Analysis of a quasi-reversibility method for a terminal value quasi-linear parabolic problem with measurements

Peer-reviewed author version

Huy Tuan Nguyen; ANH-KHOA, Vo & Van Au Vo (2019) Analysis of a quasi-reversibility method for a terminal value quasi-linear parabolic problem with measurements. In: Siam Journal on Mathematical Analysis, 51 (1), p. 60-85.

DOI: 10.1137/18M1174064
Handle: http://hdl.handle.net/1942/30249

ANALYSIS OF A QUASI-REVERSIBILITY METHOD FOR A TERMINAL VALUE QUASI-LINEAR PARABOLIC PROBLEM WITH MEASUREMENTS*

NGUYEN HUY TUAN†, VO ANH KHOA‡, AND VO VAN AU§

Abstract. This paper presents a modified quasi-reversibility method for computing the exponentially unstable solution of a nonlocal terminal-boundary value parabolic problem with noisy data. Based on data measurements, we perturb the problem by the so-called filter regularized operator to design an approximate problem. Different from recently developed approaches that consist in the conventional spectral methods, we analyze this new approximation in a variational framework, where the finite element method can be applied. To see the whole skeleton of this method, our main results lie in the analysis of a semi-linear case and we discuss some generalizations where this analysis can be adapted. As is omnipresent in many physical processes, there is likely a myriad of models derived from this simpler case, such as source localization problems for brain tumors and heat conduction problems with nonlinear sinks in nuclear science. With respect to each noise level, we benefit from the Faedo-Galerkin method to study the weak solvability of the approximate problem. Relying on the energy-like analysis, we provide detailed convergence rates in L^2-H^1 of the proposed method when the true solution is sufficiently smooth. Depending on the dimensions of the domain, we obtain an error estimate in L^r for some $r > 2$. Proof of the backward uniqueness for the quasi-linear system is also depicted in this work. To prove the regularity assumptions acceptable, several physical applications are discussed.

Key words. Quasi-linear parabolic problems, Ill-posed problems, Uniqueness, Faedo-Galerkin method, Quasi-reversibility method, Convergence rates.

AMS subject classifications. 65J05, 65J20, 35K92

1. Introduction.

1.1. Background of the terminal value model.

This paper is concerned with a general construction of a modified quasi-reversibility method for a quasi-linear parabolic reaction-diffusion system of the following form

$$u_t + \nabla \cdot (-a(x,t;u;\nabla u)\nabla u) = F(x,t;u;\nabla u) \quad \text{in } Q_T := \Omega \times (0,T), \tag{1}$$

where the vector of concentrations $u = u(x,t) \in \mathbb{R}^N$ is unknown with $N \in \mathbb{N}^*$ being the number of equations involved in (1). Here, the domain of interest $\Omega \subset \mathbb{R}^d$ for $d \in \mathbb{N}^*$ and the final time of observation $0 < T < \infty$ are assumed. Furthermore, Ω is open, connected and bounded with a sufficiently smooth boundary $\partial\Omega$. The nonlocal diffusion coefficient $a \in \mathbb{R}^{N\times N}$ and the nonlinearity $F \in \mathbb{R}^N$ are explicitly density- and gradient-dependent.

As met in practical applications, we associate (1) either with the homogeneous Dirichlet boundary condition ($u = 0$ on $\partial\Omega$) or with the homogeneous Neumann

---

*Submitted to the editors DATE.

Funding: The work of the first author and the third author was supported by the Institute for Computational Science and Technology Ho Chi Minh City under the project named "Some ill-posed problems for partial differential equations" (312/QD-KHCNTT). The work of the second author was supported by a postdoctoral fellowship of the Research Foundation-Flanders (FWO).

†Institute for Computational Science and Technology Ho Chi Minh City, Vietnam, and Faculty of Mathematics and Computer Science, University of Science, Vietnam National University, 227 Nguyen Van Cu, District 5, Ho Chi Minh City, Vietnam (nhtuan@hcmus.edu.vn).

‡Author for correspondence. Institute for Numerical and Applied Mathematics, University of Goettingen, 37083 Goettingen, Germany, and Faculty of Sciences, Hasselt University, Campus Diepenbeek, Agoralaan Building D, BE3590 Diepenbeek, Belgium (vakhoa.hcmus@gmail.com).

§Faculty of General Sciences, Can Tho University of Technology, Can Tho City, Vietnam (vvau@ctuet.edu.vn).

boundary condition ($-a(x,t;u;\nabla u)\nabla u \cdot \mathrm{n} = 0$ on $\partial\Omega$). Given the terminal data

$$u(x,T) = u_f(x) \quad \text{in } \Omega, \tag{2}$$

we would like to seek in this work the initial value $u(x,0) = u_0(x)$ in a stable way since the solution to this type of problems is highly unstable (cf. e.g. [25]).

The motivation behind the consideration of (1)-(2) basically follows the identification of source location for brain tumor that has been investigated in [24]. It is worth mentioning that reconstructing the initial densities of tumor cells provides a substantial contribution to predicting tumorigenicity in connection with genetic events (see this possible correlation studied in e.g. [56]). The spirit of studying the terminal data (2) also arises in the theory of Kolmogorov backward equations, carried out by nonlocal transformations in [3] e.g., to integrate the expected value of the payout from future values. Therefore, the problem under consideration here is viewed as a prototypical framework which can be adapted to particular applicable contexts and be extended to other theoretical approaches.

The existing literature on quasi-linear reaction-diffusion systems is very huge to be singled out here. Since the diffusion tensor a in (1) is nonlinear with included self- and/or cross diffusion types, there are of course numerous distinctive aspects concerning different types of forward problems considered here. For example, discussions on well-posedness, spectrum analysis and behaviors of travelling waves have been detailed in [41] and references cited therein, see e.g. [50, 40, 15]. We also wish to mention here the works [6, 55, 23] for addressing more complex scenarios related to either theoretical or numerical standpoints of reaction-diffusion type systems.

1.2. Goals and novelty. The purpose here is to follow up on our earlier work [52], where we have proposed a regularization strategy in the vein of the classical *quasi-reversibility* (QR) method which specifically solves ill-posed problems of elliptic and parabolic types. Observe that the identification of population density for a single-species model in [52] is well-suited to the concept of source localization. In this sense, behaviors of the tumor cell density are influenced not only by certain proliferation and/or extinction rates, but also by their transport processes with convection/advection, and the total population in local movements. The novelty we present here is the careful adaptation of the *filter regularized operator*, which we have briefly studied in [17], to the modified QR method in a variational framework. Remarkably, this setting enables us to interact with certain reaction-diffusion problems with spatial nonlinear diffusion; compared to our spectral-based regularization methods (cf. e.g. [52, 54, 45]) that have been studied so far.

The motivation for using the QR type method stems from our wish to design a regularization approach that can deal with a quite general class of parabolic problems due to the limitation of regularization theory. As is known, regularization of many simpler models has been deduced so far, such as the heat conduction problem (e.g. [53, 47] and references therein), the parabolic problems in image restoration (cf. [11, 8, 12, 10]), the Burger equation in fluid mechanics [9] and even the Navier-Stokes equation [34]. However, it is impossible to find papers working on time-reversed quasi-linear systems (1)-(2) except our previous work [52] that has been mentioned above. Our major contribution here is thus coping with problems that remain unsolved until now. We stress that (1) includes not only popular semi-linear types (e.g. equations/systems named Fitzhugh-Nagumo, Fisher-KPP, Zeldovich, Lengyel-Epstein, de Pillis-Radunskaya and Frank-Kamenetsky; see [39, 19, 18, 14, 2, 46] for the background of deterministic models), but also certain nonlocal types in e.g. [49, 37, 13, 26]. In principle, our mathematical results derived here are helpful in fostering interests in the branch of inverse and ill-posed problems for partial differential equations. Alternative approaches to design a regularized problem can be the quasi-boundary value method (commenced in [51] and numerically discussed in [28] e.g.), the truncation method (see, e.g. [45, 48]) and the recently developed Tikhonov method based on Carleman weight functions (cf. [32]). In addition, backward problems with impulsive and random noise have been investigated in [35] by the generalized Tikhonov regularization and in [29] by a QR-based statistical approach, respectively.

We accentuate that this work is not aimed to improve the conventional convergence rates of this method, but to complete the theoretical error analysis of this direction. Together with rigorous L^2-H^1 error estimates, we obtain an L^r-type rate ($r > 2$) of convergence, which we believe that this is the first time it is explored in this direction. Theoretically, our work also unravels the problem of finding the global-in-time error estimate. To be more specific, we recall the analysis in [31], where a linear case of (1) was considered through a version of the QR method. Proofs of the stability and error estimates in [31] are concretely based on the massive Carleman estimate, but it is well known that this method often requires T sufficiently small; see [31, Theorem 5.1]. This price is also manifested here when we prove the backward uniqueness result for (1) using a Carleman-type estimate with a suitable non-increasing weight function; see Lemma 5.1. Here, the rates of convergence we obtain for the semi-linear case are similar to [31, Theorem 5.4], but are uniform in time, requiring a very high smoothness of the true solution somewhat in terms of Gevrey spaces. To prove this, an exponentially decreasing weight function is used to get rid of large parameters appearing in the difference problem. Essentially, this largeness is driven by the magnitude stability of the regularized problem, which goes to infinity when the measurement parameter tends to zero; see e.g. [44] for various types of the magnitude in the past. In accordance with the existence result of the regularized problem, this way the proofs of convergence would be simpler using a large amount of energy-like estimates.

1.3. Contemporary history of the QR method. The QR method was first proposed by Lattès and Lions in the monograph [38]. This method, when applied to the context of linear backward parabolic problems, basically perturbs the spatial second-order operator by the addition of a fourth-order term. It is, on the other hand, going with a leading parameter which is positive and small enough to get the convergence. Additionally, the sign of this extra term is chosen such that the perturbed/regularized problem is well-posed with respect to the leading parameter, as time evolves back to the initial point. In the community of regularization, this parameter is referred to as the *regularization parameter*. Let us also note that since our work aims to prepare the playground to handle real-world models, the presence of noise on the terminal data is evident. Accordingly, the smallness of the regularization parameter here depends strictly on such noise levels, which makes our scheme applicable in reality.

Having massive research interests for five decades, the literature of the QR method and its modifications (e.g. the stabilized QR method in [42]), nowadays, is vast from the vantage point of theoretical and numerical analysis. As some of concrete references for elliptic equations, a dual-based QR method for the Cauchy problem with noisy data is designed in [5] and some numerical approaches have been postulated in [7]. On the other hand, the error analysis is very attractive and has been investigated, for example, in [33] with the Hölder-type rate and in [4] with a logarithmic rate. This method is also extended to deal with inverse problems for parabolic and hyperbolic equations; see e.g. [16, 30] for a brief overview of this field with sharp error estimates and convergence results in H^1. On top of that, the reader can be referred to [31] as a survey of applications of Carleman estimates to proofs of convergence of the QR method for a wide class of ill-posed problems for PDEs.

1.4. Outline of the paper. From a mathematical point of view, the nonlinearities a and F involved in (1) are undoubtedly the major challenges. Here we aim at showing the general setting of the QR method and thereupon explaining the ideas on a simpler case while leaving the more general case of (1) to future works in this inception stage. For this reason, we introduce in (10) the regularized problem for the general system (1), while we reduce ourselves to the analysis of a semi-linear case with single-species mode. In this regard, we will not also detail any further technical assumption of the diffusion tensor a, except the ellipticity condition in the general form of (1) that serves the convergence analysis. Notice that proofs of our main results are done with the zero Dirichlet boundary condition, which plays a key role in the variational framework we choose. As shortly discussed in the last part of subsection 5.1 these results are also obtainable for the zero Neumann boundary condition.

Except the notation and necessary assumptions on the input of the problem in the general form are present in section 2, our main themes in this paper can be summarized as follows:

- Detailed settings of the modified QR method are studied in section 3..
- Weak solvability of the regularized problem is investigated in subsection 4.1; see Theorem 4.3 and Theorem 4.4. Detailed rates of convergence are obtained in subsection 4.2, where the main results are reported in Theorem 4.5, Theorem 4.6 and Corollary 4.7, respectively.
- Some particular extensions follow in subsection 5.1, including the uniqueness result for the system (1)-(2).

Finally, some working applications are present in Appendix A.

2. Preliminaries. Let $\langle \cdot, \cdot \rangle$ be either the scalar product in L^2 or the dual pairing of a continuous linear functional and an element of a function space. The notation $\|\cdot\|_X$ stands for the norm in the Banach space X. We call X' the dual space of X. We denote by $L^p(0,T;X)$, $1 \le p \le \infty$ for the Banach space of real-valued functions $u : (0,T) \to X$ measurable, provided that

$$\|u\|_{L^p(0,T;X)} = \left(\int_0^T \|u(t)\|_X^p \, dt \right)^{1/p} < \infty \quad \text{for } 1 \le p < \infty,$$

while

$$\|u\|_{L^\infty(0,T;X)} = \operatorname*{ess\,sup}_{t \in (0,T)} \|u(t)\|_X < \infty \quad \text{for } p = \infty.$$

We denote the norm of the function space $C^k([0,T];X)$, $0 \le k \le \infty$ by

$$\|u\|_{C^k([0,T];X)} = \sum_{n=0}^{k} \sup_{t \in [0,T]} \left\| u^{(n)}(t) \right\|_X < \infty.$$

We denote by $H_0^1(\Omega)$ for the Hilbert space of weakly differentiable functions $u : \Omega \to \mathbb{R}$ that vanishes on the boundary in the sense of trace. On the other hand, $W^{p,q}(\Omega)$ for $p \in \mathbb{N}$ denotes the Sobolev space of functions with index of differentiability p and of integrability q (if $q \in \mathbb{N}$) or, in the case $q = \infty$, whose essential supremum exists.

Depending on the situation, we denote by $|\cdot|$ either the absolute value of a function or the finite-dimensional Euclidean norm of a vector. There are several assumptions needed for the analysis below:

(A_1) The diffusion tensor $a = (a_{ij})_{1\le i,j\le N}$ is such that the mapping $(\mathbf{p},\mathbf{q}) \mapsto a(x,t;\mathbf{p};\mathbf{q})$ is continuous for $(\mathbf{p},\mathbf{q}) \in [L^2(\Omega)]^N \times \left[L^2(\Omega)\right]^{Nd}$ and the mapping $(x,t) \mapsto a(x,t;\mathbf{p};\mathbf{q})$ is continuously differentiable for $(x,t) \in \overline{Q_T}$. Moreover, there exists a positive constant $\overline{M}$ such that

$$0 < \sum_{i,j=1}^{N} a_{ij}(x,t;\mathbf{p};\mathbf{q})\,\xi_i\xi_j \le \overline{M}\,|\xi|^2 \text{ for all } \xi \in \mathbb{R}^N, (\mathbf{p},\mathbf{q}) \in [L^2(\Omega)]^N \times \left[L^2(\Omega)\right]^{Nd}.$$

(A_2) There exists a tensor $A(x,t;\mathbf{p},\mathbf{q}) \in \mathbb{R}^{N\times N}$ such that $A_{ij} = \overline{M} - a_{ij}$ for $1 \le i,j \le N$. Then there exists a positive constant $\underline{M}$ satisfying

$$0 < \underline{M}\,|\xi|^2 \le \sum_{i,j=1}^{N} A_{ij}(x,t;\mathbf{p};\mathbf{q})\,\xi_i\xi_j \le \overline{M}\,|\xi|^2,$$

for all $\xi \in \mathbb{R}^N, (\mathbf{p},\mathbf{q}) \in [L^2(\Omega)]^N \times \left[L^2(\Omega)\right]^{Nd}$.

(A_3) For any $(x,t) \in \overline{Q_T}$, the source function F is measurable and locally Lipschitz-continuous in the sense that for $1 \le i \le N$

$$|F(x,t;\mathbf{p};\mathbf{q}) - F(x,t;\mathbf{r};\mathbf{s})| \le L_F(\ell)\left(|\mathbf{p}-\mathbf{r}| + |\mathbf{q}-\mathbf{s}|\right),$$

for $\max\{|\mathbf{p}|, |\mathbf{r}|, |\mathbf{q}|, |\mathbf{s}|\} \le \ell$ for some $\ell > 0$.

(A_4) There exists a measurement of u_f, denoted by u_f^ε, in $[L^2(\Omega)]^N$ such that

$$\left\| u_f - u_f^\varepsilon \right\|_{[L^2(\Omega)]^N} \le \varepsilon,$$

where $\varepsilon > 0$ represents the noise level.

REMARK 2.1. *It follows from (A_3) that we can take*

$$L_F(\ell) := \sup\left\{ \frac{|F(x,t;\mathbf{p};\mathbf{q}) - F(x,t;\mathbf{r};\mathbf{s})|}{|\mathbf{p}-\mathbf{r}| + |\mathbf{q}-\mathbf{s}|} \right.$$
$$\left. : (x,t) \in Q_T, \mathbf{p} \ne \mathbf{r}, \mathbf{q} \ne \mathbf{s} \text{ and } |\mathbf{p}|, |\mathbf{r}|, |\mathbf{q}|, |\mathbf{s}| \le \ell \right\} < \infty,$$

for $\ell > 0$. Moreover, we introduce the cut-off function F_ℓ, as follows:

$$(3) \quad F_\ell(x,t;\mathbf{p};\mathbf{q}) := \begin{cases} F(x,t;\ell;\ell) & \text{if } \max_{1\le j\le N, 1\le k\le d}\{\mathbf{p}_j, \mathbf{q}_{jk}\} \in (\ell,\infty), \\ F(x,t;\mathbf{p};\mathbf{q}) & \text{if } \max_{1\le j\le N, 1\le k\le d}\{\mathbf{p}_j, \mathbf{q}_{jk}\} \in [-\ell,\ell], \\ F(x,t;-\ell;-\ell) & \text{if } \max_{1\le j\le N, 1\le k\le d}\{\mathbf{p}_j, \mathbf{q}_{jk}\} \in (-\infty,-\ell). \end{cases}$$

Due to the cut-off function, for any $\ell > 0$ it holds

$$(4) \qquad |F_\ell(x,t;\mathbf{p};\mathbf{q}) - F_\ell(x,t;\mathbf{r};\mathbf{s})| \le L_F(\ell)\left(|\mathbf{p}-\mathbf{r}| + |\mathbf{q}-\mathbf{s}|\right),$$

for all $(x,t) \in Q_T$ and $\mathbf{p},\mathbf{r} \in [L^2(\Omega)]^N, \mathbf{q},\mathbf{s} \in [L^2(\Omega)]^{Nd}$.

The proof of (4) is trivial. For brevity, we sketch out the proof in the following cases and omit the details:

Case 1: $\max_{1\le j\le N,1\le k\le d}\{\mathbf{p}_j, \mathbf{q}_{jk}\} < -\ell$ *and* $\max_{1\le j\le N,1\le k\le d}\{\mathbf{r}_j, \mathbf{s}_{jk}\} < -\ell$;

Case 2: $\max_{1\le j\le N,1\le k\le d}\{\mathbf{p}_j, \mathbf{q}_{jk}\} < -\ell \le \max_{1\le j\le N,1\le k\le d}\{\mathbf{r}_j, \mathbf{s}_{jk}\} \le \ell$;

Case 3: $\max_{1\le j\le N,1\le k\le d}\{\mathbf{p}_j, \mathbf{q}_{jk}\} < -\ell < \ell \le \max_{1\le j\le N,1\le k\le d}\{\mathbf{r}_j, \mathbf{s}_{jk}\}$;

Case 4: $-\ell < \max_{1\le j\le N,1\le k\le d}\{\mathbf{p}_j, \mathbf{q}_{jk}\}, \max_{1\le j\le N,1\le k\le d}\{\mathbf{r}_j, \mathbf{s}_{jk}\} \le \ell$;

Case 5: $\max_{1\le j\le N,1\le k\le d}\{\mathbf{p}_j, \mathbf{q}_{jk}\} > \ell$ *and* $\max_{1\le j\le N,1\le k\le d}\{\mathbf{r}_j, \mathbf{s}_{jk}\} > \ell$.

3. General frameworks for the QR method. This is the moment to establish a regularized problem for the system (1)-(2) with measured data u_f^ε. For $\varepsilon > 0$, we denote by $\beta := \beta(\varepsilon) \in (0,1)$ the regularization parameter satisfying

$$\lim_{\varepsilon\to 0^+} \beta(\varepsilon) = 0,$$

and then consider the function $\gamma : [0,T] \times (0,1)$ such that for any $\beta > 0$, there holds

$$\gamma(T,\beta) \geq 1, \quad \lim_{\beta\to 0^+} \gamma(t,\beta) = \infty \quad \text{for all } t \in (0,T].$$

Compared to [17], we do not require the fundamental multiplicative-like identities with respect to the first argument in γ. With the function γ at hands, we define the following operators.

Definition 3.1 (Perturbing operator). *The linear mapping $\mathbf{Q}_\varepsilon^\beta : [L^2(\Omega)]^N \to [L^2(\Omega)]^N$ is said to be a perturbing operator if there exist a function space $\mathbb{W} \subset [L^2(\Omega)]^N$ and an ε-independent constant $C_0 > 0$ such that*

$$\left\|\mathbf{Q}_\varepsilon^\beta u\right\|_{[L^2(\Omega)]^N} \leq C_0 \|u\|_{\mathbb{W}} / \gamma(T,\beta) \quad \text{for any } u \in \mathbb{W}. \tag{5}$$

Definition 3.2 (Stabilized operator). *The linear mapping $\mathbf{P}_\varepsilon^\beta : [L^2(\Omega)]^N \to [L^2(\Omega)]^N$ is said to be a stabilized operator if there exists an ε-independent constant $C_1 > 0$ such that*

$$\left\|\mathbf{P}_\varepsilon^\beta u\right\|_{[L^2(\Omega)]^N} \leq C_1 \log(\gamma(T,\beta)) \|u\|_{[L^2(\Omega)]^N} \quad \text{for any } u \in [L^2(\Omega)]^N. \tag{6}$$

In principle, the way we define these two terms $\mathbf{P}_\varepsilon^\beta$ and $\mathbf{Q}_\varepsilon^\beta$ is in line with the classical quasi-reversibility method. In this sense, we obtain the regularized problem by adding the perturbing operator $\mathbf{Q}_\varepsilon^\beta$ to the original problem. Then the stabilized operator will be derived from this addition by a linear mapping, whenever the leading coefficients of operators, which are targeted to be stabilized, are essentially bounded. Hence, in this work we simply take $\mathbf{P}_\varepsilon^\beta := \overline{M}\Delta + \mathbf{Q}_\varepsilon^\beta$. Interestingly, this enables us to consider a very simple eigenvalue problem regardless of the complex structure involved in the diffusion coefficient.

At the moment, we do not know the optimal bounds of (5) and (6), which are altogether related. We deliberately present the logarithmic stability estimate (6) for the stabilized operator due to the typical logarithmic convergence usually obtained after the regularization of a backward parabolic model. In other words, this upper bound is essential and decisive in the convergence analysis in subsection 4.2. The decay behavior of the perturbing operator (cf. (5)) is directly governed by the so-called source condition that measures the high smoothness of the true solution. In the following example, we mimic the stochastic gradient descent algorithm in machine learning schemes to show the existence of these operators.

Example 3.3. *Consider $N = 1$ for a single-species model. It is well known that for any bounded subset of $\mathbb{R}^d$ with a smooth boundary, there exists an orthonormal basis of $L^2(\Omega)$, denoted by $\{\phi_p\}_{p\in\mathbb{N}}$, satisfying $\phi_p \in H_0^1(\Omega) \cap C^\infty(\overline{\Omega})$ and $-\Delta\phi_p(x) = \mu_p\phi_p(x)$ for $x \in \Omega$. The (Dirichlet and Neumann) eigenvalues $\{\mu_p\}_{p\in\mathbb{N}}$ form an infinite sequence which goes to infinity, viz.*

$$0 \leq \mu_0 < \mu_1 \leq \mu_2 \leq \ldots, \text{ and } \lim_{p\to\infty} \mu_p = \infty. \tag{7}$$

We choose

$$\mathbf{Q}_{\varepsilon}^{\beta}u=\frac{1}{T}\sum_{p\in\mathbb{N}}\log\left(1+\gamma^{-1}\left(T,\beta\right)e^{\overline{M}T\mu_{p}}\right)\left\langle u,\phi_{p}\right\rangle\phi_{p}\quad\text{for } u\in L^{2}\left(\Omega\right). \tag{8}$$

Using the elementary inequality $\log\left(1+a\right)\le a$ *for* $a>0$*, then by Parseval's identity it gives*

$$\begin{aligned}\left\|\mathbf{Q}_{\varepsilon}^{\beta}u\right\|_{L^{2}(\Omega)}^{2}&=\frac{1}{T^{2}}\sum_{p\in\mathbb{N}}\log^{2}\left(1+\gamma^{-1}\left(T,\beta\right)e^{\overline{M}T\mu_{p}}\right)\left|\left\langle u,\phi_{p}\right\rangle\right|^{2}\\&\le\frac{\gamma^{-2}\left(T,\beta\right)}{T^{2}}\left\|e^{\overline{M}T(-\Delta)}u\right\|_{L^{2}(\Omega)}^{2}.\end{aligned}$$

The norm $\left\|e^{\overline{M}T(-\Delta)}u\right\|_{L^{2}(\Omega)}$ *is characterized by the so-called Gevrey class of real-analytic functions. In this case, it is also performed as a Hilbert space and then contained in* $L^{2}(\Omega)$*. Fruitful discussions on this typical space are preferably in section 5 and Appendix A. It now remains to deduce the estimate for the operator* $\mathbf{P}_{\varepsilon}^{\beta}$*. In fact, it follows from its own structure that*

$$\mathbf{P}_{\varepsilon}^{\beta}u=\sum_{p\in\mathbb{N}}\left(\frac{1}{T}\log\left(1+\gamma^{-1}\left(T,\beta\right)e^{\overline{M}T\mu_{p}}\right)-\overline{M}\mu_{p}\right)\left\langle u,\phi_{p}\right\rangle\phi_{p}.$$

Thanks to the inequality $\log\left(1+ab\right)\le\log\left(b\left(1+a\right)\right)\le\log\left(1+a\right)+\log\left(b\right)$ *for* $a>0,b\ge1$*, we have*

$$\log\left(1+\gamma^{-1}\left(T,\beta\right)e^{\overline{M}T\mu_{p}}\right)-\overline{M}T\mu_{p}\le\log\left(1+\gamma^{-1}\left(T,\beta\right)\right)\quad\text{for all } p\in\mathbb{N}.$$

Consequently, by Parseval's identity we get

$$\left\|\mathbf{P}_{\varepsilon}^{\beta}u\right\|_{L^{2}(\Omega)}\le\frac{1}{T}\log\left(\gamma\left(T,\beta\right)\right)\left\|u\right\|_{L^{2}(\Omega)}.$$

Now, we detail the regularized problem: For each $\varepsilon>0$, let $\ell^{\varepsilon}:=\ell\left(\varepsilon\right)\in\left(0,\infty\right)$ be a cut-off parameter satisfying

$$\lim_{\varepsilon\to0^{+}}\ell^{\varepsilon}=\infty, \tag{9}$$

then we consider the following problem:

$$u_{t}^{\varepsilon}+\nabla\cdot\left(-a\left(x,t;u^{\varepsilon};\nabla u^{\varepsilon}\right)\nabla u^{\varepsilon}\right)-\mathbf{Q}_{\varepsilon}^{\beta}u^{\varepsilon}=F_{\ell^{\varepsilon}}\left(x,t;u^{\varepsilon};\nabla u^{\varepsilon}\right)\quad\text{in } Q_{T}, \tag{10}$$

associated with the Dirichlet boundary condition and the terminal noisy data

$$u^{\varepsilon}=0\ \text{ on }\partial\Omega\times\left(0,T\right),\quad u^{\varepsilon}\left(x,T\right)=u_{f}^{\varepsilon}\left(x\right)\ \text{ in }\Omega. \tag{11}$$

4. Analysis of the QR method. Some certain cases of the general system (1) can be solved by the QR scheme we have proposed. Nevertheless, this mathematical over-generality merely leads to extra steps of proofs. Thereby, this curtails the core idea behind the regularization. In this section we only consider a rather simplified version of (1), while we will briefly discuss the result of the general system in subsection 5.1. We take into account the following reaction-diffusion equation with $N = 1$:

$$u_t - a\Delta u = F(x,t;u) \quad \text{in } Q_T, \tag{12}$$

endowed with the zero Dirichlet boundary condition and the terminal condition (2).

This means we will use the assumptions (A_1)-(A_4) with $N = 1$. Notice that (12) also implies the following reduction through our analysis:

- $a(x,t;u;\nabla u) = a > 0$ – the method only needs to use the strict upper bound of the diffusion coefficient, saying that $a < \overline{M}$ (reduced from (A_1)). Corresponding to (A_2), this way we take $A = \overline{M} - a \in (\overline{M} - M_1, \overline{M})$ for $a < M_1 < \overline{M}$ by the completeness of real numbers.
- $F(x,t;u;\nabla u) = F(x,t;u)$ is globally Lipschitz-continuous in u, i.e. $F_\ell = F$ and $L_F(\ell) = L_F$ is independent of all involved parameters; see (A_3) and (4) with a typical example $F(u) = \sin u$. We will come back to the locally Lipschitz-continuous case of F in subsection 5.1. This case is significantly more difficult to estimate due to the blow-up profile of the cut-off parameter ℓ^ε; see (9).

Hereby, when we recall these assumptions, i.e. (A_1)-(A_4), in the analysis, it is understood that the correspondingly reduced versions are considered. We below scrutinize the existence result for the regularized problem and the convergence analysis obtained after applying the QR scheme (10)-(11) to the semi-linear case (12). When doing so, proofs of our results are based on several energy-like estimates using an auxiliary parameter, denoted either by ρ_ε or by ρ_β, depending on whether the regularization parameter β is involved. In this spirit, we technically seek fine energy controls for the "scaled" problems obtained by the weight function $e^{\rho_\varepsilon(t-T)}$. The choice of this parameter is definitely dependent on every single aspect of analysis, but it will at least include the magnitude of stability of the regularized problem. Thus, its behavior obeys

$$\lim_{\varepsilon\to 0^+} \rho_\varepsilon = \infty.$$

4.1. Existence result for the regularized problem. For each $\varepsilon > 0$, we put $v^\varepsilon(x,t) = e^{\rho_\varepsilon(t-T)} u^\varepsilon(x,t)$. Under a suitable choice of such a parameter, we obtain the existence result for the regularized problem in the framework of Faedo-Galerkin procedures. Using (A_3), the regularized problem (10) for the semi-linear case (12) can be rewritten as

$$u_t^\varepsilon + A\Delta u^\varepsilon = F(x,t;u^\varepsilon) + \mathbf{P}_\varepsilon^\beta u^\varepsilon. \tag{13}$$

Multiplying this equation by $e^{\rho_\varepsilon(t-T)}$, it becomes

$$v_t^\varepsilon + A\Delta v^\varepsilon - \rho_\varepsilon v^\varepsilon = e^{\rho_\varepsilon(t-T)} F(x,t;u^\varepsilon) + \mathbf{P}_\varepsilon^\beta v^\varepsilon, \tag{14}$$

by virtue of the linearity of the operator $\mathbf{P}_\varepsilon^\beta$.

Note that the boundary and terminal conditions of (14) remain the same as (11) due to the structural definition of v^ε. Henceforward, multiplying (14) by a test function $\psi \in H_0^1(\Omega)$ we define a weak formulation of (11) in the following standard type.

Definition 4.1. *For each $\varepsilon > 0$, a function v^ε is said to be a weak solution of (14) if*

$$v^\varepsilon \in L^2\left(0,T;H_0^1(\Omega)\right) \cap L^\infty\left(0,T;L^2(\Omega)\right)$$

and it holds

$$(15)\quad \frac{d}{dt}\langle v^{\varepsilon},\psi\rangle - A\int_{\Omega}\nabla v^{\varepsilon}\cdot\nabla\psi dx - \rho_{\varepsilon}\langle v^{\varepsilon},\psi\rangle$$
$$= e^{\rho_{\varepsilon}(t-T)}\left\langle F\left(\cdot,t;e^{\rho_{\varepsilon}(T-t)}v^{\varepsilon}\right),\psi\right\rangle + \left\langle \mathbf{P}_{\varepsilon}^{\beta}v^{\varepsilon},\psi\right\rangle \text{ for all } \psi\in H_0^1(\Omega).$$

Let $\mathbb{S}_n$ be the space generated by $\phi_1,\phi_2,...,\phi_n$ for $n=1,2,...$ where in general $\{\phi_j\}$ is a Schauder basis of $H^1(\Omega)$ (so it can be the eigenfunctions mentioned in Example 3.3), then let

$$(16)\qquad v_n^{\varepsilon}(x,t) = \sum_{j=1}^{n} V_{jn}^{\varepsilon}(t)\,\phi_j(x)$$

be the weak solution of the following approximate problem, corresponding to (14):

$$(17)\qquad \left\langle (v_n^{\varepsilon})_t,\psi\right\rangle - A\int_{\Omega}\nabla v_n^{\varepsilon}\cdot\nabla\psi dx - \rho_{\varepsilon}\langle v_n^{\varepsilon},\psi\rangle$$
$$= e^{\rho_{\varepsilon}(t-T)}\left\langle F\left(\cdot,t;e^{\rho_{\varepsilon}(T-t)}v_n^{\varepsilon}\right),\psi\right\rangle + \left\langle \mathbf{P}_{\varepsilon}^{\beta}v_n^{\varepsilon},\psi\right\rangle,$$

for all $\psi\in\mathbb{S}_n$, with the final condition

$$(18)\qquad v_n^{\varepsilon}(T) = v_{fn}^{\varepsilon} = \sum_{j=1}^{n}\left(V_f^{\varepsilon}\right)_{jn}\phi_j \to u_f^{\varepsilon} \text{ strongly in } L^2(\Omega) \text{ as } n\to\infty.$$

To derive the nonlinear ordinary differential equations with respect to the time argument for $V_{jn}(t)$, it follows from (17) with using $\psi=\phi_j$ that for $1\le j\le n$,

$$\left(V_{jn}^{\varepsilon}\right)_t - (A+\rho_{\varepsilon})V_{jn}^{\varepsilon} = e^{\rho_{\varepsilon}(t-T)}\left\langle F\left(\cdot,t;e^{\rho_{\varepsilon}(T-t)}v_n^{\varepsilon}\right),\phi_j\right\rangle + \left\langle \mathbf{P}_{\varepsilon}^{\beta}v_n^{\varepsilon},\phi_j\right\rangle,$$

and $V_{jn}^{\varepsilon}(T) = \left(V_f^{\varepsilon}\right)_{jn}$.

By using the Newton-Liebniz formula, one has

$$(19)\qquad V_{jn}^{\varepsilon}(t) = \left(V_f^{\varepsilon}\right)_{jn} - (A+\rho_{\varepsilon})\int_t^T V_{jn}^{\varepsilon}(s)\,ds$$
$$-\int_t^T\left[e^{\rho_{\varepsilon}(s-T)}\left\langle F\left(\cdot,s;e^{\rho_{\varepsilon}(T-s)}v_n^{\varepsilon}\right),\phi_j\right\rangle + \left\langle \mathbf{P}_{\varepsilon}^{\beta}v_n^{\varepsilon}(s),\phi_j\right\rangle\right]ds.$$

LEMMA 4.2. *Suppose that (6) holds. For any fixed $n\in\mathbb{N}$ and for each $\varepsilon>0$, the system (17)-(18) has a unique solution $V_{jn}^{\varepsilon}\in C([0,T])$.*

Proof. The proof of this lemma is standard. Here we sketch out some important steps because it seems pertinent to see more detailed impact of ρ_{ε} on all the analysis. We define the norm in the Banach space $Y=C([0,T];\mathbb{R}^n)$ as follows:

$$\|c\|_Y := \sup_{t\in[0,T]}\sum_{j=1}^{n}|c_j(t)| \quad \text{with } c=(c_j)_{1\le j\le n}.$$

By virtue of (19), we can define a Volterra-type integral equation and then set the operator $\mathcal{G}: C([0,T];\mathbb{R}^n)\to C([0,T];\mathbb{R}^n)$ by

$$\mathcal{G}(V^{\varepsilon})(t) = H^{\varepsilon} - \int_t^T K^{\varepsilon}(s,V^{\varepsilon})\,ds,$$

where in the vector form, V^ε and H^ε indicate V_j^ε and $\left(V_f^\varepsilon\right)_j$, respectively, and K^ε stands for the right-hand side of (19) under the integration in time.

Observe that when summing (19) with respect to j up to n, we have

$$(20)\quad \sum_{j=1}^{n} V_j^\varepsilon(t) = \sum_{j=1}^{n}\left(V_f^\varepsilon\right)_j - (A+\rho_\varepsilon)\sum_{j=1}^{n}\int_t^T V_j^\varepsilon(s)\,ds - \sum_{j=1}^{n}\int_t^T \left[e^{\rho_\varepsilon(s-T)}\left\langle F\left(\cdot,s;e^{\rho_\varepsilon(T-s)}v_n^\varepsilon\right),\phi_j\right\rangle + \left\langle \mathbf{P}_\varepsilon^\beta v_n^\varepsilon(s),\phi_j\right\rangle\right]ds.$$

For $V^\varepsilon \in C([0,T];\mathbb{R}^n)$ and $W^\varepsilon \in C([0,T];\mathbb{R}^n)$ we have the following estimates. With the aid of the Lipschitz assumption (A$_3$) we easily get

$$(21)\quad \left|\left\langle F\left(\cdot,s;e^{\rho_\varepsilon(T-s)}v_n^\varepsilon\right) - F\left(\cdot,s;e^{\rho_\varepsilon(T-s)}w_n^\varepsilon\right),\phi_j\right\rangle\right| \le CL_F e^{\rho_\varepsilon(T-s)}\sum_{k=1}^{n}\left|V_k^\varepsilon - W_k^\varepsilon\right|,$$

and in the same vein, using (6) implies that

$$(22)\quad \left|\left\langle \mathbf{P}_\varepsilon^\beta v_n^\varepsilon(s),\phi_j\right\rangle - \left\langle \mathbf{P}_\varepsilon^\beta w_n^\varepsilon(s),\phi_j\right\rangle\right| \le CC_1\log\left(\gamma(T,\beta)\right)\sum_{k=1}^{n}\left|V_k^\varepsilon - W_k^\varepsilon\right|.$$

Grouping (21) and (22), it follows from (1) that the following estimate can be obtained:

$$\begin{aligned}&\left|\mathcal{G}(V^\varepsilon) - \mathcal{G}(W^\varepsilon)\right| \\ &\le C(T-t)\left(\overline{M}+\rho_\varepsilon+nL_F+C_1\log\left(\gamma(T,\beta)\right)\right)\left\|V^\varepsilon - W^\varepsilon\right\|_Y,\end{aligned}$$

and furthermore, by induction we deduce

$$\begin{aligned}&\left|\mathcal{G}^m(V^\varepsilon) - \mathcal{G}^m(W^\varepsilon)\right| \\ &\le \frac{(T-t)^m}{m!}C^m\left(\overline{M}+\rho_\varepsilon+L_F+C_1\log\left(\gamma(T,\beta)\right)\right)^m\left\|V^\varepsilon - W^\varepsilon\right\|_Y,\end{aligned}$$

where we denote by $\mathcal{G}^m(V^\varepsilon) = \mathcal{G}(\mathcal{G}...\mathcal{G}(V^\varepsilon))$.

Since for each $\varepsilon>0$ and $n\in\mathbb{N}$, there exists $m_0\in\mathbb{N}$ such that

$$\frac{(T-t)^{m_0}}{m_0!}C^{m_0}\left(\overline{M}+\rho_\varepsilon+L_F+C_1\log\left(\gamma(T,\beta)\right)\right)^{m_0} < 1,$$

then $\mathcal{G}^{m_0}$ is a contraction mapping from $C([0,T];\mathbb{R}^n)$ onto itself. By the Banach fixed-point argument, there exists a unique solution V^ε in Y such that $\mathcal{G}^{m_0}(V^\varepsilon) = V^\varepsilon$. Combining this with the fact that $\mathcal{G}^{m_0}(\mathcal{G}(V^\varepsilon)) = \mathcal{G}(\mathcal{G}^{m_0}(V^\varepsilon)) = \mathcal{G}(V^\varepsilon)$, the integral equation $\mathcal{G}(V^\varepsilon) = V^\varepsilon$ admits a unique solution in $C([0,T];\mathbb{R}^n)$. ∎

From here on, we state the existence result in the following theorem.

THEOREM 4.3. *For each $\varepsilon>0$, the regularized problem (14) has a weak solution v^ε in the sense of Definition 4.1. Moreover, it satisfies $v^\varepsilon \in C\left([0,T];L^2(\Omega)\right)$ and $v_t^\varepsilon \in L^2\left(0,T;\left(H^1(\Omega)\right)'\right)$.*

Proof. We now give some *a priori* estimates for the solution of the problem (14). When doing so, we choose $\psi = v_n^\varepsilon$ in (17) to get

$$\frac{1}{2}\frac{d}{dt}\left\|v_n^\varepsilon\right\|_{L^2(\Omega)}^2 - A\left\|\nabla v_n^\varepsilon\right\|_{[L^2(\Omega)]^d}^2 - \rho_\varepsilon\left\|v_n^\varepsilon\right\|_{L^2(\Omega)}^2 = \underbrace{e^{\rho_\varepsilon(t-T)}\left\langle F\left(\cdot,t;e^{\rho_\varepsilon(T-t)}v_n^\varepsilon\right),v_n^\varepsilon\right\rangle}_{:=I_3} + \underbrace{\left\langle \mathbf{P}_\varepsilon^\beta v_n^\varepsilon, v_n^\varepsilon\right\rangle}_{:=I_4}. \tag{23}$$

Note from the resulting structural condition of F in (A_3) that

$$e^{\rho_\varepsilon(t-T)}\left|F\left(x,t;e^{\rho_\varepsilon(T-t)}v_n^\varepsilon\right) - F\left(x,t;0\right)\right| \le L_F\left|v_n^\varepsilon\right|,$$

one thus has

$$\begin{aligned} I_3 &\ge -\frac{e^{2\rho_\varepsilon(t-T)}}{2L_F}\left\|F\left(\cdot,t;e^{\rho_\varepsilon(T-t)}v_n^\varepsilon\right)\right\|_{L^2(\Omega)}^2 - \frac{L_F}{2}\left\|v_n^\varepsilon\right\|_{L^2(\Omega)}^2 \\ &\ge -\frac{e^{2\rho_\varepsilon(t-T)}}{2L_F}\left\|F\left(\cdot,t;0\right)\right\|_{L^2(\Omega)}^2 - \frac{1}{2}\left(1+L_F\right)\left\|v_n^\varepsilon\right\|_{L^2(\Omega)}^2. \end{aligned}$$

Similarly, based on the structural definition of $\mathbf{P}_\varepsilon^\beta$ in (6), it yields

$$\begin{aligned} \left\langle \mathbf{P}_\varepsilon^\beta v_n^\varepsilon, v_n^\varepsilon\right\rangle &\ge -\frac{1}{2}\left(\left\|\mathbf{P}_\varepsilon^\beta v_n^\varepsilon\right\|_{L^2(\Omega)}^2 + \left\|v_n^\varepsilon\right\|_{L^2(\Omega)}^2\right) \\ &\ge -\frac{1}{2}\left(C_1^2\log^2\left(\gamma\left(T,\beta\right)\right)+1\right)\left\|v_n^\varepsilon\right\|_{L^2(\Omega)}^2. \end{aligned}$$

Then, (23) can be estimated by

$$\begin{aligned} &\frac{d}{dt}\left\|v_n^\varepsilon\right\|_{L^2(\Omega)}^2 + \frac{e^{2\rho_\varepsilon(t-T)}}{L_F}\left\|F\left(\cdot,t;0\right)\right\|_{L^2(\Omega)}^2 \\ &\ge 2\underline{M}\left\|\nabla v_n^\varepsilon\right\|_{[L^2(\Omega)]^d}^2 + \left(2\rho_\varepsilon - \left(1+L_F\right) - C_1^2\log^2\left(\gamma\left(T,\beta\right)\right) - 1\right)\left\|v_n^\varepsilon\right\|_{L^2(\Omega)}^2, \end{aligned}$$

where we have used the assumption (A_3).

Hereby, for each $\varepsilon > 0$ we choose $2\rho_\varepsilon = L_F + C_1^2\log^2\left(\gamma\left(T,\beta\right)\right) + 2 > 0$, then integrate the resulting estimate from t to T to obtain

$$\begin{aligned} &\left\|v_n^\varepsilon\left(T\right)\right\|_{L^2(\Omega)}^2 + \frac{e^{-2T\rho_\varepsilon}}{L_F}\int_t^T\left\|F\left(\cdot,s;0\right)\right\|_{L^2(\Omega)}^2 ds \\ &\qquad \ge \left\|v_n^\varepsilon\left(t\right)\right\|_{L^2(\Omega)}^2 + \underline{M}\int_t^T\left\|\nabla v_n^\varepsilon(s)\right\|_{[L^2(\Omega)]^d}^2 ds. \end{aligned}$$

Since $v_n^\varepsilon\left(T\right) \to u_f^\varepsilon$ in $L^2\left(\Omega\right)$ (cf. (18)), we can find an ε-independent constant $\bar{c} > 0$ such that

$$\left\|v_n^\varepsilon\left(t\right)\right\|_{L^2(\Omega)}^2 + \underline{M}\int_t^T\left\|\nabla v_n^\varepsilon\right\|_{[L^2(\Omega)]^d}^2 ds \le \bar{c}.$$

As byproduct, we have

$$v_n^\varepsilon \text{ is bounded in } L^\infty\left(0,T;L^2\left(\Omega\right)\right) \text{ and in } L^2\left(0,T;H_0^1\left(\Omega\right)\right). \tag{24}$$

Observe that

$$\left(v_n^\varepsilon\right)_t + A\Delta v_n^\varepsilon - \rho_\varepsilon v_n^\varepsilon = e^{\rho_\varepsilon(t-T)}F\left(x,t;e^{\rho_\varepsilon(T-t)}v_n^\varepsilon\right) + \mathbf{P}_\varepsilon^\beta v_n^\varepsilon \in \left(H^1\left(\Omega\right)\right)',$$

which provides

$$(v_n^\varepsilon)_t \text{ is bounded in } L^2\left(0,T;\left(H^1(\Omega)\right)'\right). \tag{25}$$

Thanks to the Banach-Alaoglu theorem, the uniform bounds with respect to n, as obtained in (24)-(25), imply that one can extract a subsequence (which we relabel with the index n if necessary) such that for each $\varepsilon > 0$,

$$v_n^\varepsilon \to v^\varepsilon \text{ weakly} - * \text{ in } L^\infty\left(0,T;L^2(\Omega)\right), \tag{26}$$

$$v_n^\varepsilon \to v^\varepsilon \text{ weakly in } L^2\left(0,T;H_0^1(\Omega)\right), \tag{27}$$

$$(v_n^\varepsilon)_t \to v_t^\varepsilon \text{ weakly in } L^2\left(0,T;\left(H^1(\Omega)\right)'\right). \tag{28}$$

Furthermore, by the Aubin-Lions compactness theorem in combination with the Gelfand triple $H_0^1(\Omega) \subset L^2(\Omega) \subset \left(H^1(\Omega)\right)'$, one gets from (26) and (28) that

$$v_n^\varepsilon \to v^\varepsilon \text{ strongly in } L^2(Q_T) \text{ and so a.e. in } Q_T \text{ for a further subsequence.} \tag{29}$$

Note also that due to (6), one has for each $\varepsilon > 0$,

$$\mathbf{P}_\varepsilon^\beta v_n^\varepsilon \to \mathbf{P}_\varepsilon^\beta v^\varepsilon \text{ strongly in } L^2(Q_T) \text{ and so a.e. in } Q_T \text{ for a further subsequence.} \tag{30}$$

In the same manner, one has for each $\varepsilon > 0$,

$$F\left(e^{\rho_\varepsilon(T-t)}v_n^\varepsilon\right) \to F\left(e^{\rho_\varepsilon(T-t)}v^\varepsilon\right) \text{ strongly in } L^2(Q_T). \tag{31}$$

From here on, by grouping (26)-(28) and (29)-(31) we can pass to the limit in (17) to show that v^ε satisfies the problem (14) in the weak sense (15). On top of that, due to (24) and (25), we have

$$v^\varepsilon \in C\left([0,T];L^2(\Omega)\right),$$

where we have applied the embeddings $H_0^1(\Omega) \subset L^2(\Omega) \subset \left(H^1(\Omega)\right)'$ and $H^1(0,T) \subset C[0,T]$.

Now, it remains to verify the terminal data. In fact, we take a function $\vartheta \in C^1[0,T]$ with $\vartheta(0) = 0$ and $\vartheta(T) = 1$. As a consequence of the convergence (25), one has

$$\int_0^T \langle (v_n^\varepsilon)_t, \psi\rangle \vartheta dt \to \int_0^T \langle v_t^\varepsilon, \psi\rangle \vartheta dt \quad \text{for all } \psi \in L^2(\Omega),$$

and by integration by parts together with the Newton-Liebniz formula, it becomes

$$-\int_0^T \langle v_n^\varepsilon, \psi\rangle \vartheta_t dt + \langle v_n^\varepsilon(T), \psi\rangle \vartheta(T) \to -\int_0^T \langle v^\varepsilon, \psi\rangle \vartheta_t dt + \langle v^\varepsilon(T), \psi\rangle \vartheta(T), \tag{32}$$

for all $\psi \in L^2(\Omega)$. Consequently, the weak convergence (27) allows us to obtain $\langle v_n^\varepsilon(T), \psi\rangle \to \langle v^\varepsilon(T), \psi\rangle$ for all $\psi \in H_0^1(\Omega)$ from (32). Combining this convergence with the fact already known that $v_n^\varepsilon(T)$ converges strongly to u_f^ε in $L^2(\Omega)$; see (18). We thus get $\langle v_n^\varepsilon(T), \psi\rangle \to \left\langle u_f^\varepsilon, \psi\right\rangle$ for all $\psi \in H_0^1(\Omega)$. Due to the uniqueness of the limit, it reveals that $\langle v^\varepsilon(T), \psi\rangle = \left\langle u_f^\varepsilon, \psi\right\rangle$ for all $\psi \in H_0^1(\Omega)$ and thus $v^\varepsilon(T) = u_f^\varepsilon$ a.e. in Ω. ∎

Now we show the positivity and boundedness of solution to the regularized problem (14). In the following theorem, if the measured inputs of the concentrations are positive and essentially bounded in a spatial environment, their distributions that obey the proposed approximation remain the same properties therein by a suitable choice of the auxiliary parameter ρ_ε. In other words, the behavior of the regularized solution strictly depends on the way ρ_ε being taken.

THEOREM 4.4. *Let v^ε be a weak solution of the problem (14) as deduced in Theorem 4.3. For each $\varepsilon > 0$, suppose that $0 \le u_f^\varepsilon \in L^\infty(\Omega)$ and $F(x,t;0) \equiv 0$ for a.e. $(x,t) \in Q_T$. Moreover, for all real-valued constant $C > 0$ we assume $\mathbf{P}_\varepsilon^\beta C = \mathbf{Q}_\varepsilon^\beta C \ge 0$. Then, $0 \le v^\varepsilon \le \left\| u_f^\varepsilon \right\|_{L^\infty(\Omega)}$ for a.e. $(x,t) \in Q_T$.*

Proof. Let $v^\varepsilon := v^{\varepsilon,+} - v^{\varepsilon,-}$ where $f^+ := \max\{f,0\}$ and $f^- := \max\{-f,0\}$. In (15), we now take the test function $\psi = -v^{\varepsilon,-}$. Then, by ($\mathrm{A}_3$), ($\mathrm{A}_4$) and (6) we have

$$
\begin{aligned}
\frac{d}{dt}\left\|v^{\varepsilon,-}\right\|_{L^2(\Omega)}^2 &\ge \underline{M}\left\|\nabla v^{\varepsilon,-}\right\|_{[L^2(\Omega)]^d}^2 \\
&\quad + \left(\rho_\varepsilon - L_F - C_1\log\left(\gamma(T,\beta)\right) - 1\right)\left\|v^{\varepsilon,-}\right\|_{L^2(\Omega)}^2,
\end{aligned}
\tag{33}
$$

inspired very much the way we have estimated (23).

Choosing $\rho_\varepsilon = L_F + C_1\log\left(\gamma(T,\beta)\right) + 1 > 0$ and observing that $v^{\varepsilon,-}|_{t=T} = 0$, we integrate (33) from t to T to get $\|v^{\varepsilon,-}\|_{L^2(\Omega)}^2 \le 0$, which indicates the positivity of v^ε.

To prove the upper bound, we take the test function $\psi = (v^\varepsilon - B)^+$ in (15) where $B \ge \left\|u_f^\varepsilon\right\|_{L^\infty(\Omega)}$. Thus, we arrive at

$$
\begin{aligned}
&\frac{d}{dt}\left\|(v^\varepsilon - B)^+\right\|_{L^2(\Omega)}^2 \\
&\ge \underline{M}\left\|\nabla(v^\varepsilon - B)^+\right\|_{[L^2(\Omega)]^d}^2 + \rho_\varepsilon\left\|(v^\varepsilon - B)^+\right\|_{L^2(\Omega)}^2 + \rho_\varepsilon\left\langle B,(v^\varepsilon - B)^+\right\rangle \\
&+ \left\langle \mathbf{P}_\varepsilon^\beta (v^\varepsilon - B)^+, (v^\varepsilon - B)^+\right\rangle + \left\langle \mathbf{P}_\varepsilon^\beta B, (v^\varepsilon - B)^+\right\rangle \\
&+ \underbrace{e^{\rho_\varepsilon(t-T)}\left\langle F\left(e^{\rho_\varepsilon(T-t)}v^\varepsilon\right),(v^\varepsilon - B)^+\right\rangle}_{:=I_5}.
\end{aligned}
\tag{34}
$$

Here, taking into account the structural condition of F we get

$$
\begin{aligned}
I_5 &\ge -L_F\left|\left\langle |v^\varepsilon|, (v^\varepsilon - B)^+\right\rangle\right| \\
&\ge -L_F\left(\left\|(v^\varepsilon - B)^+\right\|_{L^2(\Omega)}^2 + \left\langle B,(v^\varepsilon - B)^+\right\rangle\right).
\end{aligned}
$$

At this stage, we proceed as in the proof of the positivity. By choosing $\rho_\varepsilon = C_1\log\left(\gamma(T,\beta)\right) + L_F > 0$, it follows from (34) that $(v^\varepsilon - B)^+ = 0$, provided that $\left.(v^\varepsilon - B)^+\right|_{t=T} = 0$. Hence, we complete the proof of the theorem. ∎

4.2. Convergence analysis. We are now going to derive the convergence rates obtained when the regularized solution u_β^ε of (10)-(11) is applied to approximate the solution u of (12)-(2) in the presence of noise on the final data. Note that in the previous subsection we only write u^ε as the regularized solution since the parameter ε is already fixed. Instead, we denote in this part u_β^ε due to the choice of the regularization parameter $\beta(\varepsilon)$ which plays a vital role in this analysis.

Although Example 3.3 shows that $C_1 = \frac{1}{T}$, for an arbitrary $C_1 > 0$ we need $C_1 T \le 1$ in our main results to gain strong convergences. At some points, this is in the same spirit of the terminology *small solution* defined in [9].

4.2.1. Statement of the results. Here we state our main results as Theorem 4.5 and Theorem 4.6; their solid proofs are deferred to subsection 4.2.2 and subsection 4.2.3, respectively. Moreover, proof of Corollary 4.7 is given in subsection 4.2.4.

In the following, let $\gamma(t,\beta)$ for $t \in [0,T]$ and $\beta := \beta(\varepsilon)$ be as in section 3. We choose

$$\lim_{\varepsilon \to 0^+} \gamma^{C_1 T}(T,\beta)\,\varepsilon = K \in (0,\infty). \tag{35}$$

THEOREM 4.5. **(Error estimates for $0 < t \le T$)**
Assume that the problem (12)-(2) *admits a unique solution*

$$u \in C([0,T];\mathbb{W}), \tag{36}$$

where the precise structure of $\mathbb{W}$ depends on the choice of the operator $\mathbf{Q}_\varepsilon^\beta$ in (5). For a suitable choice of the operator $\mathbf{P}_\varepsilon^\beta$ in (6), we consider $u_\beta^\varepsilon \in C\left([0,T];L^2(\Omega)\right)$ as a solution of (13)-(11) *corresponding to the measured data u_f^ε. Then the following error estimate holds*

$$\begin{aligned}
&\left\|u_\beta^\varepsilon(\cdot,t) - u(\cdot,t)\right\|_{L^2(\Omega)} + \sqrt{2\underline{M}}\int_t^T \left\|\nabla u_\beta^\varepsilon(\cdot,s) - \nabla u(\cdot,s)\right\|_{[L^2(\Omega)]^d} ds \\
&\le \gamma^{-C_1 t}(T,\beta)\left(K + \sqrt{2T}C_0\gamma^{C_1T-1}(T,\beta)\|u\|_{C([0,T];\mathbb{W})}\right)e^{TC_2},
\end{aligned}$$

for $t \in (0,T)$ and $C_i > 0$ ($i \in \{0,1,2\}$) independent of ε.

THEOREM 4.6. **(Error estimate for $t = 0$)**
Under the assumptions of Theorem 4.5, we assume further that

$$u \in C([0,T];\mathbb{W}) \cap C^1\left(0,T;L^2(\Omega)\right). \tag{37}$$

Then, for $\varepsilon > 0$ small enough we can find a unique $t^\varepsilon \in (0,T)$ such that

$$\begin{aligned}
\left\|u_\beta^\varepsilon(\cdot,t^\varepsilon) - u(\cdot,0)\right\|_{L^2(\Omega)} \le \Big[&\left(K + \sqrt{2T}C_0\gamma^{C_1T-1}(T,\beta)\|u\|_{C([0,T];\mathbb{W})}\right)e^{TC_2} \\
&+ \|u_t\|_{C(0,T;L^2(\Omega))}\Big]\frac{1}{\sqrt{C_1}\log^{\frac{1}{2}}(\gamma(T,\beta))},
\end{aligned}$$

where $C_i > 0$ ($i \in \{0,1,2\}$) are independent of ε.

COROLLARY 4.7. *Under the assumptions of Theorem 4.5, one has for any $0 < t < T$, u_β^ε is strongly convergent to u in $L^2(t,T;L^r(\Omega))$ for some $r > 2$ with the same rate as in Theorem 4.5.*

4.2.2. Proof of Theorem 4.5. For an auxiliary parameter $\rho_\beta > 0$, we put $w_\beta^\varepsilon(x,t) := e^{\rho_\beta(t-T)}\left[u_\beta^\varepsilon(x,t) - u(x,t)\right]$. Then, we compute that

$$\begin{aligned}
&\frac{\partial w_\beta^\varepsilon}{\partial t} + A\Delta w_\beta^\varepsilon - \rho_\beta w_\beta^\varepsilon \\
&= \mathbf{P}_\varepsilon^\beta w_\beta^\varepsilon + e^{\rho_\beta(t-T)}\mathbf{Q}_\varepsilon^\beta u + e^{\rho_\beta(t-T)}\left[F\left(x,t;u_\beta^\varepsilon\right) - F(x,t;u)\right].
\end{aligned} \tag{38}$$

This equation is associated with the zero Dirichlet boundary condition $w_\beta^\varepsilon = 0$ on $\partial\Omega \times (0,T)$ and the following terminal condition:

$$w_\beta^\varepsilon (x,T) = u_{\beta f}^\varepsilon (x) - u_f (x) \quad \text{for } x \in \Omega.$$

Multiplying (38) by w_β^ε and then integrating the resulting equation over Ω, we arrive at

$$\begin{aligned}(39)\qquad &\frac{1}{2}\frac{d}{dt}\left\|w_\beta^\varepsilon\right\|_{L^2(\Omega)}^2 - A\left\|\nabla w_\beta^\varepsilon\right\|_{[L^2(\Omega)]^d}^2 - \rho_\beta \left\|w_\beta^\varepsilon\right\|_{L^2(\Omega)}^2 \\ &= \underbrace{\left\langle \mathbf{P}_\varepsilon^\beta w_\beta^\varepsilon, w_\beta^\varepsilon\right\rangle}_{:=\mathcal{I}_1} + \underbrace{e^{\rho_\beta (t-T)}\left\langle \mathbf{Q}_\varepsilon^\beta u, w_\beta^\varepsilon\right\rangle}_{:=\mathcal{I}_2} + \underbrace{e^{\rho_\beta (t-T)}\left\langle F\left(u_\beta^\varepsilon\right) - F(u), w_\beta^\varepsilon\right\rangle}_{:=\mathcal{I}_3}.\end{aligned}$$

To investigate the convergence analysis, we need to bound from below the right-hand side of (39). Relying on the structural property of the operator $\mathbf{P}_\varepsilon^\beta$ (cf. (6)), $\mathcal{I}_1$ can be estimated by

$$(40)\qquad \mathcal{I}_1 \ge -C_1 \log\left(\gamma(T,\beta)\right)\left\|w_\beta^\varepsilon\right\|_{L^2(\Omega)}^2,$$

with the aid of Hölder's inequality.

Using the Young inequality and the structural property of the operator $\mathbf{Q}_\varepsilon^\beta$ (cf. (5)), $\mathcal{I}_2$ can be estimated by

$$(41)\qquad \mathcal{I}_2 \ge -C_0^2\gamma^{-2}(T,\beta)\left\|u\right\|_{\mathbb{W}}^2 - \frac{1}{4}\left\|w_\beta^\varepsilon\right\|_{L^2(\Omega)}^2.$$

From now on, taking also into account the Lipschitz constant L_F and choosing an appropriate Young inequality, we get the estimate of $\mathcal{I}_3$ as follows:

$$\begin{aligned}(42)\qquad \mathcal{I}_3 &\ge -\frac{e^{2\rho_\beta (t-T)}}{8L_F^2}\left\|F\left(u_\beta^\varepsilon\right) - F(u)\right\|_{L^2(\Omega)}^2 - 2L_F^2\left\|w_\beta^\varepsilon\right\|_{[L^2(\Omega)]^N}^2 \\ &\ge -\left(\frac{1}{4} + 2L_F^2\right)\left\|w_\beta^\varepsilon\right\|_{L^2(\Omega)}^2.\end{aligned}$$

Plugging (40), (41) and (42) into (39), and then integrating the resulting estimate from t to T we obtain, after some rearrangement, that

$$\begin{aligned}(43)\qquad &\left\|w_\beta^\varepsilon(T)\right\|_{L^2(\Omega)}^2 + 2(T-t)C_0^2\gamma^{-2}(T,\beta)\left\|u\right\|_{\mathbb{W}}^2 \\ &\qquad\ge \left\|w_\beta^\varepsilon(t)\right\|_{L^2(\Omega)}^2 + 2\underline{M}\int_t^T \left\|\nabla w_\beta^\varepsilon(s)\right\|_{[L^2(\Omega)]^d}^2 ds,\end{aligned}$$

by putting $\rho_\beta = C_1 \log\left(\gamma(T,\beta)\right) + \frac{1}{2} + 2L_F^2 > 0$.

Note here that the existence of $u_\beta^\varepsilon \in L^2\left(0,T;H_0^1(\Omega)\right)$ has already been obtained in subsection 4.1. Due to (A_4) the first norm on the left-hand side of (43) is bounded from above by ε^2. By the back-substitution $w_\beta^\varepsilon(x,t) := e^{\rho_\beta (t-T)}\left[u_\beta^\varepsilon(x,t) - u(x,t)\right]$ and the choice of ρ_β, we thus conclude that

$$\begin{aligned}(44)\qquad &\left\|u_\beta^\varepsilon(\cdot,t) - u(\cdot,t)\right\|_{L^2(\Omega)}^2 + 2\underline{M}\int_t^T \left\|\nabla u_\beta^\varepsilon(\cdot,s) - \nabla u(\cdot,s)\right\|_{[L^2(\Omega)]^d}^2 ds \\ &\le \gamma^{2C_1(T-t)}(T,\beta)\left(\varepsilon^2 + 2(T-t)C_0^2\gamma^{-2}(T,\beta)\left\|u\right\|_{C([0,T];\mathbb{W})}^2\right)e^{2(T-t)C_2},\end{aligned}$$

where we have denoted by

$$C_2 := \frac{1}{2} + 2L_F^2. \tag{45}$$

Together with the ε-dependent blow-up rate of γ in (35), this ends the proof of the theorem.

4.2.3. Proof of Theorem 4.6. It is clear that in Theorem 4.5 the convergence does not hold at $t = 0$. Taking a number $t^\varepsilon \in (0,T)$, we prove that for each $\varepsilon > 0$, there exists $t^\varepsilon > 0$ such that $u_\beta^\varepsilon (x, t = t^\varepsilon)$ is a good approximation candidate of $u(x, t = 0)$. Indeed, if the source condition (37) holds true, we get

$$\begin{aligned}
&\left\| u_\beta^\varepsilon (\cdot, t^\varepsilon) - u(\cdot, 0) \right\|_{L^2(\Omega)} \\
&\le \left\| u_\beta^\varepsilon (\cdot, t^\varepsilon) - u(\cdot, t^\varepsilon) \right\|_{L^2(\Omega)} + \left\| u(\cdot, t^\varepsilon) - u(\cdot, 0) \right\|_{L^2(\Omega)} \\
&\le \gamma^{-C_1 t^\varepsilon} (T, \beta) \left(K + \sqrt{2T} C_0 \gamma^{C_1 T - 1} (T, \beta) \|u\|_{C([0,T];\mathbb{W})} \right) e^{TC_2} \\
&+ t^\varepsilon \|u_t\|_{C(0,T;L^2(\Omega))} .
\end{aligned}$$

Observe that the error bound $\left\| u_\beta^\varepsilon (\cdot, t^\varepsilon) - u(\cdot, 0) \right\|_{L^2(\Omega)}$ is essentially decided by the infimum of $\frac{1}{2}\left(\gamma^{-C_1 t^\varepsilon} (T,\beta) + t^\varepsilon \right)$ with respect to $t^\varepsilon > 0$. We find that the term $\gamma^{-C_1 t^\varepsilon}(T,\beta)$ is decreasing and t^ε obviously possesses a linear growth. Therefore, for every $\beta := \beta(\varepsilon) > 0$ there exists a unique $t^\varepsilon \in (0,T)$ such that

$$\begin{cases} \lim\limits_{\varepsilon \to 0^+} t^\varepsilon = 0, \\ t^\varepsilon = \gamma^{-C_1 t^\varepsilon} (T, \beta), \end{cases} \tag{46}$$

and the second equation can be rewritten as

$$\frac{\log(t^\varepsilon)}{t^\varepsilon} = -C_1 \log(\gamma(T,\beta)). \tag{47}$$

Using the elementary inequality $\log(a) > -a^{-1}$ for all $a > 0$, it follows from (47) that

$$t^\varepsilon < \sqrt{\frac{1}{C_1 \log(\gamma(T,\beta))}}.$$

Henceforward, for t^ε sufficiently small we complete the proof of the theorem.

4.2.4. Proof of Corollary 4.7. In this part, we rely on the Gagliardo-Nirenberg interpolation inequality for functions having zero trace to derive the error estimate. Essentially, it reads as

$$\begin{aligned}
&\int_t^T \left\| u_\beta^\varepsilon (\cdot, s) - u(\cdot, s) \right\|_{L^r(\Omega)}^2 ds \\
&\le C_\Omega^2 \left\| u_\beta^\varepsilon - u \right\|_{C([t,T];L^2(\Omega))}^{2\alpha} \int_t^T \left\| \nabla \left(u_\beta^\varepsilon - u \right)(\cdot, s) \right\|_{[L^2(\Omega)]^d}^{2(1-\alpha)} ds,
\end{aligned} \tag{48}$$

where $C_\Omega > 0$ is a generic constant that only depends on the geometry of Ω, and the involved parameters should hold with: $r > 2$ and $0 < \alpha < 1$ satisfying

$$\frac{1}{r} > \frac{d-2}{2d} \text{ and } \frac{1}{r} = \frac{\alpha}{2} + \frac{(1-\alpha)(d-2)}{2d}.$$

Note that (48) is available because of the existence of $u_\beta^\varepsilon \in L^2\left(0,T;H_0^1\left(\Omega\right)\right) \cap L^\infty\left(0,T;L^2\left(\Omega\right)\right)$ leading to $C\left(\left[0,T\right];L^2\left(\Omega\right)\right)$; see subsection 4.1, and the compact embedding $H^1\left(\Omega\right) \subset L^r\left(\Omega\right)$. The special case of (48) in two and three-dimensional versions ($d = 2,3$) is the well known Ladyzhenskaya inequality.

Using Hölder's inequality we can write (48) as

$$\begin{aligned}(49)\quad &\int_t^T \left\|u_\beta^\varepsilon\left(\cdot,s\right) - u\left(\cdot,s\right)\right\|_{L^r(\Omega)}^2 ds \\ &\le C_\Omega^2 \left(T-t\right)^\alpha \left\|u_\beta^\varepsilon - u\right\|_{C([t,T];L^2(\Omega))}^{2\alpha} \left(\int_t^T \left\|\nabla\left(u_\beta^\varepsilon - u\right)\left(\cdot,s\right)\right\|_{[L^2(\Omega)]^d}^2 ds\right)^{1-\alpha}.\end{aligned}$$

We remark that in (49) we are only able to get the convergence until the near zero point of time, i.e. it merely holds for $0 < t < T$. Accordingly, it is straightforward to obtain the rate in L^r from (62). Thus, we complete the proof of the corollary.

5. Discussions.

5.1. Some remarks on the system (1).

Having completed main results for the semi-linear case (12), it now suffices to provide some amendable remarks surrounding the general system (1) and its regularization (10).

Uniqueness result. It is discernible that the regularized problem may have many solutions but those regularized solutions (if they exist) must converge to a unique true solution. Here we introduce collectively important steps, included in Lemma 5.1, to prove the uniqueness result for the time-reversed system (1) with the zero Dirichlet boundary condition. Then, from now onwards we will not come back to this issue in future publications for the regularization of this system. The technique we follow is mainly from [21, Chapter 6], which was used to study the large-time behavior of solutions to a linear class of initial-boundary value parabolic equations. Detailed proofs of the following results can be inspired from [27] for the observations in the semi-linear case (12) with Hölder nonlinearities and the nonlinear Robin-type boundary condition.

Setting the function space

$$W_T\left(\Omega\right) := C\left(\left[0,T\right];H_0^1\left(\Omega\right)\cap W^{2,\infty}\left(\Omega\right)\right) \cap L^\infty\left(0,T;H^2\left(\Omega\right)\right) \cap C^1\left(0,T;L^2\left(\Omega\right)\right),$$

we denote by $P_T\left(\Omega\right)$ the set of functions in $W_T\left(\Omega\right)$ such that they vanish on the boundary $\partial\Omega$ and at the moments $t \in \{0,T\}$, i.e.

$$P_T\left(\Omega\right) := \left\{u \in W_T\left(\Omega\right) : \left.u\right|_{\partial\Omega} = 0, \left.u\right|_{t=T} = 0, \left.u\right|_{t=0} = 0\right\}.$$

Then, for $\eta > 0$ we set

$$\lambda\left(t\right) = t - T - \eta. \tag{50}$$

In what follows, this function plays a prime factor to prove the backward uniqueness result. According to solid proofs in [27], it is also worth noting that Lemma 5.1 is essentially a Carleman estimate with the weight $\lambda^{-\frac{m}{k}}$; see [31] for the observation in this spirit.

Lemma 5.1. *Assume the diffusion $a_{ij}(x,t,\cdot,\cdot) \in C^1(\overline{Q_T})$ for $1 \le i,j \le N$ is such that it satisfies the strict ellipticity condition and the mapping $(\mathbf{p},\mathbf{q}) \mapsto a\left(x,t;\mathbf{p};\mathbf{q}\right)$*

is sesquilinear for $(\mathbf{p},\mathbf{q}) \in [L^2(\Omega)]^N \times \left[L^2(\Omega)\right]^{Nd}$. *For any* $v \in [P_T(\Omega)]^N$, *for any positive* m *and any positive real* k, *one has*

$$\left\|\lambda^{-\frac{m}{k}}\left(\nabla\cdot\left(a\left(v;\nabla v\right)\nabla v\right)-v_t\right)\right\|^2_{[L^2(Q_T)]^N} \geq \frac{m}{k}\left\|\lambda^{-\frac{m}{k}-1}v\right\|^2_{[L^2(Q_T)]^N} - D\left\|\lambda^{-\frac{m}{k}}\nabla v\right\|^2_{[L^2(Q_T)]^{Nd}},$$

where D *depends only on the bounds of* $\partial_t a$. *Moreover, if* $0 < T \leq \mu$ *for* $0 < \mu \leq \mu_0$ *and* $0 < \eta \leq \eta_0$ *sufficiently small, there exists a positive* K *independent of* m *such that*

$$K\left\|\lambda^{-\frac{m}{k}}\left(\nabla\cdot\left(a\left(v;\nabla v\right)\nabla v\right)-v_t\right)\right\|^2_{[L^2(Q_T)]^N} \geq \left\|\lambda^{-\frac{m}{k}-1}v\right\|^2_{[L^2(Q_T)]^N} + \frac{1}{2}\left\|\lambda^{-\frac{m}{k}}\nabla v\right\|^2_{[L^2(Q_T)]^{Nd}}, \tag{51}$$

for m *sufficiently large.*

Let u and v be the two solutions of the backward problem (1)-(2) in $[W_T(\Omega)]^N$. The difference system for $w = u - v$ reads as

$$\begin{aligned} w_t + \nabla\cdot\left(-a\left(x,t;w;\nabla w\right)\nabla w\right) = {} & F\left(x,t;u;\nabla u\right) - F\left(x,t;v;\nabla v\right) \\ & + \nabla\cdot\left(a\left(x,t;u;\nabla u\right)\nabla u\right) - \nabla\cdot\left(a\left(x,t;v;\nabla v\right)\nabla v\right) \\ & - \nabla\cdot\left(a\left(x,t;w;\nabla w\right)\nabla w\right), \end{aligned} \tag{52}$$

endowed with the zero Dirichlet boundary condition and the zero terminal condition.

Under the assumptions that a, F are Lipschitz-continous with respect to the nonlinear arguments $\mathbf{p},\mathbf{q}$ and that a satisfies the strict ellipticity condition, we can find a positive constant C such that from (52) the following differential inequality holds

$$\left|w_t + \nabla\cdot\left(-a\left(x,t;w;\nabla w\right)\nabla w\right)\right|^2 \leq C\left(|w|^2 + |\nabla w|^2\right). \tag{53}$$

Observe that $w \in [P_T(\Omega)]^N$, we can obtain the uniqueness result in $[P_T(\Omega)]^N$ for (1) by using (51), (53) and by choosing appropriately small values of μ_0 and η_0.

Nonlocal diffusion. We could meliorate the existence result (cf. Theorem 4.3) when the diffusion a in the system (1) is of the following physical types:

- $a = a(x,t)$ typically accounting for the anisotropic diffusion and possibly taxis processes;
- $a = a(t;u) = \max\left\{\theta_0, \theta_1 + \left|\int_\Omega u(x,t)\,dx\right|\right\} + \theta_2$ for some $\theta_0,\theta_1,\theta_2 > 0$. The diffusion in this form is controlled by the local movements of species involved in the evolution equation (see e.g. [1, 52] for the concrete biological motivation of this equation);
- $a = a\left(t;\|\nabla u\|^2_{L^2(\Omega)}\right) = \theta_3 + \int_\Omega |\nabla u|^2\,dx$ for some $\theta_3 > 0$ indicating a Kirchhoff-type diffusion model for e.g. flows through porous media.

Using the same argument in Theorem 4.3, it is worth mentioning that the convergence results obtained in (27) and (29) are sufficient to passing to the limit in the diffusion term involving the aforementioned forms. Consequently, the existence result remains true in these cases for any spatial dimensions d. However, this technique is not valid for the p-Laplacian equation inspired from the power-law type of Ohm's law in conductivity of electricity, which reads as

$$u_t - \nabla\cdot\left(|\nabla u|^{p-2}\nabla u\right) = F(u) \quad \text{for } p \geq 2, \tag{54}$$

due to the failure of passage to the limit. When $d = 1$, there is a possibility of proving this solvability by the embedding $H_0^1(\Omega) \subset L^\infty(\Omega)$.

Since we use the boundedness of the diffusion term as a key point in the convergence analysis, a slight improvement of Theorem 4.5 and Theorem 4.6 can be obtained when a is dependent of the gradient. In fact, assuming the source condition (compared to (37))

$$u \in C([0,T];\mathbb{W}) \cap L^\infty(0,T;W^{1,\infty}(\Omega)) \cap C^1\left(0,T;L^2(\Omega)\right), \tag{55}$$

one could suppose that $\underline{M} \ge \eta \|\nabla u\|_{L^\infty(Q_T)}$ for some $\eta > 0$ sufficiently small, somewhat similar to the concept of large diffusion in terms of A, to gain similar error bounds. Technically, the reason behind this assumption is to preserve the positivity of the gradient term in (43). In some physical problems, the small diffusion a would fit this circumstance because $\overline{M}$ now can be taken sufficiently large and then choosing $\underline{M}$ large is possible.

Locally Lipschitz-continuous nonlinearities. From now on, we extend the convergence analysis when the source term F locally depends on u and ∇u. In this scenario, we need the estimate (4) for the cut-off function F_ℓ introduced in Remark 2.1. Essentially, there are two main difficulties in the proofs.

- When exploring the difference equation in proof of Theorem 4.5 we confront with the difference term $F_{\ell^\varepsilon}(u_\beta^\varepsilon;\nabla u_\beta^\varepsilon) - F(u;\nabla u)$. Thus, estimating $\mathcal{I}_3$ in (39) would be problematic.
- This moment the constant C_2 in (44) and given by (45) would depend on ℓ^ε. Observe that the behavior of ℓ^ε should be increasing (when $\varepsilon \to 0$) as it approximates the source function F in (3). Therefore, this parameter must be formulated in a clear manner to ensure the convergence of our QR scheme.

These issues are really needed to elucidate because, as particularly mentioned in subsection 1.1, the local Lipschitz continuity of F is encountered in most of the significant equations in real-life applications. Here we sketch out some essential ideas that we can adapt to the proof of Theorem 4.5. Note that here we need the aid of the source condition (55).

At first, we choose the cut-off parameter $\ell^\varepsilon > 0$ such that

$$\ell^\varepsilon \ge \|u\|_{L^\infty(0,T;W^{1,\infty}(\Omega))}. \tag{56}$$

This way we solve the first issue because $F_{\ell^\varepsilon}(x,t;u;\nabla u) = F(x,t;u;\nabla u)$; cf. (3).

Taking into account the Lipschitz constant $L_F(\ell^\varepsilon) > 0$ and choosing an appropriate Young inequality, we get the estimate of $\mathcal{I}_3$ as follows:

$$\begin{aligned}\mathcal{I}_3 &\ge -\frac{e^{2\rho_\beta(t-T)}\underline{M}}{8L_F^2(\ell^\varepsilon)}\left\|F_{\ell^\varepsilon}\left(u_\beta^\varepsilon;\nabla u_\beta^\varepsilon\right) - F_{\ell^\varepsilon}(u;\nabla u)\right\|_{L^2(\Omega)}^2 - \frac{2L_F^2(\ell^\varepsilon)}{\underline{M}}\left\|w_\beta^\varepsilon\right\|_{L^2(\Omega)}^2 \\ &\ge -\frac{\underline{M}}{4}\left(\left\|w_\beta^\varepsilon\right\|_{L^2(\Omega)}^2 + \left\|\nabla w_\beta^\varepsilon\right\|_{[L^2(\Omega)]^d}^2\right) - \frac{L_F^2(\ell^\varepsilon)}{\underline{M}}\left\|w_\beta^\varepsilon\right\|_{L^2(\Omega)}^2.\end{aligned} \tag{57}$$

Henceforward, (43) remains the same when we put $\rho_\beta = C_1\log(\gamma(T,\beta)) + \frac{\underline{M}+1}{4} + \frac{L_F^2(\ell^\varepsilon)}{\underline{M}}$. With this choice, the constant C_2 in (45) is ε-dependent and given by

$$C_2(\ell^\varepsilon) := \frac{\underline{M}+1}{4} + \frac{L_F^2(\ell^\varepsilon)}{\underline{M}}. \tag{58}$$

Now observe (44) with this new C_2 in (58) and have in mind that the error estimate at $t=0$ (cf. Theorem 4.6) is of the order $\mathcal{O}\left(\log^{-\frac{1}{2}}(\gamma(T,\beta))\right)$. We only need to find a fine control of the term $e^{\frac{(T-t)L_F^2(\ell^\varepsilon)}{\underline{M}}}$ in such a way that its growth does not ruin the logarithmic rate of convergence. To do so, our strategy is the following: We take

$$\varrho := \varrho(\beta) = \sqrt{\frac{\underline{M}}{T}\log\left(\log^{\kappa}\left(\gamma\left(T,\beta\right)\right)\right)} > 0, \tag{59}$$

for some ε-independent constant $\kappa > 0$ being selected later. Then, we have

$$\lim_{\varepsilon\to 0^+}\varrho(\beta) = \infty. \tag{60}$$

If we choose $\Lambda^\beta := \sup L_F^{-1}\left\{(-\infty,\varrho(\beta)]\right\}$, then $L_F\left(\Lambda^\beta\right) = \varrho(\beta)$ and we also obtain

$$e^{\frac{(T-t)L_F^2\left(\Lambda^\beta\right)}{\underline{M}}} \le \log^{\kappa}\left(\gamma\left(T,\beta\right)\right). \tag{61}$$

Note also that by (60), $L_F^{-1}\left\{(-\infty,\varrho(\beta)]\right\} \neq 0$ and $\Lambda^\beta \in (0,\infty)$ is well-defined. Moreover, we can prove that $\lim_{\varepsilon\to 0^+}\Lambda^\beta = \infty$. Indeed, we suppose that there exists $C > 0$ such that $\Lambda^\beta \le C$ for β near the zero point. Since L_F is non-decreasing with respect to ℓ^ε, it holds $L_F(C) \ge L_F\left(\Lambda^\beta\right) = \varrho(\beta)$, which contradicts the fact already known (60). Now, for $\ell^\varepsilon \in \left(0,\Lambda^\beta\right]$ we deduce that

$$e^{\frac{(T-t)L_F^2(\ell^\varepsilon)}{\underline{M}}} \le \log^{\kappa}\left(\gamma\left(T,\beta\right)\right),$$

resulted from (61). This also indicates that we have identified a fine upper bound of the ℓ^ε-dependent Lipschitz constant L_F, and the error estimate (44) now becomes

$$\begin{aligned} &\left\|u_\beta^\varepsilon(\cdot,t) - u(\cdot,t)\right\|_{L^2(\Omega)}^2 + 2\underline{M}\int_t^T \left\|\nabla u_\beta^\varepsilon(\cdot,s) - \nabla u(\cdot,s)\right\|_{[L^2(\Omega)]^d}^2 ds \\ &\le \log^{2\kappa}\left(\gamma\left(T,\beta\right)\right)\gamma^{2C_1(T-t)}\left(T,\beta\right)\left(\varepsilon^2 + 2TC_0^2\gamma^{-2}\left(T,\beta\right)\|u\|_{C([0,T];\mathbb{W})}^2\right)e^{2TC_3}, \end{aligned} \tag{62}$$

where $C_3 := \frac{\underline{M}+1}{4}$ is no longer dependent of ℓ^ε.

Similar to proof of Theorem 4.6, we inherit from (62) to gain the error estimate at $t=0$ with the order $\mathcal{O}\left(\log^{\kappa-\frac{1}{2}}(\gamma(T,\beta))\right)$. Hence, together with (62) we choose $\kappa := \kappa(t) = \min\left\{C_1 t, \frac{1}{2}\right\} > 0$ to complete the convergence analysis in this case. On top of this, the choice of the cut-off parameter can be summarized by (56) and (59), working with sufficiently small values of ε.

No-flux boundary condition. Since our problem (1)-(2) is also present in population dynamics, the zero Neumann condition should be analyzed. In this case, we associate the regularized problem (13) with the boundary condition $-a\nabla u^\varepsilon\cdot\mathrm{n} = 0$, taking the place of the zero Dirichlet boundary condition in (11). Under this setting, the techniques used in the proofs of our main results can be applied in the same manner, focusing on the same structure of the weak formulation we have in (15) (where the test function ψ now belongs to the closed subspace of $H^1(\Omega)$ that satisfies the zero Neumann boundary condition) and the key equation (39) for the convergence analysis. Accordingly, the rates of convergence derived in Theorem 4.5 and Theorem 4.6 remain unchanged. Moreover, the strong convergence on the boundary is confirmed for $0 < t < T$ by the following trace inequality:

$$\int_t^T \left\| u_\beta^\varepsilon (\cdot,s) - u(\cdot,s) \right\|_{[L^2(\partial\Omega)]^N}^2 ds$$

$$\leq C_\Omega \left(\left\| u_\beta^\varepsilon - u \right\|_{[C([t,T];L^2(\Omega))]^N}^2 + \int_t^T \left\| \nabla \left(u_\beta^\varepsilon - u \right)(\cdot,s) \right\|_{[L^2(\Omega)]^{Nd}}^2 ds \right),$$

which yields the same rate as in Theorem 4.5.

5.2. Possible future generalizations of above results.

Gevrey class. It is worth noting that the property (7) remains true up to a compact Riemannian manifold, which is generally called the Sturm-Liouville decomposition. As a prominent example, the standard eigen-elements for a d-torus $\mathbb{T}^d = \mathbb{R}^d / (2\pi\mathbb{Z})^d$ are

$$\phi_p(x) = \prod_{j=1}^d e^{2\pi i p_j x_j}, \quad \mu_p = \sum_{j=1}^d (2\pi p_j)^2, \quad p_j \in \mathbb{N}, 1 \leq j \leq d, \ i = \sqrt{-1}.$$

In this scenario, Gevrey classes are popular in micro-local analysis for the propagation of wavefront set and in the study of analytic regularity for nonlinear evolution equations with periodic boundary data. A famous result of the Gevrey solvability for nonlinear analytic parabolic equations is recalled in an example of Appendix A. Here, our discussions focus on the preasymptotic error bounds for *approximation numbers* of periodic Gevrey-type spaces of analytic functions with connection to the Galerkin method.

For $0 < \alpha, p, q < \infty$, we denote by $\mathbb{G}_\alpha^{p,q}(\mathbb{T}^d)$ the Gevrey space that consists of all functions in $C^\infty(\mathbb{T}^d)$ and satisfies

$$\|u\|_{\mathbb{G}_\alpha^{p,q}(\mathbb{T}^d)} := \left(\sum_{k\in\mathbb{Z}^d} \exp\left(2\alpha \|k\|_p^q\right) \hat{u}_k \right)^{1/2} < \infty,$$

where $\hat{u}_k$ denotes the Fourier coefficient of u with respect to the frequency vector $k \in \mathbb{Z}^d$. By this definition, the norm $\left\| e^{\overline{M}T(-\Delta)} u \right\|_{L^2(\mathbb{T}^d)}$ in Example 3.3 is essentially $\|u\|_{\mathbb{G}^{2,2}_{\overline{M}T}(\mathbb{T}^d)}$. For $q \in (0,1)$, this space is the classical Gevrey classes that contain non-analytic functions, whilst for $q \geq 1$ all functions are real-analytic therein.

In approximation theory for Hilbert spaces, approximation numbers represent the worst-case error obtained when approximating a class of functions by projecting them onto the optimal finite-dimensional subspace. The basic reason lies in the information-based complexity that requires the rank $n \in \mathbb{N}$ of the optimal projection operator is sufficiently large ($n > 2^d$) to gain the classical error bounds, which is not substantially practical for high dimensions. Therefore, approximation numbers can be an excellent candidate to handle this context. In a nutshell, the connection between such approximation numbers and Galerkin schemes for a classical variational problem, where a parabolic problem can be involved, is clearly present in [36, Subsection 1.5] with references cited therein for a background of Gevrey classes.

Definition 5.2 (Approximation numbers). *Let X and Y be two Banach spaces. The norm of an operator $\mathcal{A} : X \to Y$ is denoted by $\|\mathcal{A}\|_{X\to Y}$. The nth approximation* *number ($n \in \mathbb{N}$) of an operator $\mathcal{T} : X \to Y$ is defined by*

$$a_n\left(\mathcal{T} : X \to Y\right) := \inf_{rank(\mathcal{A})<n} \left\|\mathcal{T} - \mathcal{A}\right\|_{X\to Y}.$$

Taking into account the Gevrey space that have been mentioned above, the approximation numbers of the embedding $\mathrm{Id} : \mathbb{G}^{2,2}_{\overline{M}T}(\mathbb{T}^d) \to L^2(\mathbb{T}^d)$ are bounded by

$$n^{-\frac{c_1\overline{M}T}{\log_2(1+d/\log_2(n))}} \le a_n\left(\mathrm{Id} : \mathbb{G}^{2,2}_{\overline{M}T}\left(\mathbb{T}^d\right) \to L^2\left(\mathbb{T}^d\right)\right) \le n^{-\frac{c_2\overline{M}T}{\log_2(1+d)}},$$

for $d \le n \le 2^d$ and $c_1, c_2 > 0$. This rigorous estimate is almost identical to the preasymptotic estimate for approximation numbers of the classical embedding $\mathrm{Id} : H^1\left(\mathbb{T}^d\right) \to L^2\left(\mathbb{T}^d\right)$, albeit $\mathbb{G}^{2,2}_{\overline{M}T}(\mathbb{T}^d)$ obviously contains smoother functions than $H^1\left(\mathbb{T}^d\right)$. On the other hand, the approximation numbers for the embedding $\mathrm{Id} : \mathbb{G}^{2,2}_{\overline{M}T}(\mathbb{T}^d) \to H^1\left(\mathbb{T}^d\right)$ are asymptotically identical when $1 \le n \le d$, while for the embedding $\mathrm{Id} : W^{1,\infty}(\mathbb{T}^d) \to L^2(\mathbb{T}^d)$, they are completely identical whenever $n \le 2^d$.

Eventually, all these numbers indicate that there is a possibility to choose a combination of an optimal dimensional subspace and a linear finite element algorithm such that an approximate numerical solution by Galerkin methods is a good candidate in L^2 and H^1 for the true solution satisfying (36). Consequently, the worst-case *a priori* error for the (low) n-dimensional subspace in this context behaves like that of the standard finite element methods (FEMs).

5.3. Concluding remarks. We have extended a modified quasi-reversibility (QR) method for backward quasi-linear parabolic systems with noise. Several rates of convergence have been derived, especially the rigorous error estimates in $L^r(\Omega)$ $(r \ge 2)$ and $H^1(\Omega)$, albeit many open questions remain unsolved. Although the spectral method that takes into account Duhamel's principle is not used, settings for filter regularized operators still rely on existence of the space $\mathbb{W}$, which usually plays a role as a class of Gevrey spaces in the existing trend of regularization for time-reversed nonlinear parabolic equations.

Our present contribution gives rise to some further interesting questions. Recently, we have only done with several error estimates which indicate the strong convergence of the regularization scheme. This typical convergence is not expected to be applied in the stochastic setting, but it can be designed to obtain an approximate solution in the FEM framework. In this sense, our theoretical analysis can be a key ingredient to establish regularized multiscale FEM schemes which deal with models in certain complex domains because spatial environments where population densities take place are usually not nice (e.g. porous media). Other open perspectives include the effective iterative QR method and also the presence of the Robin-type boundary condition describing e.g. the surface reaction in more complex scenarios.

Appendix A. Applications to existing models. Here, we examine four types of backward problems arising in many physical applications to show the applicability of our theoretical analysis. In order to show existing arguments on the *a priori* information (55) where $\mathbb{W}$ stands for a class of Gevrey spaces demonstrated in Example 3.3, we specify below the possible regularity assumptions for different models chosen from simple to complex, based on the analysis of the forward models. Note that $1 \le d \le 3$ are only considered due to the practical meaning.

A.1. Fisher-Burger equation. In a finite interval $[0,l]$ with periodic boundary condition, one concerns the following equation:

$$u_t + Cuu_x = Du_{xx} + Bu\left(1-u\right),$$

with B, C, D being positive constants, for simplicity.

This problem is performed as a combination between the classical Fisher and Burger equations. Here we can further consider the real analytic cases with respect to u of the nonlinear F which imply several types of modelling interactions between particles. We know that in [20] the authors proved the local weak solvability of the forward problem. In this sense, if the initial condition is sufficiently smooth, viz. $u_0 \in H^1(\Omega)$, then we obtain a unique solution $u \in \mathbb{G}_t^{2,2}$ for any $t \in [0,T^*]$ with T^* sufficiently small. This not only verifies that the Gevrey regularity on the true solution could be valid in some certain models, but also agrees with the mild restriction of time in the convergence results.

A.2. p-Laplacian equation. In a bounded domain with a Hölder boundary, we take into account the equation (54) with the zero Dirichlet boundary condition. Cf. [43], we can obtain the classical solution in $L^\infty(0,T;W_0^{1,\infty}(\Omega))$ when $u_0 \in W^{1,\infty}(\Omega)$. Together with the Fisher-Burger equation, we remark that these forward problems have interesting phenomena including e.g. profiles of extinction and blow-up in finite time, the instantaneous shrinking of the support from the diffusion coefficient. Depending on the situation one may need appropriate choices of the auxiliary parameter ρ_ε involved in the regularized problem to keep track of the arisen phenomena. Therefore, rigorous analysis of the regularized problem (10)-(11) will be considered in the forthcoming works.

A.3. Gray-Scott-Klausmeier model. Based on the one-dimensional setting with $\Omega = \mathbb{R}$ in [41], we set $u = (u_1, u_2)$ with $u_1 > 0$ to guarantee the positive-definite diffusion $a(u)$. Then the closed-form nonlinearities are

$$a\left(u\right) = \begin{pmatrix} 2u_1 & 0 \\ 0 & D \end{pmatrix}, \quad F\left(u, u_x\right) = \begin{pmatrix} Cu_{1x} + A\left(1-u_1\right) - u_1 u_2^2 \\ -Bu_2 + u_1 u_2^2 \end{pmatrix},$$

where the involved parameters A, B, C, D are positive.

This model describes the interaction between water u_1 and plant biomass u_2 in semiarid landscapes. The local well-posedness in $H^2(\mathbb{R})$ (cf. [41, Theorem 2.2]) enjoys the possibility of taking $W^{1,\infty}$ in (55) due to the embedding $W^{1,1}(\mathbb{R}) \subset L^\infty(\mathbb{R})$.

A.4. Shigesada-Kawasaki-Teramoto model. In a three-dimensional setting with no-flux boundary condition, we consider

$$a\left(u\right) = \begin{pmatrix} a_{10} + 2a_{11}u_1 + a_{12}u_2 & a_{12}u_1 \\ a_{21}u_2 & a_{20} + 2a_{22}u_2 + a_{21}u_1 \end{pmatrix},$$

where the non-negative coefficients a_{ij} satisfy $8a_{11} \ge a_{12}$ and $8a_{22} \ge a_{21}$ to fulfill the positive-definiteness of diffusion. The source term is taken as the Lotka-Volterra functions, which reads as

$$F\left(u\right) = \begin{pmatrix} \left(b_{10} - b_{11}u_1 - b_{12}u_2\right)u_1 \\ \left(b_{20} - b_{21}u_1 - b_{22}u_2\right)u_2 \end{pmatrix},$$

where the coefficients b_{ij} are non-negative.

This famous model plays a vital role in population dynamics for multi-species systems in which self- and cross-diffusion effects are participated. An included example is the Keller-Segel model for cell aggregation, structured by setting $a_{10} = a_{20} = 1$, $a_{11} = a_{12} = a_{21} = a_{22} = 0$, $a_{12} = -1$ with $F(u) = (0, u_1 - u_2)^T$. It is important to see that a does not need to be symmetric in this context. Concerning the forward problem, the existence of bounded weak solution, i.e. $u_i \in L^\infty(0,T;L^\infty(\Omega))$, is proven in [22] if the initial data $u_i^0 \in L^\infty(\Omega)$ for $i = 1,2$. Moreover, if $\nabla u_i \in L^\infty(0,T;L^\infty(\Omega))$, the uniqueness result is obtained. Essentially, observe that we can adapt the *a priori* argument $L^\infty(0,T;W^{1,\infty}(\Omega))$ in (55) to this model.

Acknowledgments. The authors would like to thank the anonymous referee(s) and the handling editor(s) for fruitful comments through the significant improvement of this paper. V.A.K acknowledges Prof. Daniel Lesnic (Leeds, UK) and Prof. Thorsten Hohage (Göttingen, Germany) for recent supports to his research inception.

REFERENCES

[1] R. Almeida, S. Antontsev, J. Duque, and J. Ferreira, *A reaction-diffusion model for the non-local coupled system: existence, uniqueness, long-time behaviour and localization properties of solutions*, IMA Journal of Applied Mathematics, 81 (2016), pp. 344–364.
[2] J. Bebernes and D. Eberly, *Mathematical Problems from Combustion Theory*, vol. 83 of Applied Mathematical Sciences, Springer-Verlag, 1989.
[3] G. Bluman and V. Shtelen, *Nonlocal transformations of Kolmogorov equations into the backward heat equation*, Journal of Mathematical Analysis and Applications, 291 (2004), pp. 419–437.
[4] L. Bourgeois and J. Darde, *About stability and regularization of ill-posed elliptic Cauchy problems: the case of Lipschitz domains*, Applicable Analysis, 89 (2009), pp. 1745–1768.
[5] L. Bourgeois and J. Darde, *A duality-based method of quasi-reversibility to solve the Cauchy problem in the presence of noisy data*, Inverse Problems, 26 (2010). 095016.
[6] V. Calvez, L. Gosse, and M. Twarogowska, *Travelling chemotactic aggregates at mesoscopic scale and bistability*, SIAM Journal on Applied Mathematics, 77 (2017), pp. 2224–2249.
[7] H. Cao, M. V. Klibanov, and S. V. Pereverzev, *A Carleman estimate and the balancing principle in the quasi-reversibility method for solving the Cauchy problem for the Laplace equation*, Inverse Problems, 25 (2009). 035005.
[8] A. Carasso, *Error bounds in the final value problem for the heat equation*, SIAM Journal on Mathematical Analysis, 7 (1976), pp. 195–199.
[9] A. Carasso, *Computing small solutions of Burgers' equation backwards in time*, Journal of Mathematical Analysis and Applications, 59 (1977), pp. 169–209.
[10] A. S. Carasso, *Overcoming Hölder continuity in ill-posed continuation problems*, SIAM Journal on Numerical Analysis, 31 (1994), pp. 1535–1557.
[11] A. S. Carasso, *Reconstructing the past from imprecise knowledge of the present: Effective non-uniqueness in solving parabolic equations backward in time*, Mathematical Methods in the Applied Sciences, 36 (2013), pp. 249–261.
[12] 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), pp. 344–367.
[13] R. Cherniha, J. R. King, and S. Kovalenko, *Lie symmetry properties of nonlinear reaction-diffusion equations with gradient-dependent diffusivity*, Communications in Nonlinear Science and Numerical Simulation, 36 (2016), pp. 98–108.
[14] M. Chipot and P. Quittner, *Handbook of Differential Equations: Stationary Partial Differential Equations*, vol. 1, Elsevier, 2004.
[15] J. Cholewa and T. Dlotko, *Global Attractors in Abstract Parabolic Problems*, Cambridge University Press, Cambridge, UK, 2000.
[16] C. Clason and M. V. Klibanov, *The quasi-reversibility method for thermoacoustic tomography in a heterogeneous medium*, SIAM Journal on Scientific Computing, 30 (2007), pp. 1–23.
[17] 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 (2018),

pp. 3–12.

[18] J. C. M. Duque, R. Almeida, S. N. Antontsev, and J. Ferreira, *The Euler-Galerkin finite element method for a nonlocal coupled system of reaction-diffusion type*, Journal of Computational and Applied Mathematics, 296 (2016), pp. 116–126.

[19] A. Eladdadi and P. Kim, *Mathematical Models of Tumor-Immune System Dynamics*, Springer Proceedings in Mathematics & Statistics, Springer, 2014.

[20] A. B. Ferrari and E. S. Titi, *Gevrey regularity for nonlinear analytic parabolic equations*, Communications in Partial Differential Equations, 23 (1998), pp. 424–448.

[21] A. Friedman, *Partial Differential Equations of Parabolic Type*, Prentice-Hall, 1982.

[22] G. Galiano, *On a cross-diffusion population model deduced from mutation and splitting of a single species*, Computers & Mathematics with Applications, 64 (2012), pp. 1927–1936.

[23] M. Hieber and J. Rehberg, *Quasilinear parabolic systems with mixed boundary conditions on nonsmooth domains*, SIAM Journal on Mathematical Analysis, 40 (2008), pp. 292–305.

[24] R. Jaroudi, G. Baravdish, F. Åström, and B. T. Johansson, *Source localization of reaction-diffusion models for brain tumors*, in Pattern Recognition, B. Rosenhahn and B. Andres, eds., vol. 9796 of Lecture Notes in Computer Science, Springer, Cham, 2016, pp. 414–425.

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

[26] N. I. Kavallaris, J. Lankeit, and M. Winkler, *On a degenerate nonlocal parabolic problem describing infinite dimensional replicator dynamics*, SIAM Journal on Mathematical Analysis, 49 (2017), pp. 954–983.

[27] V. A. Khoa, T. T. Hung, and D. Lesnic, *Uniqueness result for an age-dependent reaction-diffusion problem*, ArXiv e-prints, (2018), https://arxiv.org/abs/1806.03442.

[28] V. A. Khoa, M. T. N. Truong, N. H. M. Duy, and N. H. Tuan, *The Cauchy problem of coupled elliptic sine-Gordon equations with noise: Analysis of a general kernel-based regularization and reliable tools of computing*, Computers and Mathematics with Applications, 73 (2017), pp. 141–162.

[29] M. Kirane, E. Nane, and N. H. Tuan, *On a backward problem for multidimensional Ginzburg-Landau equation with random data*, Inverse Problems, 34 (2017). 015008.

[30] M. V. Klibanov, *Estimates of initial conditions of parabolic equations and inequalities via lateral Cauchy data*, Inverse Problems, 22 (2006).

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

[32] M. V. Klibanov, *Carleman weight functions for solving ill-posed Cauchy problems for quasilinear PDEs*, Inverse Problems, 31 (2015). 125007.

[33] M. V. Klibanov and F. Santosa, *A computational quasi-reversibility method for Cauchy problems for Laplace's equation*, SIAM Journal on Applied Mathematics, 51 (1991), pp. 1653–1675.

[34] R. J. Knops and L. E. Payne, *On the stability of solutions of the Navier-Stokes equations backwards in time*, Archive for Rational Mechanics and Analysis, 29 (1968), pp. 331–335.

[35] C. König, F. Werner, and T. Hohage, *Convergence rates for exponentially ill-posed inverse problems with impulsive noise:*, SIAM Journal on Numerical Analysis, 54 (2016), pp. 341–360.

[36] T. Kühn, S. Mayer, and T. Ulrich, *Counting via entropy: New preasymptotics for the approximation numbers of Sobolev embeddings*, SIAM Journal on Numerical Analysis, 54 (2016), pp. 3625–3647.

[37] J. Lankeit, *Equilibration of unit mass solutions to a degenerate parabolic equation with a nonlocal gradient nonlinearity*, Nonlinear Analysis, 135 (2016), pp. 236–248.

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

[39] K. J. Mahasa, R. Ouifki, A. Eladdadi, and L. de Pillis, *Mathematical model of tumor-immune surveillance*, Journal of Theoretical Biology, 404 (2016), pp. 312–330.

[40] M. Meyries, *Local well-posedness and instability of travelling waves in a chemotaxis model*, Advances in Differential Equations, 16 (2011), pp. 31–60.

[41] M. Meyries, J. D. M. Rademacher, and E. Siero, *Quasi-linear parabolic reaction-diffusion systems: A user's guide to well-posedness, spectra, and stability of travelling waves*, SIAM Journal on Applied Dynamical Systems, 13 (2014), pp. 249–275.

[42] K. Miller, *Stabilized quasi-reversibility and other nearly-best-possible methods for non-well-posed problems*, in Symposium on Non-Well-Posed Problems and Logarithmic Convexity, vol. 316 of Lecture Notes in Mathematics, Springer, 1967, pp. 161–176.

[43] M. Nakao, *On the initial-boundary value problem for some quasilinear parabolic equations of divergence form*, Journal of Differential Equations, 263 (2017), pp. 8565–8580.

[44] P. T. Nam, *An approximate solution for nonlinear backward parabolic equations*, Journal of Mathematical Analysis and Applications, 367 (2010), pp. 337–349.

[45] P. T. Nam, D. D. Trong, and N. H. Tuan, *The truncation method for a two-dimensional nonhomogeneous backward heat problem*, Applied Mathematics and Computation, 216 (2010), pp. 3423–3432.

[46] C. V. Pao, *Nonlinear Parabolic and Elliptic Equations*, Springer, 1992.

[47] T. Regińska and L. Eldén, *Solving the sideways heat equation by a wavelet-Galerkin method*, Inverse Problems, 13 (1997), pp. 1093–1106.

[48] T. Regińska and K. Regiński, *Approximate solution of a Cauchy problem for the Helmholtz equation*, Inverse Problems, 22 (2006), pp. 975–989.

[49] J. Rubinstein and P. Sternberg, *Nonlocal reaction-diffusion equations and nucleation*, IMA Journal of Applied Mathematics, 48 (1992), pp. 249–264.

[50] B. Sandstede and A. Scheel, *On the structure of spectra of modulated travelling waves*, Mathematische Nachrichten, 232 (2001), pp. 39–93.

[51] R. E. Showalter, *Cauchy Problem for Hyper-parabolic Partial Differential Equations*, Trends in the Theory and Practice of Non-Linear Analysis, Elsevier, 1983.

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

[53] N. H. Tuan, B. T. Duy, N. D. Minh, and V. A. Khoa, *Hölder stability for a class of initial inverse nonlinear heat problem in multiple dimension*, Communications in Nonlinear Science and Numerical Simulation, 23 (2015), pp. 89–114.

[54] N. H. Tuan and P. H. Quan, *Some extended results on a nonlinear ill-posed heat equation and remarks on a general case of nonlinear terms*, Nonlinear Analysis: Real World Applications, 12 (2011), pp. 2973–2984.

[55] D. Wei, *Existence, uniqueness, and numerical analysis of solutions of a quasilinear parabolic problem*, SIAM Journal on Numerical Analysis, 29 (1992), pp. 484–497.

[56] M. C. Zlatescu, A. TehraniYazdi, H. Sasaki, J. F. Megyesi, R. A. Betensky, D. N. Louis, and J. G. Cairncross, *Tumor location and growth pattern correlate with genetic signature in oligodendroglial neoplasms*, Cancer Research, 61 (2001), pp. 6713–6715.