma 4.3 we have

\[ \tilde{F}_1 \leq C \left( \| y_0 \|_{0,\Omega} + \| g \|_{L^2(0,T;L^\infty(\Omega))} \| \omega \|_{M(\Omega)} \right). \]

Similar to (4.22) we can conclude from (4.19), (4.21) that

\[ |\tilde{F}_2| \leq C \left( h^{2 - \frac{d}{2}} + h^{\frac{1}{2}} \right) \| \psi \|_{2,1}, \]
thus

\[
|\bar{E}_1| \leq C(h^{2-\frac{\gamma}{2}} + k^{\frac{1}{2}})\|\psi\|_{L^2(\Omega)} \left(\|y_0\|_{L^2(\Omega)} + \|g\|_{L^2(\Omega)} + \|\omega\|_{L^2(\Omega)}\right)
\]

\[
\leq C(h^{2-\frac{\gamma}{2}} + k^{\frac{1}{2}})\left(\|y_0\|_{L^2(\Omega)} + \|g\|_{L^2(\Omega)} + \|\omega\|_{L^2(\Omega)}\right).
\]

Then from (4.30), (4.34), (4.35) and (4.39) we have

\[
\|y - Y_h\|_{L^2(\Omega)} = \sup_{f \in L^2(\Omega)} \frac{(f, y - Y_h)_{L^2(\Omega)}}{\|f\|_{L^2(\Omega)}} \leq C(h^{2-\frac{\gamma}{2}} + k^{\frac{1}{2}})\left(\|y_0\|_{L^2(\Omega)} + \|g\|_{H^1(\Omega)} + \|\omega\|_{L^2(\Omega)}\right),
\]

which completes the proof.

5. Numerical examples

In this section we will carry out some numerical experiments to confirm our theoretical findings. For the computation the software package AFEPack [17] has been used. To validate the estimates developed in the previous section, we show the convergence order by separating the discretization errors. At first we consider the behavior of the error for a sequence of discretizations with different mesh sizes and fixed time steps. Then we show the behavior of the error for different time steps but a fixed spatial triangulation.

In the following numerical examples, we define an error functional to show the experimental order of convergence by

\[
\text{rate} = \frac{\log E(h_1) - \log E(h_2)}{\log h_1 - \log h_2},
\]

where \(E(h)\) denotes the error on triangulation with mesh size \(h\) or time step \(k\). Then it is easy to see that “rate = \(\gamma\)” means that “error \(= O(h^\gamma)\)”.

5.1. Parabolic equations with measure data in time. At first we consider the following parabolic equation with Dirac right-hand side in time:

\[
\begin{cases}
\partial_t y - \Delta y + g \omega & \text{in } \Omega_T, \\
y = 0 & \text{on } \Gamma_T, \\
y(\cdot, 0) = y_0 & \text{in } \Omega,
\end{cases}
\]

where \(g(x, t) \in C([0, T]; L^2(\Omega))\) and \(\omega(t) \in M[0, T]\). For ease of constructing examples we may admit some additional regular parts to appear in the right-hand side.

Example 5.1. The first example is a modification from the example presented in [22]. Let \(\Omega_T = B(0, 1) \times [0, 1]\), where \(B(0, 1)\) is the unit circle centered at zero with radius 1, \(\gamma \in (0, 1)\) and \(\lambda \in \mathbb{R}\). Let

\[
\epsilon(t) = (e^{-\lambda t} - e^{-\frac{1}{2}}),
\]

we take the exact solution as

\[
y(x, t) = \sin(\pi x^2) \frac{e^{\lambda t}}{\lambda(1 - \gamma)} \begin{cases}
\epsilon(0)^{1-\gamma}, & t \geq \frac{1}{2}, \\
\epsilon(0)^{1-\gamma} - \epsilon(t)^{1-\gamma}, & t < \frac{1}{2}
\end{cases}
\]
Table 1. Error of $y$ for Example 5.1 with $\lambda = 1$.

<table>
<thead>
<tr>
<th>Dofs</th>
<th>$|y - Y_h|_{L^2(\Omega_T)}$</th>
<th>rate</th>
<th>$N$</th>
<th>$|y - Y_h|_{L^2(\Omega_T)}$</th>
<th>rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td>1.805765645861</td>
<td>_</td>
<td>3</td>
<td>0.229100247225</td>
<td>_</td>
</tr>
<tr>
<td>81</td>
<td>0.47953557806</td>
<td>1.9127</td>
<td>9</td>
<td>0.197363792883</td>
<td>0.1357</td>
</tr>
<tr>
<td>289</td>
<td>0.119994521144</td>
<td>1.9988</td>
<td>27</td>
<td>0.131680471328</td>
<td>0.3683</td>
</tr>
<tr>
<td>1089</td>
<td>0.031089301244</td>
<td>1.9127</td>
<td>81</td>
<td>0.080625762381</td>
<td>0.4465</td>
</tr>
<tr>
<td>4225</td>
<td>0.010045681219</td>
<td>1.9988</td>
<td>243</td>
<td>0.048726635543</td>
<td>0.4584</td>
</tr>
<tr>
<td>16641</td>
<td>0.006520318504</td>
<td>0.6236</td>
<td>729</td>
<td>0.029977569575</td>
<td>0.4422</td>
</tr>
</tbody>
</table>

Table 2. Error of $y$ for Example 5.1 with $\lambda = 2$.

<table>
<thead>
<tr>
<th>Dofs</th>
<th>$|y - Y_h|_{L^2(\Omega_T)}$</th>
<th>rate</th>
<th>$N$</th>
<th>$|y - Y_h|_{L^2(\Omega_T)}$</th>
<th>rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td>2.405100524743</td>
<td>_</td>
<td>3</td>
<td>0.283293492668</td>
<td>_</td>
</tr>
<tr>
<td>81</td>
<td>0.636075956291</td>
<td>1.9188</td>
<td>9</td>
<td>0.157102357581</td>
<td>0.5367</td>
</tr>
<tr>
<td>289</td>
<td>0.158595488032</td>
<td>2.0038</td>
<td>27</td>
<td>0.106189907672</td>
<td>0.3565</td>
</tr>
<tr>
<td>1089</td>
<td>0.041037524463</td>
<td>1.9503</td>
<td>81</td>
<td>0.067208827352</td>
<td>0.4164</td>
</tr>
<tr>
<td>4225</td>
<td>0.013412438652</td>
<td>1.6134</td>
<td>243</td>
<td>0.041682864989</td>
<td>0.4348</td>
</tr>
<tr>
<td>16641</td>
<td>0.008934235708</td>
<td>0.5862</td>
<td>729</td>
<td>0.026138785274</td>
<td>0.4248</td>
</tr>
</tbody>
</table>

After a simple calculation we have

$$
\mu(x, t) = \sin(\pi|x|^2)\delta(t) + (\sin(\pi|x|^2) \cdot \frac{e^{\lambda t}}{1 - \gamma} + (-4\pi \cos(\pi|x|^2))
+ 4\pi^2|x|^2 \sin(\pi|x|^2) \cdot \frac{e^{\lambda t}}{\lambda(1-\gamma)}) \cdot
\begin{cases}
\epsilon(0)^{1-\gamma}, & t \geq \frac{1}{2}, \\
\epsilon(0)^{1-\gamma} - \epsilon(t)^{1-\gamma}, & t < \frac{1}{2},
\end{cases}
$$

where

$$
\delta(t) = \begin{cases}
0, & t \geq \frac{1}{2}, \\
\epsilon(t)^{-\gamma}, & t < \frac{1}{2}.
\end{cases}
$$

To confirm our theoretical results we test the convergence order with respect to space discretization and time discretization, respectively. To investigate the convergence order with respect to the space discretization we fixed the time discretization with $N = 33333$ for $\lambda = 1, 2$ and $N = 3333$ for $\lambda = 10$, while the space discretization is fixed with 16641 Dofs to investigate the convergence order with respect to the time discretization, the results for $\lambda = 1$, $\lambda = 2$ and $\lambda = 10$ are listed in Tables 1, 2 and 3, respectively.

We can see from Tables 1–3 that the convergence order with respect to the spatial discretization is almost 2, which is optimal and better than our predicted order $O(h)$. While the convergence order with respect to the time discretization is almost $O(k^{1/2})$ for the cases $\lambda = 1$ and $\lambda = 2$, which is consistent with our theoretical results. The convergence order for the time discretization is $O(k)$ for larger $\lambda$, as presented in Table 3.

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use
Example 5.2. The second example is constructed inspired by Example 4.2 of [9]. Let $\Omega_T = B(0,1) \times [0,1]$, and we take the exact solution as
\[ y(x,t) = \sin(\pi|x|^2) \cdot \{ t^2, \quad t < 0.5, \]
\[ \quad t^2 + 2t, \quad t \geq 0.5. \]

We know that $\partial_t y = \sin(\pi|x|^2) \cdot \delta_t(\frac{1}{2}) + \sin(\pi|x|^2) \cdot \gamma(t)$, where $\delta_t(z)$ denotes the Dirac measure with respect to the variable $t$ concentrated at $t = z$, and
\[ \gamma(t) = \begin{cases} 2t, & t < 0.5, \\ 2t + 2, & t \geq 0.5. \end{cases} \]

Thus, after simple calculation we have
\[ \mu(x,t) = \sin(\pi|x|^2) \cdot \delta_t(\frac{1}{2}) + \sin(\pi|x|^2) \cdot \gamma(t) \]
\[ + (-4\pi \cos(\pi|x|^2) + 4\pi^2|x|^2 \sin(\pi|x|^2))) \cdot \begin{cases} t^2, & t < 0.5, \\ t^2 + 2t, & t \geq 0.5. \end{cases} \]

At first we fixed the time discretization with 4096 time steps to investigate the behavior of error with respect to the spatial discretization, then the space discretization is fixed with 16641 Dofs to investigate the convergence order with respect to the time discretization. The results are listed in Table 3. From Table 3 we found that the convergence orders with respect to the spatial and time discretization are $O(h^2)$ and $O(k)$, respectively, both of them are higher than our predicted results which are $O(h)$ and $O(k^{1/2})$, respectively.

### Table 3. Error of $y$ for Example 5.1 with $\lambda = 10$.

<table>
<thead>
<tr>
<th>Dofs</th>
<th>$|y - Y_h|_{L^2(\Omega_T)}$</th>
<th>rate</th>
<th>$N$</th>
<th>$|y - Y_h|_{L^2(\Omega_T)}$</th>
<th>rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td>788.214049388618</td>
<td>_</td>
<td>3</td>
<td>2421.450238197359</td>
<td>_</td>
</tr>
<tr>
<td>81</td>
<td>200.611180708993</td>
<td>1.9742</td>
<td>9</td>
<td>614.302084069153</td>
<td>1.2485</td>
</tr>
<tr>
<td>289</td>
<td>48.311558519598</td>
<td>2.0540</td>
<td>27</td>
<td>162.023290177949</td>
<td>1.2131</td>
</tr>
<tr>
<td>1089</td>
<td>11.935206551992</td>
<td>2.0171</td>
<td>81</td>
<td>48.670234674881</td>
<td>1.0947</td>
</tr>
<tr>
<td>4225</td>
<td>2.934940200289</td>
<td>2.0238</td>
<td>243</td>
<td>14.900101550467</td>
<td>1.0774</td>
</tr>
<tr>
<td>16641</td>
<td>0.885011153476</td>
<td>1.7296</td>
<td>729</td>
<td>5.036596568855</td>
<td>0.9873</td>
</tr>
</tbody>
</table>

### Table 4. Error of $y$ for Example 5.2.

<table>
<thead>
<tr>
<th>Dofs</th>
<th>$|y - Y_h|_{L^2(\Omega_T)}$</th>
<th>rate</th>
<th>$N$</th>
<th>$|y - Y_h|_{L^2(\Omega_T)}$</th>
<th>rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td>1.003408828629</td>
<td>_</td>
<td>2</td>
<td>0.067209558965</td>
<td>_</td>
</tr>
<tr>
<td>81</td>
<td>0.265418710871</td>
<td>1.9186</td>
<td>4</td>
<td>0.033600480166</td>
<td>1.0002</td>
</tr>
<tr>
<td>289</td>
<td>0.065958680971</td>
<td>2.0086</td>
<td>8</td>
<td>0.016860913555</td>
<td>0.9948</td>
</tr>
<tr>
<td>1089</td>
<td>0.016567151789</td>
<td>1.9932</td>
<td>16</td>
<td>0.008466148990</td>
<td>0.9939</td>
</tr>
<tr>
<td>4225</td>
<td>0.004146042659</td>
<td>1.9985</td>
<td>32</td>
<td>0.002276249156</td>
<td>0.9818</td>
</tr>
<tr>
<td>16641</td>
<td>0.001034571264</td>
<td>2.0027</td>
<td>64</td>
<td>0.002276249156</td>
<td>0.9132</td>
</tr>
</tbody>
</table>

5.2. Parabolic equations with measure data in space. Since our theoretical results are also valid for Neumann boundary conditions, we consider in this subsection the following parabolic equation with Dirac source term in space and Neumann...
Table 5. Error of $y$ for Example 5.3 with respect to space and time.

<table>
<thead>
<tr>
<th>Dofs</th>
<th>$|y - Y_h|_{L^2(\Omega_T)}$</th>
<th>rate</th>
<th>$N$</th>
<th>$|y - Y_h|_{L^2(\Omega_T)}$</th>
<th>rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>31</td>
<td>0.033467303580</td>
<td>_</td>
<td>2</td>
<td>0.069700966171</td>
<td>_</td>
</tr>
<tr>
<td>105</td>
<td>0.007254281003</td>
<td>2.50</td>
<td>4</td>
<td>0.030875705044</td>
<td>1.17</td>
</tr>
<tr>
<td>385</td>
<td>0.002842650160</td>
<td>1.44</td>
<td>8</td>
<td>0.014389164719</td>
<td>1.10</td>
</tr>
<tr>
<td>1473</td>
<td>0.001350586630</td>
<td>1.17</td>
<td>16</td>
<td>0.006927239146</td>
<td>1.05</td>
</tr>
<tr>
<td>5761</td>
<td>0.000551335197</td>
<td>1.10</td>
<td>32</td>
<td>0.003398471622</td>
<td>1.02</td>
</tr>
<tr>
<td>22785</td>
<td>0.000194881287</td>
<td>1.05</td>
<td>64</td>
<td>0.001688698129</td>
<td>1.00</td>
</tr>
</tbody>
</table>

boundary condition:

\[
\begin{align*}
    y_t - \Delta y + y &= \mu \quad \text{in } \Omega_T, \\
    \frac{\partial y}{\partial n} &= 0 \quad \text{on } \Gamma_T, \\
    y(0) &= y_0 \quad \text{in } \Omega,
\end{align*}
\]

where \( \mu = g(x,t)\omega(x) \), \( g(x,t) \in L^2(0,T;C(\overline{\Omega})) \) and \( \omega(x) \in \mathcal{M}(\overline{\Omega}) \). For ease of constructing examples we also admit some additional regular parts to appear in the right-hand side.

Example 5.3. The third example is a modification from Example 4.2 of [9]. Let \( \Omega_T = [0,1]^2 \times [0,1] \), we take the exact solution as

\[
y(x,t) = (e^t + 1) \cdot \left\{ \begin{array}{ll}
    0.5 - x_1^2, & x_1 < 0.5, \\
    0.25, & x_1 \geq 0.5,
\end{array} \right.
\]

since \( y \) does not depend on the spatial variable \( x_2 \), we find that \( \Delta y = \Delta x_1 y = (\delta_{x_1}(\frac{1}{2}) - \psi(x_1)) \cdot (e^t + 1) \), where

\[
\psi(x_1) = \begin{cases} 
2, & x_1 < 0.5, \\
0, & x_1 \geq 0.5,
\end{cases}
\]

and \( \delta_{x_i}(z) \) denotes the Dirac measure with respect to the variable \( x_i \) concentrated at \( x_i = z \). Then after simple calculation we have

\[
\mu(x,t) = y(x,t) - (e^t + 1) \cdot (\delta_{x_1}(\frac{1}{2}) - \psi(x_1)) + e^t \cdot \left\{ \begin{array}{ll}
    0.5 - x_1^2, & x_1 < 0.5, \\
    0.25, & x_1 \geq 0.5.
\end{array} \right.
\]

To investigate the convergence order with respect to the space discretization we fix the time discretization with \( N = 2048 \), while the spatial discretization is fixed with 22785 Dofs to investigate the convergence order with respect to the time discretization, the results are listed in Table 5 and Figure 1 presents the numerical result at time \( t = 0.5 \) for a grid with 22785 Dofs. We can see from Table 5 that the convergence order w.r.t the space discretization is almost 1, which is consistent with our theoretical results. While the convergence order w.r.t the time discretization is 1, which is better than our predicted result of order \( k^{\frac{1}{2}} \).

Example 5.4. The fourth example is a modification from Example 4.1 of [9]; see also [1]. Let \( \Omega_T = B(0,1) \times [0,1] \), we take the exact solution as

\[
y(x,t) = -\frac{1}{2\pi} \log |x| \cdot (e^t + 1),
\]
then after simple calculation we have

\[ \mu(x, t) = (e^t + 1)\delta_0 + y(x, t) - \frac{1}{2\pi} \log |x| \cdot e^t, \]

where \( \delta_0 \) is the Dirac function at \( x = (0, 0) \).

To investigate the convergence order with respect to the space discretization we fixed the time discretization with \( N = 2048 \), while the space discretization is fixed with 66049 Dofs to investigate the convergence order with respect to the time discretization, the results are listed in Table 6 and Figure 2 presents the numerical result at time \( t = 0.5 \) for a grid with 66049 Dofs. We can see from Table 6 that the convergence order w.r.t the space discretization is 1, which is consistent with our theoretical results. While the convergence order w.r.t the time discretization is also 1, which is better than our predicted result of order \( k^{\frac{1}{2}} \).
Table 6. Error of y for Example 5.4 with respect to space and time.

<table>
<thead>
<tr>
<th>Dofs</th>
<th>(|y - Y_h|_{L^2(\Omega_T)})</th>
<th>rate</th>
<th>(N)</th>
<th>(|y - Y_h|_{L^2(\Omega_T)})</th>
<th>rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td>0.051725167740</td>
<td>\emptyset</td>
<td>2</td>
<td>0.029719956266</td>
<td>\emptyset</td>
</tr>
<tr>
<td>81</td>
<td>0.026461768120</td>
<td>0.9670</td>
<td>4</td>
<td>0.013208571995</td>
<td>1.1700</td>
</tr>
<tr>
<td>289</td>
<td>0.013361104038</td>
<td>0.9859</td>
<td>8</td>
<td>0.006210402376</td>
<td>1.0887</td>
</tr>
<tr>
<td>1089</td>
<td>0.006716533360</td>
<td>0.9923</td>
<td>16</td>
<td>0.003085031620</td>
<td>1.0094</td>
</tr>
<tr>
<td>4225</td>
<td>0.003366984764</td>
<td>0.9963</td>
<td>32</td>
<td>0.001685293247</td>
<td>0.8723</td>
</tr>
<tr>
<td>16641</td>
<td>0.001685479572</td>
<td>0.9983</td>
<td>64</td>
<td>0.001112777391</td>
<td>0.5988</td>
</tr>
</tbody>
</table>

6. Conclusion and future work

In this paper we study the a priori error estimates for the finite element approximations of parabolic equations with measure data separately in time and space, respectively. The space discretization is done using piecewise linear and continuous finite elements, whereas the time discretization is based on the backward Euler method. We derive the a priori error estimates for the semidiscretization problems and the fully discrete problems, respectively. Numerical results are provided at the end of the paper to confirm our theoretical findings.

To the best of author’s knowledge, this paper is among the few contributions to finite element method for partial differential equations involving measure data. The results obtained in this paper especially constitute the crucial ingredients to derive the error estimates for some kind of parabolic optimal control problems with state constraints. The traditional approach of error analysis for parabolic optimal control problems with state constraints is to avoid the error estimates for adjoint state equation, which is caused by the lack of error estimates for parabolic equations involving measure data (see, e.g., [10], [15] and [22]). We believe that the results in this paper provide a shortcut for the error analysis of such kinds of problems.

Moreover, the results obtained in this paper can be viewed as the first step, but a crucial step, for the error analysis of the finite element approximation to parabolic optimal control problems with pointwise control, where control acts on finitely many points of the domain. The state equation has the form of (2.13) with right-hand side \(\mu = u(t)\omega(x)\) (see (1.6)), where \(u(t)\) denotes the control variable, the details can be found in [11], [14] and [25]. The finite element approximation of such kinds of problems and corresponding error analysis will be addressed in our future work. In addition, as pointed out in the introduction, parabolic equations with measure data find many applications in optimal control theory. But generally the Lagrange multiplier \(\mu\) associated to the state constraints for parabolic optimal control problems with pointwise state constraints belongs to \(\mathcal{M}(\Omega_T)\) (see (1.5)), so it is also very interesting to study the finite element approximation of parabolic problems with measure data in both space and time (see [4] for the analytical setting), the approaches developed in the current paper seems to be inapplicable for such extreme cases.

On the other hand, since the solutions of parabolic problems involving measure data have lower regularities, only reduced convergence order can be expected by standard finite element approximation. Thus the a posteriori error estimate and adaptive finite element method for such kinds of problems are necessary and deserve further study. Araya et al. ([1]) have studied a posteriori error estimates for elliptic
problems with Dirac delta source terms, the applications of their approaches to our setting will be postponed to our future work.

ACKNOWLEDGEMENTS

The author is grateful to the anonymous referees for their careful reviews and many valuable suggestions that led to a considerably improved paper, he also acknowledges professor Ningning Yan from AMSS, CAS for a careful reading of the preliminary version of this article. The author would like to thank the Alexander von Humboldt Foundation for support during his stay at the University of Hamburg, Germany. He is very grateful to the Department of Mathematics, University of Hamburg for the hospitality and support.

REFERENCES


Schwerpunkt Optimierung und Approximation, Universität Hamburg, Bundesstrasse 55, 20146, Hamburg, Germany

E-mail address: wgong@lsec.cc.ac.cn

Current address: LSEC, Institute of Computational Mathematics and Scientific/Engineering Computing, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China