Electronic Transactions on Numerical Analysis.
Volume 64, pp. 80–108, 2025.

ISSN 1068–9613.
DOI: 10.1553/etna_vol64s80

ETNA
Kent State University and
Johann Radon Institute (RICAM)

BAYESIAN IDENTIFICATION OF MATERIAL PARAMETERS IN VISCOELASTIC STRUCTURES AS AN INVERSE PROBLEM IN A SEMIGROUP SETTING*

REBECCA ROTHERMEL† AND THOMAS SCHUSTER‡

Abstract. This paper considers the nonlinear inverse problem of identifying the material parameters in viscoelastic structures based on a generalized Maxwell model. The aim is to reconstruct the model parameters from stress data acquired from a relaxation experiment, where the number of Maxwell elements, and thus the number of material parameters themselves, is assumed to be unknown. This implies that the forward operator acts on a Cartesian product of a semigroup (of integers) and a Hilbert space, and thus demands an extension of existing regularization theory. We develop a stable reconstruction procedure by applying Bayesian inversion to this setting. We use an appropriate binomial prior that takes the integer setting for the number of Maxwell elements into account, and at the same time computes the underlying material parameters. We extend the regularization theory for inverse problems to this special setup, and prove the existence, stability, and convergence of the computed solution. The theoretical results are evaluated by extensive numerical tests.

Key words. viscoelastic material, Bayesian inversion, semigroup, Maxwell model, binomial prior, inverse problem

AMS subject classifications. 65L09, 74H75

1. Introduction.

We consider the problem of identifying material parameters in viscoelastic structures. Parameter identification represents a challenging class of inverse problems having important and demanding real-world applications. These include, for example, the identification of the distortion energy density of hyperelastic materials [2, 19, 30, 49, 50, 57], the surface-enthalpy-dependent heat fluxes of steel plates [43, 44], inverse scattering problems [7], the estimation of parameters from waveform information [10], the inverse kinematic problem [31], electrical impedance tomography [4], or terahertz tomography [55] to name only a few. All these problems have in common that they are usually nonlinear and (locally) ill-posed. This means that even small errors in the measured data lead to large inaccuracies in the computed solution, if no regularization is applied. There exists a vast amount of literature for solving inverse problems in Hilbert and Banach spaces, such as, e.g., [12, 27, 29, 33, 40, 48]. The problem that is considered in the present paper needs an extension of the existing theory to spaces that show only the structure of a semigroup.

Understanding how a material deforms when a force is applied is essential for many industrial applications, ranging from food processing, additive manufacturing, structural health monitoring [15], seismics [41], to product design [3].

There are several publications dealing with the identification of viscoelastic parameters. However, these differ from the present work by several features. In [11, 52] viscoelastic structures are considered and their material parameters identified, but instead of the generalized Maxwell model they use different modeling. A comparable model is used in [9]. There, however, only the stiffness of a viscoelastic material is reconstructed. The relaxation times as well as the number of Maxwell elements are assumed to be known. Relaxation experiments in combination with cyclic tests are used in [46] to determine the basic stiffness and the material parameters of four Maxwell elements. There, the number of Maxwell elements is known a priori; see also [14, 37, 51]. Babaei et al. [1] propose two methods for solving the underlying

*Received April 7, 2024. Accepted December 28, 2025. Published online on January 19, 2026. Recommended by Frank Werner.
†Department of Mathematics, Saarland University, 66123 Saarbrücken, Germany
‡Corresponding author. Department of Mathematics, Saarland University, 66123 Saarbrücken, Germany (thomas.schuster@num.uni-sb.de).

ETNA
Kent State University and
Johann Radon Institute (RICAM)

inverse problem. The so-called *ad hoc method* first guesses the number of Maxwell elements n and subsequently reconstructs the material parameters for fixed n. Our approach differs from this procedure, since the computation of n is part of the algorithm and the computation of the parameters is done simultaneously. Their second approach [1], the *discrete spectral approach*, distributes about 1000 relaxation times equidistantly in a logarithmic scale over the interval $[10^{-1}, 10^3]$. This drastically simplifies the reconstruction and only requires solving a system of linear equations to calculate the stiffnesses. Then, the dominant parameters are identified and the correct number of Maxwell elements is found, although this process and the handling of the remaining parameters are not described in more detail.

We use a generalized Maxwell model, which is characterized by an unknown number of Maxwell elements n and material parameters (relaxation times τ_j and stiffnesses μ, μ_j). This means that the exact solution determines at the same time the number of parameters to be determined, a feature that has to be included in the modeling process. The forward operator, whose construction is outlined in Section 2.2, maps the material parameters and the number of Maxwell elements to the stress function. This function describes the time history of stress in the material during a relaxation experiment, where a strain is applied to the material and kept constant. The forward operator acts on a Cartesian product of a semigroup (the integers $\mathbb{N}$) and a Hilbert space ($\ell^2(\mathbb{N})$). In Section 2.3 we define the inverse problem of determining the number of Maxwell elements and simultaneously the material parameters from the stress function.

Unfortunately, the use of a large number of parameters often leads to overfitting a noisy data term and thus to unavoidable errors in the parameters. In a previous paper [45], the authors developed a clustering algorithm adapted to this problem. But it shows a strong error susceptibility to noisy data. As a solution, we propose in Section 3 a novel method using statistical Bayesian inversion theory. This uses a binomial prior to estimate the number of Maxwell elements. The developed algorithm alternately searches for a suitable solution for n and the material parameters by minimizing an appropriate Tikhonov functional. The minimization is done in $\mathbb{N} \times \ell^2(\mathbb{N})$ and, hence, the standard theory for regularizing Tikhonov functionals does not apply, since it uses Hilbert and Banach space settings (cf. [12, 20, 22, 48]). To this end we extend the regularization theory for Tikhonov functionals to this particular setting, where we consider the integers $\mathbb{N}$ as a topological semigroup endowed with the discrete topology. Proofs for the convergence and stability of our proposed algorithm is the subject of Section 4.

Numerical validation of the theoretical framework is done in Section 5. For this purpose, we introduce the clustering algorithm from [45] in order to compare the reconstruction results of the different algorithms. We perform experiments using different exact and noise-perturbed data sets. We analyze different displacement rates of the strain function in the relaxation experiment, and consider the effect of the success probability associated with the binomial distribution of the prior. Additionally, we introduce different penalty terms with respect to the material parameters and analyze their influence on the reconstruction results.

Summarizing, the paper contains the following innovations:

- extending regularization theory to forward operators acting on the Cartesian product $\mathbb{N} \times \ell^2(\mathbb{N})$, where the integers form a topological semigroup,
- dependence of the number of material parameters on the number of Maxwell elements n, and thus on part of the inverse problem solution itself,
- Bayesian inversion approach using a binomial prior, and
- proof of regularization property (existence, stability, and convergence of solutions) for our approach.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

2. A rheological model for viscoelastic material behavior.

2.1. Viscoelastic materials. We introduce the rheological model of a viscoelastic material, which we use as the starting point. For comprehensive introductions to the phenomenological behavior and modeling of viscoelasticity, we refer to textbooks such as [53, 56]. Figure 2.1 shows the typical shape of a standard specimen that is clamped at the thick ends and loaded in the direction of the arrows. According to Saint Venant's principle (see [54]), the disturbances caused by the clamping at the ends of the specimen decay after a short distance. Therefore, we can assume that the strain and stress state at the center of the specimen is homogeneous and we can use a one-dimensional model.

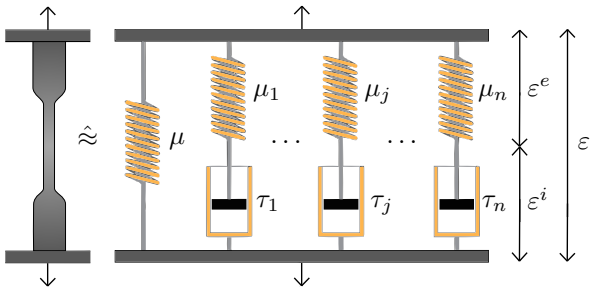

FIG. 2.1. *Standard specimen and generalized Maxwell model with unknown number of Maxwell elements n and $2n+1$ material parameters.*

The force and extension of the specimen are measured and can be used to make direct calculations from the strain and stress. We consider the following relaxation experiment: For a given strain rate $\dot{\varepsilon}$ and a maximum strain value $\bar{\varepsilon}$, the strain in a time interval $t \in [0, T]$ is given as

$$\varepsilon(t) = \begin{cases} \dot{\varepsilon}\, t, & 0 \le t \le \bar{\varepsilon}/\dot{\varepsilon}, \\ \bar{\varepsilon}, & \bar{\varepsilon}/\dot{\varepsilon} < t \le T. \end{cases}$$

The function is plotted in Figure 2.2 and describes the following procedure: The material is stretched until a maximum strain value $\bar{\varepsilon}$ is reached at a strain rate $\dot{\varepsilon}$. The strain rate is $\dot{\varepsilon} = \dot{\varepsilon}_u / l_0$ with the displacement rate $\dot{\varepsilon}_u$, given by the testing device, and specimen length l_0. Thus, the maximum strain is reached at time $t = \bar{\varepsilon}/\dot{\varepsilon}$. After that, the applied strain is kept constant.

The simplest model of a viscoelastic solid is a three-parameter model consisting of a parallel combination of a Maxwell element with a spring (compare [17, 25, 46]). If one applies a strain to such a solid model, both springs will stretch. If the strain is then kept constant during the rest of the experiment, the damper expands according to its relaxation time. To incorporate this behavior into our model, we will use an arbitrary, unknown number of Maxwell elements instead of a single one. This leads us to the generalized Maxwell model, also known as the *Maxwell–Wiechert model* (compare Figure 2.1). The number of Maxwell elements can be expanded to any number n, with the relaxation time τ_j in each of the dampers and the stiffness of the spring μ_j in each of the Maxwell elements. The stiffness of the individual spring is denoted by μ. If all Maxwell elements relax to zero stress value, the equilibrium position is

ETNA
Kent State University and
Johann Radon Institute (RICAM)

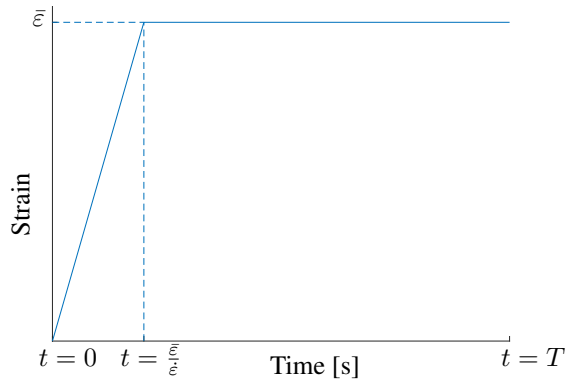

FIG. 2.2. *Strain curve $\varepsilon(t)$ with strain rate $\dot{\varepsilon}$ and maximum strain value $\bar{\varepsilon}$.*

reached, and the single spring with stiffness μ represents the basic stiffness of the material, which ensures that the material does not behave like a liquid.

The deformation of the material represented by the strain value ε is divided into an elastic component ε_j^e and an inelastic component ε_j^i in each of the Maxwell elements. The elastic component corresponds to the strain of the spring and the inelastic component to the strain of the damper. Since the strain of the damper depends on its relaxation time, an evolution equation based on the entropy principle is used to model the change of the inelastic strain with time [24, 39]. For small deformations, the evolution equation is given as

$$\dot{\varepsilon}_j^i(t) = \frac{\varepsilon(t) - \varepsilon_j^i(t)}{\tau_j/2}, \tag{2.1}$$

where $\dot{\varepsilon}_j^i$ represents the time derivative. The total stress generated in the system is then given by the sum of the stresses induced in each of the springs. Assuming linear elastic behavior of the springs, this results in

$$\sigma(t) = \mu\,\varepsilon(t) + \sum_{j=1}^{n} \mu_j(\varepsilon(t) - \varepsilon_j^i(t)). \tag{2.2}$$

The total stress (2.2) is the sum of the stress of the individual spring, as well as the stresses in the various Maxwell elements. The latter depend on the inelastic components of the strain, which are determined by the evolution equation (2.1).

Apart from mechanics [39], the evolution equation (2.1) has also been used in other applications such as magnetic particle imaging [8] to model relaxation. Its solution can be formulated analytically as

$$\varepsilon_j^i(t) = \int_0^t \varepsilon(\tilde{t})\frac{2}{\tau_j}\exp\left(-2\frac{t-\tilde{t}}{\tau_j}\right)d\tilde{t}.$$

If we use this solution, we can also represent the total stress analytically (cf. [45]). The stress σ_j of the jth Maxwell element with $j > 1$ and the corresponding stiffness μ_j and relaxation time τ_j at time $t \in [0, T]$ is then represented as

$$\sigma_j(t) = \frac{\mu_j \tau_j \dot{\varepsilon}}{2}\,\eta(t, \tau_j), \tag{2.3}$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

with

$$\eta(t,\tau) := \left(\exp\left(-\frac{2}{\tau}\left[t-\frac{\bar{\varepsilon}}{\dot{\varepsilon}}\right]^{+}\right) - \exp\left(-\frac{2}{\tau}t\right)\right) \tag{2.4}$$

where $[x]^+ := \max\{0, x\}$. The function $\eta(t,\tau)$ has properties that will prove useful later on and are summarized in the following lemma.

LEMMA 2.1. *The function $\eta(t,\tau)$ is bounded on $[0,T]\times(0,+\infty)$, with $|\eta(t,\tau)| \le \bar{\eta}$. For fixed $t \in [0,T]$, it is uniformly continuous on all closed intervals $[\gamma_1,\gamma_2]$ with $0<\gamma_1<\gamma_2<+\infty$ and differentiable on $(0,+\infty)$. Furthermore, $\eta(\cdot,\tau)$ is continuous on $[0,T]$ for all $\tau\in(0,+\infty)$.*

The stress of the single spring, denoted by σ_0, can be specified with the corresponding stiffness μ as

$$\sigma_0(t) = \begin{cases} \mu\dot{\varepsilon}t, & 0 \le t \le \bar{\varepsilon}/\dot{\varepsilon}, \\ \mu\bar{\varepsilon}, & \bar{\varepsilon}/\dot{\varepsilon} < t \le T. \end{cases} \tag{2.5}$$

According to these calculations, the total stress can be written as the sum of the stresses of the single spring and of the Maxwell elements, i.e.,

$$\sigma(t) = \sum_{j=0}^{n} \sigma_j(t). \tag{2.6}$$

In view of Lemma 2.1, $\sigma(t)$ is continuous on $[0,T]$. The inverse problem, finally, consists of identifying the material parameters from the measured stress using equation (2.2). The material parameters include the stiffness of the individual spring μ, as well as the stiffnesses $\mu_1,\dots,\mu_n$ and relaxation times $\tau_1,\dots,\tau_n$ of the Maxwell elements. We emphasize that in our problem setting the number of Maxwell elements n is assumed to be unknown and has to be computed as part of the solution, which, on the other hand, affects the number of material parameters to be determined. In this view the investigated inverse problem goes beyond existing approaches. The strain ε is known and the inelastic strain ε_j^i must be determined for all Maxwell elements $i=1,\dots,n$ from equation (2.1).

REMARK 2.2. We note that existing and well-established methods for parameter identification in exponential sums such as Prony's method, MUSIC, or ESPRIT (see, e.g., [21, 38, 42]) cannot be extended to our situation in a straightforward way, since $\sigma(t)$ in (2.6) is not a pure exponential sum. Moreover, the mentioned methods need a known number of Maxwell elements n. Hence, our method goes beyond the setting used in these methods.

2.2. The forward operator. The forward mapping associated with the inverse problem described in Section 2.1 is given by

$$F : \mathcal{D}(F) \subset \mathbb{N}\times\ell^2(\mathbb{N}) \to L^2([0,T]), \quad (n,x)\mapsto F(n,x) := \sigma,$$

where the sequence $x = (\mu,\mu_1,\tau_1,\dots,\mu_n,\tau_n,0,0,\dots) \in \ell^2(\mathbb{N})$ contains the material parameters. Hence, F maps $n\in\mathbb{N}$ and $x\in\ell^2(\mathbb{N})$ to the stress $\sigma(t)$ from (2.3)–(2.6) with the number of Maxwell elements n. Since the number of material parameters depends on n, we cannot predict a priori how many unknowns have to be determined. This is why we choose $x\in\ell^2(\mathbb{N})$ allowing for a variable number of material parameters.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

We fix a finite, physically reasonable subset $I \subset \mathbb{N}$ representing the number of Maxwell elements, and set

$$\mathcal{D}(F) = \{(n,x) \in I \times \ell^2(\mathbb{N}) : \quad x := \{x_m\}_{m\in\mathbb{N}},\ x_m \in \mathbb{R}_0^+,\ m \in \mathbb{N};\\ x_m = 0,\ m > 2n+1;\ x_{2i+1} \geq \gamma,\ i = 1,\dots,n\}$$

with an arbitrarily small but fixed number $\gamma > 0$. The stiffnesses $\mu, \mu_1, \dots, \mu_n$ are nonnegative. From (2.3) it follows that $\tau_i > 0$ for $i = 1,\dots,n$. So, there exists an (artificial) parameter $\gamma > 0$ with $\tau_i \geq \gamma$ implying that $\mathcal{D}(F)$ is closed. This will be relevant in Section 4. Without loss of generality, we stick to the convention $\tau_1 \leq \tau_2 \leq \cdots \leq \tau_n$. For fixed n, we denote by $F_n : \mathcal{D}(F_n) \subset \ell^2(\mathbb{N}) \to L^2([0,T])$ the mapping $F_n(x) := F(n,x)$. Then, we have

$$\mathcal{D}(F_n) = \{x := \{x_m\}_{m\in\mathbb{N}} \in \ell^2(\mathbb{N}) :\ x_m \in \mathbb{R}_0^+,\ m \in \mathbb{N};\\ x_m = 0,\ m > 2n+1;\ x_{2i+1} \geq \gamma,\ i = 1,\dots,n\}. \tag{2.7}$$

2.3. The inverse problem. The inverse problem is formulated as computing n and x as the solution of

$$F(n,x) = \sigma^\delta \tag{2.8}$$

from given, maybe noise-contaminated, stress measurements σ^δ with $\|\sigma^\delta - \sigma\| \leq \delta$. Note that F depends on the discrete variable n as well as on the sequence x, where n especially represents the number of material parameters to be determined corresponding to the nonzero elements in $x \in \ell^2(\mathbb{N})$. Thus the number of material parameters depends on parts of the solution. This is an unusual situation in the field of inverse problems and differs from the theory in classical textbooks [12, 27, 33, 40]. We develop an iterative solver relying on statistical inversion theory leading to the minimization of a Tikhonov functional

$$T_\alpha(n,x) := \tfrac{1}{2}\|F(n,x) - \sigma^\delta\|^2 + \alpha\Omega(n,x),$$

where $\alpha > 0$ is a regularization parameter and Ω represents an appropriate penalty term.

3. A regularization method based on Bayes inversion. In this section we use Bayes inversion theory to develop a regularization method for (2.8). We first focus on the reconstruction of n, the number of Maxwell elements.

Let Σ, N, X, and E be random variables, where the stress σ is interpreted as the realization $\Sigma = \sigma$, the random variable N attains integer values and describes the number of Maxwell elements, X are the material parameters, and E is the additive noise.[1] Thus, the inverse problem (2.8) can then be written as

$$\Sigma = F(N,X) + E.$$

We determine the maximum a posteriori (MAP) estimator using the Bayes theorem and obtain

$$n_{\text{MAP}} = \arg\max_{n\in\mathbb{N}} \rho_0(n)\rho(\sigma|n)$$

or, equivalently,

$$n_{\text{MAP}} = \arg\min_{n\in\mathbb{N}}\{-\log(\rho(\sigma|n)) - \log(\rho_0(n))\}.$$

[1] We emphasize that in Section 3 the variable σ denotes the stress and not variance.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

We introduce the notation $\Phi(n|\sigma) := -\log(\rho(n|\sigma))$, $\phi(\sigma|n) := -\log(\rho(\sigma|n))$, and $\phi_0(n) := -\log(\rho_0(n))$, turning the minimization problem into

$$n_{\text{MAP}} = \arg\min_{n\in\mathbb{N}} \Phi(n|\sigma) = \arg\min_{n\in\mathbb{N}} \{\phi(\sigma|n) + \phi_0(n)\}.$$

We assume that N and E as well as X and E are stochastically independent. Let $\rho_{\text{noise}}(e)$ be the probability density of the noise. Then,

$$\rho(\sigma|n) = \rho_{\text{noise}}(\sigma - F(n,x)).$$

In our application, we suppose $E \sim \mathcal{N}(0, a^2)$, that is,

$$\rho_{\text{noise}}(e) = \frac{1}{\sqrt{2\pi a^2}} \exp\left(\frac{-\|e\|^2}{2a^2}\right),$$

whence

$$-\log(\rho_{\text{noise}}(e)) = -\log\left(\frac{1}{\sqrt{2\pi a^2}}\right) + \frac{1}{2a^2}\|e\|^2$$

follows. Thus, the likelihood function reads as

$$\phi(\sigma|n) = \phi_{\text{noise}}(\sigma - F(n,x)) = -\log(\rho_{\text{noise}}(\sigma - F(n,x))) \propto \|\sigma - F(n,x)\|^2.$$

Minimizing $\phi(\sigma|n)$ for n is, thus, equivalent to minimizing $\|\sigma - F(n,x)\|^2$.

As prior, we use the binomial distribution. It describes the number of successes in a series of independent trials that have two possible outcomes. Let M be the number of trials and $0 < q < 1$ be the probability of success in a single trial. Then,

$$B(n|q,M) = \binom{M}{n} q^n (1-q)^{M-n}$$

is the probability of achieving n successes. The usage of a binomial distribution for n is motivated by the common strategy of subdividing the interval of relaxation times into subintervals of different decades and assuming that each decade contains a Maxwell element; see [46]. In our approach, the probability of every one of the M intervals containing an element is given by q, yielding more flexibility in the model. So, the approach can be described as guessing the number of Maxwell elements in a clever way and subsequently determining the material parameter by minimizing $\phi(\sigma|n)$ for x.

To do this, let $I \subset \mathbb{N}$ be a finite set of integers and $M := \max I$ the maximum possible number of Maxwell elements. Then, the prior is given as $\rho_0(n) = B(n|q,M)$ and $\phi_0(n) = -\log(\rho_0(n))$. Thus, minimizing

$$n_{\text{MAP}} = \arg\min_{n\in\mathbb{N}} \left\{ \|\sigma^\delta - F(n,x)\|^2 - \log\left(\binom{M}{n} q^n (1-q)^{M-n}\right)\right\}$$

with $0 < q < 1$ results in finding the maximum a posteriori distribution n_{MAP}. To evaluate this functional, x must be calculated for each considered n. Section 3.2 explains this step in more detail. To control the influence of the penalty term, we use a regularization parameter $\alpha > 0$ and change the minimization finally to

$$n_{\text{MAP}} = \arg\min_{n\in\mathbb{N}} \left\{ \|\sigma^\delta - F(n,x)\|^2 - \alpha \log\left(\binom{M}{n} q^n (1-q)^{M-n}\right)\right\}. \tag{3.1}$$

REMARK 3.1. It might be confusing that we define the prior for arbitrary I and, thus, M. But the consideration of different materials can lead to different intervals I in $\mathcal{D}(F)$, and the developed framework has to be applicable for all these choices. Another interpretation of setting the prior to $B(n|q,M)$ is that this is the distribution conditioned to the specific definition of $\mathcal{D}(F)$.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

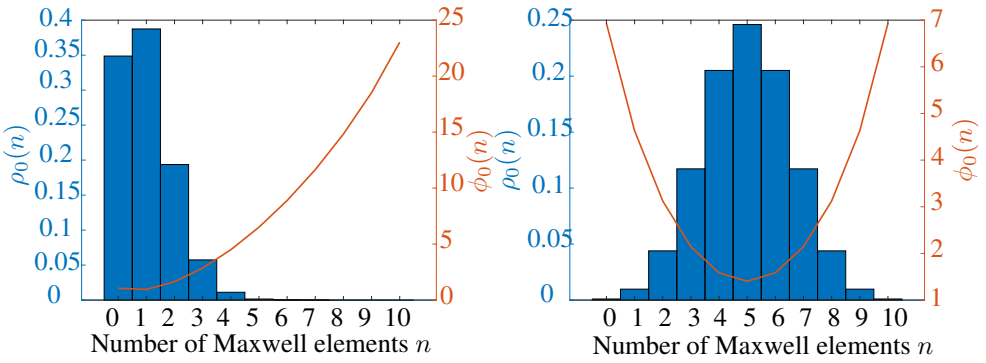

FIG. 3.1. *The binomial distribution $\rho_0(n)$ and $\phi_0(n)$ for $q = 0.1$ (left) and $q = 0.5$ (right).*

3.1. Impact of q on the minimization. The probability density of the prior ρ_0 depends on the choice of success probability $q \in (0, 1)$. We briefly investigate how different values of q affect the prior ρ_0 and, thus, ϕ_0. For this purpose, we assume a maximum number of $M = 10$ Maxwell elements. On the left-hand side of Figure 3.1 one can see the binomial distribution $\rho_0(n) = B(n|q = 0.1, M = 10)$ (blue bars), and $\phi_0(n) = -\log(B(n|0.1, 10))$ (red line) for $n = 0, \dots, M$. The right-hand side of the figure shows the corresponding plots for $q = 0.5$.

Since $\phi_0(n)$ increases with n, a large number of Maxwell elements are penalized in (3.1). Figure 3.1 shows the priors $\rho_0(n)$ and the corresponding penalties $\phi_0(n)$ for different values of $q = 0.1$ and $q = 0.5$. Obviously small success probabilities q penalize a large number of Maxwell elements significantly, whereas larger q favor larger n. So, in order to get a good fit of the measured data with only a small number of elements, it is recommended to keep q small. The numerical results in Section 5.5 will confirm that. In this view q acts as an additional regularization parameter enforcing sparsity with respect to x for small q.

3.2. The Bayes algorithm. In this section we develop an alternating iterative algorithm for the stable approximate solution of (2.8). The idea is as follows: For fixed $n \in I$ compute a minimizer $x_n^\beta \in \mathcal{D}(F_n)$ of the Tikhonov functional

$$\min_{x \in \mathcal{D}(F_n)} T_{\beta,n}(x) := \min_{x \in \mathcal{D}(F_n)} \{\tfrac{1}{2}\|F(n, x) - \sigma^\delta\|^2 + \alpha\Omega(x)\} \tag{3.2}$$

with a penalty term $\Omega : \ell^2(\mathbb{N}) \to \mathbb{R}$, which serves as an additional stabilizer, and a corresponding regularization parameter $\beta > 0$. The existence of a unique minimizer is dealt with in Section 4.2.

Numerically, we use an iterative solution method with several initial values for

$$(\mu^{(0)}, \mu_1^{(0)}, \tau_1^{(0)}, \dots, \mu_n^{(0)}, \tau_n^{(0)}, 0, 0, \dots).$$

The initial values are distributed over the possible range of parameter values. The use of multiple initial values is necessary because of the nonlinearity of the problem causing the existence of multiple local minima. For each of these initial values, the Tikhonov functional (3.2) is minimized using the functions *lsqnonlin* and *MultiStart* from the MATLAB Optimization Toolbox [34]. The function *lsqnonlin* is a subspace trust-region method and is based on the interior-reflective Newton method described in [5, 6]. In this way, we define a mapping

$$R : I \to \bigcup_{n \in I} \mathcal{D}(F_n) \subset \ell^2(\mathbb{N})$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

by $n \mapsto x_n^\beta$. After the computation of x_n^β, we minimize

$$\min_{n\in\mathbb{N}} \Phi(n|\sigma) = \min_{n\in\mathbb{N}} \left\{ \|F(n, x_n^\beta) - \sigma^\delta\|^2 - \alpha \log\left(\binom{M}{n} q^n (1-q)^{M-n} \right) \right\}.$$

Note that such a minimizer exists since I is finite. Denoting the minimizer by $\bar{n}$, we start the procedure again with n replaced by $\bar{n}$.

The algorithm needs the following input values:

- the success probability $0 < q < 1$, where $q \to 1$ favors a large number of Maxwell elements n,
- measured data σ^δ, and
- $I := \{n_1, \dots, n_m\} \subset \mathbb{N}$ with $n_1 < \cdots < n_m$, where $n_m = \max I =: M$ is the pre-selected maximum number of elements.

As initialization of the algorithm, we choose $k = 1$ and $n_k := n_1$, which correspond to the smallest possible number of Maxwell elements. Subsequently, we compute $x_1^\alpha = R(n_1)$ and calculate $\Phi(n_1|\sigma)$. In the next step, we want to determine $l \in \{1, \dots, m-k\}$ satisfying

$$\Phi(n_{k+l}|\sigma) < \Phi(n_k|\sigma). \tag{3.3}$$

This means that we have to compute $x_{k+l}^\alpha = R(n_{k+l})$ and n_{k+l} from (3.1) for each l to evaluate the expressions $\Phi(n_{k+l}|\sigma)$. If an appropriate $l \in \{1, \dots, m-k\}$ satisfying (3.3) is not found, then we stop the iteration, accepting (n_k, x_k) as the result. Otherwise we replace $k := k + l$ and repeat the procedure to find l with (3.3) until the stopping criterion is achieved. Summarizing, the algorithm reads as follows:

ALGORITHM 3.1.
Input: $0 < q < 1, M := \max I := \max\{n_1, \dots, n_m\},\ \sigma \in L^2([0,T]),\ \alpha, \beta > 0$
1) Let $k = 1$ *and* $n_k := n_1$, *compute* $x_k^\beta := R(n_k)$.
2) For $l = 1, \dots, m-k$ *and while* $\Phi(n_{k+l}|\sigma) \geq \Phi(n_k|\sigma)$ *calculate* $x_{k+l}^\beta = R(n_{k+l})$ *to evaluate* $\Phi(n_{k+l}|\sigma)$.
3) If no such l *exists, then STOP. Otherwise proceed with step 2).*
Output: (n_k, x_k^β)

Algorithm 3.1 is an iteration scheme minimizing alternately $\Phi(n|\sigma)$ by (3.1) and $T_{\alpha,n}$ with respect to x. Since I is finite, the existence of a minimizing n_k is guaranteed and we obtain that Algorithm 3.1, if we set $\alpha = \beta$, iteratively computes a minimizer of the Tikhonov functional

$$\begin{aligned} &\min_{(n,x)\in\mathcal{D}(F)} T_\alpha(n,x) \\ &\quad := \min_{(n,x)\in\mathcal{D}(F)} \left\{ \frac{1}{2}\|F(n,x) - \sigma^\delta\|^2 + \alpha \left[\Omega(x) - \log\left(\binom{M}{n} q^n (1-q)^{M-n} \right) \right] \right\} \end{aligned} \tag{3.4}$$

with a (possible generic) regularization parameter $\alpha > 0$.

4. Convergence, stability, and regularization in $\mathbb{N} \times \ell^2(\mathbb{N})$. This section is devoted to prove the convergence, stability, and regularization properties of Algorithm 3.1. We extend the well-established regularization theory to $\mathbb{N} \times \ell^2(\mathbb{N})$, which is necessary since it does not apply to discrete semigroups as $\mathbb{N}$. Using the new framework, we are able to prove the convergence, stability, and regularization properties for Algorithm 3.1.

A comprehensive treatise of regularization methods can be found in the textbooks [12, 33, 40]. Nonlinear inverse problems in particular are the subject of [27]. In this paper we

ETNA
Kent State University and
Johann Radon Institute (RICAM)

focus on Tikhonov regularization; see, e.g., [13, 27, 36, 47]. Other well-known regularization methods are the Landweber method [18, 26, 28], Newton-type methods [28], or the method of approximate inverse [32]. We follow the lines in [48, Chap. 4.1] to prove the convergence, stability, and regularization properties for Algorithm 3.1. Before that, we show that (2.8) is locally ill-posed.

4.1. Local ill-posedness of (2.8). We consider $\mathbb{N} \times \ell^2(\mathbb{N})$ with the product topologies (τ_d, τ_0) and (τ_d, τ_w). Here, τ_0 represents the strong topology in $\ell^2(\mathbb{N})$ induced by the ℓ^2-norm

$$\|x\|_{\ell^2(\mathbb{N})} := \left(\sum_{n=1}^{\infty} x_n^2 \right)^{1/2} \quad \text{for } x \in \ell^2(\mathbb{N})$$

and τ_w denotes the weak topology in $\ell^2(\mathbb{N})$. Furthermore, τ_d denotes the discrete topology induced by the discrete metric

$$d(x,y) := \begin{cases} 0, & x = y, \\ 1, & x \neq y, \end{cases} \qquad \text{for } x, y \in \mathbb{N}.$$

A sequence $\{(n_m, x_m)\}$ in $\mathbb{N} \times \ell^2(\mathbb{N})$ converges if and only if $\{n_m\}$ converges in $\mathbb{N}$ and $\{x_m\}$ converges in $\ell^2(\mathbb{N})$. A sequence $\{n_m\} \subset \mathbb{N}$ converges with respect to τ_d if and only if it is constant starting from a certain element, i.e., there exists an $m_0 \in \mathbb{N}$ with $n_m = \text{const}$ for all $m \geq m_0$.

By

$$\mathcal{U}(n) := \{n - m, \dots, n + l\}$$

for some $m, l \in \mathbb{N}_0$ we denote a τ_d-neighborhood of $n \in \mathbb{N}$ in $\mathbb{N}$.

DEFINITION 4.1. *The mapping $F(n,x) = \sigma$ is called* locally well-posed *in $(n^+, x^+) \in \mathcal{D}(F)$ if $F_n(x) = \sigma$ is locally well-posed for all $n \in \mathcal{U}(n^+) \cap I$ and all neighborhoods $\mathcal{U}(n^+)$ of n^+. The mapping $F(n,x) = \sigma$ is called* locally ill-posed *in $(n^+, x^+) \in \mathcal{D}(F)$ if there exists a neighborhood $\mathcal{U}(n^+)$ such that $F_n(x) = \sigma$ is locally ill-posed for some $n \in \mathcal{U}(n^+)$ in the sense that, for arbitrarily small radii $r > 0$, there exists a sequence $\{x_k\} \subset \bar{B}_r(x_0) \cap \mathcal{D}(F_n)$ such that*

$$F_n(x_k) \to F(x_0) \quad \text{in } L^2([0,T]) \quad \text{but} \quad x_k \not\to x_0 \quad \text{in } \mathbb{N} \times \ell^2(\mathbb{N}), \quad \text{for } n \to \infty.$$

THEOREM 4.1. *There exists a neighborhood $\mathcal{U}(n^+)$ of n^+ such that $F_n(x) = \sigma$ is locally ill-posed for some $n \in \mathcal{U}(n^+) \cap I$.*

Proof. Let $r > 0$ and $x^+ = (\mu^+, \mu_1^+, \tau_1^+, \dots, \mu_{n^+}^+, \tau_{n^+}^+, 0, 0, \dots)$ be fixed. Choose $\mathcal{U}(n^+)$ such that $n = n^+ + 1 \in \mathcal{U}(n^+)$, and

$$x_k^r = (\mu^+, \mu_1^+, \tau_1^+, \dots, \mu_{n^+}^+, \tau_{n^+}^+, r/2k, r/2, 0, 0, \dots).$$

Then, $(n, x_k^r) \in \mathcal{D}(F)$ for all $k \in \mathbb{N}$. Furthermore, it follows that

$$\|x_k^r - x^+\|_{\ell^2(\mathbb{N})}^2 = \frac{r^2}{4k^2} + \frac{r^2}{4} = \left(\frac{1}{4k^2} + \frac{1}{4} \right) r^2 < r^2 \quad \text{for all } k \in \mathbb{N}.$$

This yields $\{x_k^r\}_{k \in \mathbb{N}} \subset B_r(x^+)$ and from

$$\|x_k^r - x^+\|_{\ell^2(\mathbb{N})}^2 \to \frac{r^2}{4} \qquad \text{as } k \to \infty$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

we deduce that $x_k^r \not\rightharpoonup x^+$ as $k \to \infty$. We compute

$$\begin{aligned}\|F_n(x_k^r) - F_{n^+}(x^+)\|^2_{L^2([0,T])} &= \int_0^T \left(\frac{(r/2k)(r/2)\dot{\varepsilon}}{2}\, \eta\left(t, \frac{r}{2}\right)\right)^2 dt \\ &= \frac{r^4\dot{\varepsilon}}{8k^2} \int_0^T \eta\left(t, \frac{r}{2}\right)^2 dt \to 0\end{aligned}$$

as $k \to \infty$ since $\eta(\cdot, \tau)$ is continuous and bounded (Lemma 2.1). ∎

4.2. Algorithm 3.1 as a regularization method. As outlined in Section 3.2, we solve the inverse problem by minimizing (3.4), where $0 < q < 1$ is the success probability, M is the maximum number of Maxwell elements, α is a regularization parameter, and Ω is a penalty term satisfying Assumption 4.2. For simplicity, we define

$$\operatorname{logg}(n) := -\log\left(\binom{M}{n} q^n (1-q)^{M-n}\right).$$

For fixed $n \in I$ we consider the problem

$$F_n(x) = \sigma, \quad x \in \mathcal{D}(F_n) \subseteq \ell^2(\mathbb{N}), \quad \sigma \in F_n(\mathcal{D}(F_n)) \subseteq L^2([0,T]), \tag{4.1}$$

with $\mathcal{D}(F_n)$ the domain of F_n and

$$F_n(\mathcal{D}(F_n)) = \{\tilde{\sigma} \in L^2([0,T]) : F_n(\tilde{x}) = \tilde{\sigma} \text{ for } \tilde{x} \in \mathcal{D}(F_n)\}$$

the image of F_n. The inverse problem (4.1) is solved by minimizing the Tikhonov functional (3.2). As already mentioned, we follow the lines in [48, Chap. 4.1]. For the sake of better readability, here we repeat two assumption blocks taken from [48] (reproduced as Assumptions 4.1 and 4.2), which are essential for turning the minimization of general Tikhonov functionals

$$T_\alpha(x) := \frac{1}{p}\|F(x) - y^\delta\|^p + \alpha\Omega(x) \tag{4.2}$$

into a regularization method. The first block makes general assumptions about the Banach spaces X and Y, as well as the operator $F : \mathcal{D}(F) \subset X \to Y$ and its domain $\mathcal{D}(F)$.

ASSUMPTION 4.1.
(a) X and Y are infinite-dimensional reflexive Banach spaces.
(b) $\mathcal{D}(F)$ is a closed and convex subset of X.
(c) For $x_n \rightharpoonup x_0$ in X, $x_n \in \mathcal{D}(F)$, $n \in \mathbb{N}$, and $x_0 \in \mathcal{D}(F)$, it follows that $F(x_n) \rightharpoonup F(x_0)$ in Y, i.e., $F : \mathcal{D}(F) \subseteq X \to Y$ is weak-to-weak sequentially continuous.

The next block of assumptions are associated with the penalty term $\Omega(x)$ of the Tikhonov functional $T_\alpha(x)$ and the Ω-minimizing solution x^+ of the operator equation.

ASSUMPTION 4.2.
(a) For the exponent p in (4.2), it holds that $1 < p < \infty$.
(b) Ω is a proper, *convex, and lower semicontinuous functional, where "proper" means that the domain of Ω is nonempty. Furthermore*

$$\mathcal{D} := \mathcal{D}(F) \cap \mathcal{D}(\Omega) \neq \emptyset.$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

(c) It is assumed that Ω is a stabilizing functional in the sense that the subsets

$$\mathcal{M}_\Omega(c) := \{x \in X : \Omega(x) \le c\}$$

in X for all $c \ge 0$ are weakly sequentially precompact.

(d) There exists a Ω-minimizing solution x^+ of the equation $F(x) = y$, which belongs to the so-called Bregman domain

$$\mathcal{D}_B(\Omega) := \{x \in \mathcal{D} \subseteq X : \partial\Omega(x) \neq \emptyset\},$$

where $\partial\Omega(x) \subseteq X^*$ denotes the subdifferential of Ω at point x.

Penalty terms of the form

$$\Omega(x) := \frac{1}{q}\|x\|_X^q, \quad 1 \le q < \infty, \tag{4.3}$$

are always stabilizing functionals, if X is a reflexive Banach space, as postulated by Assumption 4.1(a). The proof of this result can be found in [48] or [35, p. 251].

We show that Assumptions 4.1 and 4.2 are satisfied for (3.2) by proving that a regularized solution x_α^δ exists, that it is stable with respect to the data, and that, given certain conditions, the regularized solution converges weakly.

Let us start by proving Assumption 4.1 for $F_n : \mathcal{D}(F_n) \subset \ell^2(\mathbb{N}) \to L^2([0,T])$:

(a) $X = \ell^2(\mathbb{N})$ and $Y = L^2([0,T])$ are infinite-dimensional Hilbert spaces and, thus, reflexive.

(b) The domain $\mathcal{D}(F_n)$ is closed and convex. Let $\{x^{(k)}\}_{k\in\mathbb{N}}$ be a sequence in $\mathcal{D}(F_n) \subset \ell^2(\mathbb{N})$. Hence, $x^{(k)} := \{x_m\}_{m\in\mathbb{N}}^{(k)}$ for $k \in \mathbb{N}$. Moreover, let $x^{(k)} \to x^*$ with $x^* \in \ell^2(\mathbb{N})$ as $k \to +\infty$. But convergence in $\ell^2(\mathbb{N})$ implies componentwise convergence. Thus, it immediately follows that $\mathcal{D}(F_n)$ is closed from its definition (2.7).

Let $\{a_m\}, \{b_m\} \in \mathcal{D}(F_n)$ and $\lambda \in [0,1]$. We prove that

$$\{c_m\} := \lambda\{a_m\} + (1-\lambda)\{b_m\} \in \mathcal{D}(F_n).$$

Since $a_m = b_m = 0$, it follows that $c_m = 0$ for $m > 2n+1$. Considering c_m for fixed $m \le 2n+1$ as a function of λ, we get $dc_m/d\lambda = a_m - b_m$. For $a_m \ge b_m$ the smallest value for $\lambda = 0$ can be found with $c_m = b_m$. For $a_m < b_m$ the minimum is $c_m = a_m$ with $\lambda = 1$. In both cases, $c_m \ge 0$ holds for all $m \le 2n+1$ and $c_{2i+1} \ge \gamma$ for all $i = 1, \dots, n$. Thus, $\{c_m\} \in \mathcal{D}(F_n)$, proving the convexity of $\mathcal{D}(F_n)$.

(c) We show that F_n is weak-to-weak sequentially continuous. Let $\{x^{(k)}\}$ be a sequence in $\mathcal{D}(F_n)$ with $x^{(k)} \rightharpoonup x^*$ in $\ell^2(\mathbb{N})$ and $x^* \in \mathcal{D}(F)$. Our aim is to prove the convergence

$$F_n(x^{(k)}) =: \sigma^{(k)} \rightharpoonup F_n(x^*) := \sigma^* \quad \text{in } L^2([0,T]).$$

By the uniform boundedness principle, every weakly convergent sequence is bounded. Moreover, in $\ell^2(\mathbb{N})$, it holds that a sequence is weakly convergent if and only if it is bounded and converges component-by-component. In $L^2([0,T])$ a sequence converges weakly, if it is bounded and converges pointwise (see [23]). We apply this in the following steps:

1. We show that F_n maps a bounded sequence $\{x^{(k)}\} \subset \mathcal{D}(F_n) \subset \ell^2(\mathbb{N})$ to a bounded sequence.
2. We show that from weak convergence in $\ell^2(\mathbb{N})$, and thus the boundedness and component-by-component convergence of $x_m^{(k)} \to x_m^*$, the pointwise convergence of $\sigma^{(k)} \xrightarrow{\text{pointwise}} \sigma^*$ in $L^2([0,T])$ follows.

From 1 and 2, we obtain the weak convergence $\sigma^{(k)} \rightharpoonup \sigma^*$.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

Proofs of 1 and 2.

1. We have $x^{(k)} = \{x_m^{(k)}\}_{m\in\mathbb{N}}$ for all $k \in \mathbb{N}$. Moreover, $\{x^{(k)}\}$ is bounded as a weakly convergent sequence, that is, there exists $S \in \mathbb{R}$, such that $\|x^{(k)}\|_{\ell^2\mathbb{N}} = \left(\sum s_{m=1}^{\infty}|x_m^{(k)}|^2\right)^{1/2} \leq S$ holds for all $k \in \mathbb{N}$. Since $x^{(k)} \in \mathcal{D}(F_n)$ we have $x_m^{(k)} = 0$ for $m > 2n+1$ and $x_m^{(k)} \geq 0$ for $m \leq 2n+1$. It follows that $x_m^{(k)} < S$ for all $m, k \in \mathbb{N}$. Let $k \in \mathbb{N}$. Then,

$$\begin{aligned}\|\sigma^{(k)}\|^2_{L^2([0,T])} &= \int_0^T \left(\sigma_0^{(k)}(t) + \sum_{j=1}^{n} \frac{x_{2j}^{(k)} x_{2j+1}^{(k)} \dot{\varepsilon}}{2}\, \eta(t, x_{2j+1}^{(k)})\right)^2 dt \\ &\leq T\left(S\bar{\varepsilon} + \frac{nS^2\dot{\varepsilon}}{2}\,\bar{\eta}\right)^2 < \infty,\end{aligned}$$

where we used (2.4) and Lemma 2.1. Thus, $\sigma^{(k)}$ is bounded for all $k \in \mathbb{N}$, implying that the sequence $\{F_n(x^{(k)})\}$ is bounded.

2. We have component-by-component convergence $x_m^{(k)} \xrightarrow{k} x_m^*$ for all $m \in \mathbb{N}$, which means $\forall\, m \in \mathbb{N},\ \forall \zeta > 0,\ \exists\, K(m) \in \mathbb{N}$, such that $\forall\, k \geq K(m)$, the following holds: $|x_m^{(k)} - x_m^*| < \zeta$. Since $x^{(k)} \in \mathcal{D}(F_n),\ k \in \mathbb{N},\ x^* \in \mathcal{D}(F)$, and thus $x_m^{(k)} = x_m^* = 0$ for $m > 2n+1,\ k \in \mathbb{N}$, we can choose $K := \max\{K(m) : m = 1, \ldots, 2n+1\}$, such that $\forall\, \zeta > 0,\ \exists\, K \in \mathbb{N}$ with $\forall\, m \in \mathbb{N},\ \forall\, k \geq K$, the following holds: $|x_m^{(k)} - x_m^*| < \zeta$. We want to show that $\sigma^{(k)} \to \sigma^*$ pointwise, which means $\forall\, t \in [0,T],\ \forall\, \omega > 0,\ \exists\, M \in \mathbb{N}$, such that $\forall\, k \geq M$, the following holds: $|\sigma^{(k)}(t) - \sigma^*(t)| < \omega$.

Let $t \in [0,T]$, $\omega < 0$ arbitrarily, and $k \geq K$. We estimate

$$\begin{aligned}&|\sigma^{(k)}(t) - \sigma^*(t)| \\ &\quad\leq \left|x_1^{(k)} - x_1^*\right|\bar{\varepsilon}T + \sum_{j=1}^{n}\left|\frac{x_{2j}^{(k)} x_{2j+1}^{(k)}\dot{\varepsilon}}{2}\,\eta(t, x_{2j+1}^{(k)}) - \frac{x_{2j}^* x_{2j+1}^*\dot{\varepsilon}}{2}\,\eta(t, x_{2j+1}^*)\right| \\ &\quad\leq \zeta\bar{\varepsilon}T + \frac{S^2\dot{\varepsilon}}{2}\sum_{j=1}^{n}\left|\eta(t, x_{2j+1}^{(k)}) - \eta(t, x_{2j+1}^*)\right|.\end{aligned}$$

Since $\eta(t,\cdot)$ is uniformly continuous on the closed interval $[\gamma, \max\{\tau_j\}]$ (Lemma 2.1), we obtain

$$|\sigma^{(k)}(t) - \sigma^*(t)| < \omega,$$

only if K is sufficiently large. This shows the pointwise convergence of $\sigma^{(k)} \to \sigma^*$ as $k \to \infty$. □

This completes the proofs of 1 and 2, and thus of all three parts of Assumption 4.1.

Next, we turn to Assumption 4.2 (applied to (3.2)). The exponent in (3.2) is $p = 2$ and, thus, Assumption 4.2(a) is satisfied. Assumptions 4.2(b) and (c) depend on the used penalty term $\Omega(x)$. For the numerical evaluations in Section 5.6 we apply the penalties

$$\Omega_1(x) = \tfrac{1}{2}\|x\|^2 \quad \text{and} \quad \Omega_2(x) = \tfrac{1}{2}x_3^2$$

for $x \in \ell^2(\mathbb{N})$ and $0 < \gamma \leq x_3 \in \mathbb{R}$. Both penalties, Ω_1 as well as Ω_2, have a nonempty domain and $\mathcal{D} := \mathcal{D}(F) \cap \mathcal{D}(\Omega_i) \neq \emptyset$ holds true for $i = 1, 2$. Both functionals are convex and continuous, which also implies the lower semicontinuity. Thus, Assumption 4.2(b) is satisfied for Ω_1 and Ω_2. As already mentioned, penalty terms $\Omega(x)$ of the form (4.3) satisfy

ETNA
Kent State University and
Johann Radon Institute (RICAM)

Assumption 4.2(c). So, Assumption 4.2(c) is fulfilled for Ω_1, but not for Ω_2. Nevertheless, we use Ω_2 in the numerical experiments in Section 5.6, since investigations in [9] have shown that the reconstruction of the smallest relaxation time is always ill-conditioned and numerical tests show that such a penalty, indeed, stabilizes this process. The existence of a Ω-minimizing solution x^+ to the problem (4.1) which belongs to the Bregman domain $\mathcal{D}_B(\Omega)$ is postulated such that Assumption 4.2(d) is also satisfied.

In the following we aim to prove the regularization property of Algorithm 3.1. To this end we need the concept of a Ω-minimizing solution.

DEFINITION 4.2. *We call* $(n^+, x^+) \in \mathcal{D}(F) \subseteq \mathbb{N} \times \ell^2(\mathbb{N})$ *a* logg-Ω-minimizing solution *of* (2.8) *if*

$$F(n^+, x^+) = \sigma$$

and

$$\operatorname{logg}(n^+) + \Omega(x^+) = \inf\{\operatorname{logg}(\tilde{n}) + \Omega(\tilde{x}) : (\tilde{n}, \tilde{x}) \in \mathcal{D}(F),\ F(\tilde{n}, \tilde{x}) = \sigma\}$$

holds true for exact data $\sigma \in L^2([0,T])$ *in* (2.8).

Since $\mathbb{N} \times \ell^2(\mathbb{N})$ is a Cartesian product of a semigroup and a Hilbert space, we extend the definition of a regularization method to such structures.

DEFINITION 4.3 (Regularization in $\mathbb{N} \times \ell^2(\mathbb{N})$). *A mapping that assigns each pair* $(\sigma^\delta, \alpha) \in L^2([0,T]) \times (0, \bar{\alpha})$ *with* $0 < \bar{\alpha} \leq +\infty$ *to well-defined elements* $(n_\alpha^\delta, x_\alpha^\delta) \in \mathcal{D}(F)$ *is called* regularization (regularization method) *for* (2.8) *if there is a suitable choice* $\alpha = \alpha(y^\delta, \delta)$ *of the regularization parameter, such that for any sequence* $\{\sigma_m\}_{m=1}^\infty \subset L^2([0,T])$ *with*

$$\|\sigma_m - \sigma\| \leq \delta_m \quad \text{and} \quad \delta_m \to 0 \quad \text{for } m \to \infty,$$

the corresponding regularized solutions $(n^{\delta_m}_{\alpha(y_m,\delta_m)}, x^{\delta_m}_{\alpha(y_m,\delta_m)})$ *converge in a well-defined sense to a solution* (n^+, x^+) *of* (2.8). *Since the latter is in general not unique, regularized solutions must converge to* logg-Ω*-minimizing solutions of* (2.8). *In the case of nonuniqueness, different subsequences of regularized solutions may converge to different* logg-Ω*-minimizing solutions of* (2.8).

For convergence, we consider primarily the product topologies (τ_d, τ_0) and (τ_d, τ_w). We have shown that for all $\alpha > 0$, $n \in I$, and $\sigma^\delta \in L^2([0,T])$, there exists a regularized solution $x_\alpha^\delta \in \mathcal{D}(F_n)$ minimizing the functional $T_{\alpha,n}(x)$ (3.2) in $\mathcal{D}(F_n)$. Since I is finite, there exists a solution $(n_\alpha^\delta, x_\alpha^\delta)$ that minimizes the functional $T_\alpha(n,x)$ (3.4) in $(n,x) \in \mathcal{D}(F)$.

THEOREM 4.2 (Stability). *For all* $\alpha > 0$, *the minimizers of* (3.4) *are stable with respect to the data* σ^δ. *Let* $\{\sigma_m\}$ *be a sequence of data with* $\lim_{n\to\infty} \|\sigma_m - \sigma^\delta\|_{L^2([0,T])} = 0$ *and let* $\{(n_m, x_m)\}$ *be the corresponding sequence of minimizers of*

$$\min_{(n,x)\in\mathcal{D}(F)} \{\tfrac{1}{2}\|F(n,x) - \sigma_m\|^2 + \alpha[\operatorname{logg}(n) + \Omega(x)]\}.$$

Then, $\{(n_m, x_m)\}$ *has a* (τ_d, τ_w)*-convergent subsequence* $\{(n_{m_k}, x_{m_k})\}$. *The limit of each such subsequence is a minimizer* $(n_\alpha^\delta, x_\alpha^\delta)$ *of* (3.4). *Moreover, for any such* (τ_d, τ_w)*-convergent subsequence*

$$\lim_{k\to\infty} (\operatorname{logg}(n_{m_k}) + \Omega(x_{m_k})) = \operatorname{logg}(n_\alpha^\delta) + \Omega(x_\alpha^\delta)$$

holds true.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

Proof. Since $\{n_m\} \subset I$ and I is finite, there must be a τ_d-convergent subsequence $\{n_{m_k}\}$, which means that there exists a $k_0 \in \mathbb{N}$ such that for all $k \geq k_0$ we have $n_{m_k} = \tilde{n}$. For $k \geq k_0$, x_{m_k} can be considered as the minimizer of

$$\tfrac{1}{2}\|F_{\tilde{n}}(x) - \sigma_m\|^2 + \alpha\Omega(x)$$

for fixed $\tilde{n}$. Then, by [48, Prop. 4.2], there exists a subsequence $\{x_{m_k}\}$ with $x_{m_k} \rightharpoonup \tilde{x}$, where $\tilde{x}$ is a minimizer of

$$T_{\alpha,\tilde{n}}(x) = \tfrac{1}{2}\|F_{\tilde{n}}(x) - \sigma^\delta\|^2 + \alpha\Omega(x)$$

and

$$\lim_{k\to\infty} \Omega(x_{m_k}) = \Omega(x_\alpha^\delta).$$

Moreover, $F(n_{m_k}, x_{m_k}) = F_{\tilde{n}}(x_{m_k}) \rightharpoonup F_{\tilde{n}}(\tilde{x})$ as $k \to \infty$, since $F_{\tilde{n}}$ is weak-to-weak sequentially continuous, and hence it follows that $F(n_{m_k}, x_{m_k}) - \sigma_{m_k} \rightharpoonup F(\tilde{n}, \tilde{x}) - \sigma^\delta$ for $k \to \infty$. Since the L^2-norm and Ω are weakly lower semicontinuous, we deduce

$$\begin{aligned}\tfrac{1}{2}\|F(\tilde{n}, \tilde{x}) - \sigma^\delta\|^2 &\leq \liminf_{k\to\infty} \tfrac{1}{2}\|F(n_{m_k}, x_{m_k}) - \sigma_{m_k}\|^2,\\ \Omega(\tilde{x}) &\leq \liminf_{k\to\infty} \Omega(x_{m_k}).\end{aligned}$$

Since for $k \geq k_0$ we have $\operatorname{logg}(n_{m_k}) = \operatorname{logg}(\tilde{n})$, it furthermore follows that

$$\operatorname{logg}(\tilde{n}) \leq \liminf_{k\to\infty} \operatorname{logg}(n_{m_k})$$

and thus

$$\begin{aligned}&\tfrac{1}{2}\|F(\tilde{n}, \tilde{x}) - \sigma^\delta\|^2 + \alpha[\operatorname{logg}(\tilde{n}) + \Omega(\tilde{x})]\\ &\quad\leq \liminf_{k\to\infty}(\tfrac{1}{2}\|F(n_{m_k}, x_{m_k}) - \sigma_{m_k}\|^2 + \alpha[\operatorname{logg}(n_{m_k}) + \Omega(x_{m_k})])\\ &\quad\leq \lim_{k\to\infty} \tfrac{1}{2}\|F(n, x) - \sigma_{m_k}\|^2 + \alpha[\operatorname{logg}(n) + \Omega(x)]\\ &\quad= \tfrac{1}{2}\|F(n, x) - \sigma^\delta\|^2 + \alpha[\operatorname{logg}(n) + \Omega(x)],\end{aligned}$$

where $(n, x) \in \mathcal{D}(F)$ is arbitrary. This shows that $(\tilde{n}, \tilde{x})$ is a minimizer of (3.4). If we choose $x = \tilde{x}$ and $n = \tilde{n}$ on the right-hand side, it follows that

$$\begin{aligned}&\tfrac{1}{2}\|F(\tilde{n}, \tilde{x}) - \sigma^\delta\|^2 + \alpha[\operatorname{logg}(\tilde{n}) + \Omega(\tilde{x})]\\ &\quad= \lim_{k\to\infty}(\tfrac{1}{2}\|F(n_{m_k}, x_{m_k}) - \sigma_{m_k}\|^2 + \alpha[\operatorname{logg}(n_{m_k}) + \Omega(x_{m_k})]).\end{aligned}$$

Moreover, $\lim_{k\to\infty} \Omega(x_{m_k}) = \Omega(x_\alpha^\delta)$ and $\lim_{k\to\infty} \operatorname{logg}(n_{m_k}) = \operatorname{logg}(n_\alpha^\delta)$ hold. Thus, we finally get

$$\lim_{k\to\infty} (\operatorname{logg}(n_{m_k}) + \Omega(x_{m_k})) = \operatorname{logg}(n_\alpha^\delta) + \Omega(x_\alpha^\delta),$$

completing the proof. ∎

We finally prove convergence of the method and, in this way, the regularization property.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

THEOREM 4.3 (Convergence). *Let* $\{\sigma_m := \sigma^{\delta_m}\} \subset L^2([0,T])$ *be a sequence of perturbed data, and* $\sigma \in F(\mathcal{D}(F))$ *be exact data with* $\|\sigma - \sigma_m\| \leq \delta_m$ *for a sequence* $\{\delta_m > 0\}$ *of noise levels converging monotonically to zero. Moreover, consider a sequence* $\{\alpha_m > 0\}$ *of regularization parameters and an associated sequence* $\{(n_m := n_{\alpha_m}^{\delta_m},\ x_m := x_{\alpha_m}^{\delta_m})\}$ *of regularized solutions which are minimizers of*

$$\tfrac{1}{2}\|F(n,x) - \sigma_m\|^2 + \alpha_m[\operatorname{logg}(n) + \Omega(x)] \quad \textit{for } (n,x) \in \mathcal{D}(F).$$

Under the conditions

$$\limsup_{m\to\infty}(\operatorname{logg}(n_m) + \Omega(x_m)) \leq \operatorname{logg}(n_0) + \Omega(x_0) \tag{4.4}$$
$$\forall\, (n_0,x_0) \in \{n \in I,\ x \in \mathcal{D} := \mathcal{D}(F_n) \cap \mathcal{D}(\Omega) : F(n,x) = \sigma\},$$

and

$$\lim_{m\to\infty} \|F(n_m,x_m) - \sigma_m\| = 0, \tag{4.5}$$

the sequence $\{(n_m,x_m)\}$ *has a* (τ_d,τ_w)*-convergent subsequence, where each limit is a* logg-Ω*-minimizing solution* $(n^+,x^+) \in \mathcal{D}(F)$ *of* (2.8)*. Additionally, if the* logg-Ω*-minimizing solution* (n^+,x^+) *is unique, we obtain the* (τ_d,τ_w)*-convergence*

$$(n_m,x_m) \to (n^+,x^+) \quad \textit{in } \mathbb{N} \times \ell^2(\mathbb{N})\,.$$

Proof. As in the proof of Theorem 4.2, we can conclude for the sequence $\{n_m\}$ that it has a convergent subsequence $\{n_{m_k}\}$ with $n_{m_k} \to \tilde{n}$. That is, there exists a $k_0 \in \mathbb{N}$ such that for all $k \geq k_0$ we have $n_{m_k} = \tilde{n}$. From (4.5) it follows from $\delta_m \to 0$ for $m \to +\infty$ that

$$\lim_{m\to\infty} \|F(n_m,x_m) - \sigma\| \leq \lim_{m\to\infty} \|F(n_m,x_m) - \sigma_m\| + \|\sigma_m - \sigma\| = 0. \tag{4.6}$$

Additionally

$$\|F(n_m,x_m)\| \leq \|\sigma_m\| + \|F(n_m,x_m) - \sigma_m\|$$

holds for all $m \in \mathbb{N}$. Since $\lim_{m\to\infty} \sigma_m = \sigma$ and $\lim_{m\to\infty}\|F(n_m,x_m) - \sigma_m\| = 0$ according to (4.5), both $\|\sigma_m\|$ and $\|F(n_m,x_m) - \sigma_m\|$ are bounded. Thus $\{F(n_m,x_m)\}$ is bounded and there exists a weakly convergent subsequence $\{F(n_{m_k},x_{m_k})\}$ in $L^2([0,T])$. So we can choose a subsequence $\{(n_{m_k},x_{m_k})\}$ such that $n_{m_k} = \tilde{n}$ is constant for all $k \in \mathbb{N}$. yielding

$$\{F(n_{m_k},x_{m_k})\} = \{F(\tilde{n},x_{m_k})\}.$$

By [48, Theorem 4.3] there exists a subsequence $\{x_{m_k}\}$ with $x_{m_k} \rightharpoonup \tilde{x}$, and (4.6) leads to

$$F(\tilde{n},\tilde{x}) = \sigma.$$

From condition (4.4) we deduce that

$$\begin{aligned}\operatorname{logg}(\tilde{n}) + \Omega(\tilde{x}) &\leq \liminf_{k\to\infty}(\operatorname{logg}(n_{m_k}) + \Omega(x_{m_k})) \leq \limsup_{k\to\infty}(\operatorname{logg}(n_{m_k}) + \Omega(x_{m_k}))\\ &\leq \operatorname{logg}(n^+) + \Omega(x^+) \leq \operatorname{logg}(n_0) + \Omega(x_0)\end{aligned}$$

for all $(n_0,x_0) \in \mathcal{D}(F)$ with $F(n_0,x_0) = \sigma$. If we set $(n_0,x_0) = (\tilde{n},\tilde{x})$ this leads to

$$\operatorname{logg}(\tilde{n}) + \Omega(\tilde{x}) = \operatorname{logg}(n^+) + \Omega(x^+)$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

and thus $(\tilde{n}, \tilde{x})$ is a logg-Ω-minimizing solution. Moreover,

$$\lim_{m\to\infty} (\operatorname{logg}(n_m) + \Omega(x_m)) = \operatorname{logg}(n^+) + \Omega(x^+).$$

If (n^+, x^+) is unique, then we have (τ_d, τ_w)-convergence to (n^+, x^+) for every subsequence of (n_m, x_m) and, thus, (τ_d, τ_w)-convergence

$$(n_m, x_m) \rightharpoonup (n^+, x^+),$$

which completes the proof. ∎

The following corollary states the regularization property under an a priori parameter choice rule:

COROLLARY 4.4 (Convergence under a priori parameter choice). *For an a priori parameter choice $\alpha_m := \alpha(\delta_m)$ based on a function $\alpha(\delta)$ satisfying*

$$\alpha(\delta) \to 0 \quad \textit{and} \quad \frac{\delta^2}{\alpha(\delta)} \to 0 \quad \textit{as } \delta \to 0$$

and associated regularized solutions $(n_m := n_{\alpha_m}^{\delta_m},\ x_m := x_{\alpha_m}^{\delta_m})$, Theorem 4.3 applies and yields (τ_d, τ_w)-convergent subsequences, where each limit is a logg-Ω*-minimizing solution $(n^+, x^+) \in \mathcal{D}(F)$ of (2.8).*

Proof. The proof follows the lines in [48, Corollary 4.6] and is adapted to the current setting. From the definition of (n_m, x_m) as a minimizer of (3.4), it follows that

$$\tfrac{1}{2}\|F(n_m, x_m) - \sigma_m\|^2 + \alpha_m[\operatorname{logg}(n_m) + \Omega(x_m)] \le \tfrac{1}{2}\delta_m^2 + \alpha_m[\operatorname{logg}(n^+) + \Omega(x^+)].$$

This implies

$$\operatorname{logg}(n_m) + \Omega(x_m) \le \frac{\delta_m^2}{2\alpha_m} + \alpha_m[\operatorname{logg}(n^+) + \Omega(x^+)]$$

and, with $\delta_m^2/\alpha_m \to 0$ as $m \to \infty$, that (4.4) is fulfilled. On the other hand, we deduce that

$$\lim_{m\to\infty} \|F(n_m, x_m) - \sigma_m\| \le (\delta_m^2 + 2\alpha_m(\operatorname{logg}(n^+) + \Omega(x^+)))^{1/2} \to 0$$

for $m \to \infty$. This yields (4.5) and Theorem 4.3 applies. ∎

Applying Morozov's discrepancy principle, we also obtain convergence under an a posteriori parameter choice.

COROLLARY 4.5 (Convergence under a posteriori parameter choice). *For an a posteriori parameter choice $\alpha_m := \alpha(\sigma_m, \delta_m)$ and associated regularized solutions $(n_m := n_{\alpha_m}^{\delta_m},\ x_m := x_{\alpha_m}^{\delta_m})$ satisfying the discrepancy principle*

$$\tau_1\delta_m \le \|F(n_m, x_m) - \sigma_m\| \le \tau_2\delta_m$$

for some $1 \le \tau_1 < \tau_2$, Theorem 4.3 applies and yields (τ_d, τ_w)-convergent subsequences, where each limit is a logg-Ω*-minimizing solution $(n^+, x^+) \in \mathcal{D}(F)$ of (2.8).*

Proof. Following the lines in the proof of Corollary 4.4 we estimate

$$\begin{aligned}\operatorname{logg}(n_m) + \Omega(x_m) &\le \frac{1}{2\alpha_m}(\delta_m^2 - \|F(n_m, x_m) - \sigma_m\|^2 + (\operatorname{logg}(n^+) + \Omega(x^+))) \\ &\le \operatorname{logg}(n^+) + \Omega(x^+),\end{aligned}$$

whence, taking the discrepancy principle into account, (4.4) as well as (4.5) follow. This completes the proof. ∎

ETNA
Kent State University and
Johann Radon Institute (RICAM)

5. Numerical validation. In this section, we validate Algorithm 3.1 by checking its performance and comparing it with the cluster algorithm published in [45] and a simple least-squares method. Note that the experiments of Sections 5.2–5.5 are performed without a penalty term in $T_{\alpha,n}$, i.e., we choose $\Omega = 0$ in (3.2) such that $R(n)$ just solves a nonlinear least-squares problem. The application of penalties $\Omega \neq 0$ is the subject of Section 5.6. The choice of regularization parameters α was done by trial and error. First we briefly recall and sketch the cluster algorithm.

5.1. The cluster algorithm. Note that the forward operator of the underlying inverse problem is not injective, that is, there exist parameters (n_1, x_1) and (n_2, x_2) with $(n_1, x_1) \neq (n_2, x_2)$ and

$$F(n_1, x_1) = F(n_2, x_2).$$

We refer to [45, Example 1] regarding an example of such pairs (n_1, x_1) and (n_2, x_2).

We overcome this nonuniqueness by the following requirement. Assume that the relaxation times are located in different decades (cf. [16]). For example, if $\tau_l \in [10, 100]$ for some $l \in \{1, \dots, n\}$, then $\tau_j \notin [10, 100]$ must hold true for all $j \in \{1, \dots, l-1, l+1, \dots, n\}$. This information is essential for the development of the cluster algorithm.

Let us fix a maximum number of Maxwell elements $N \in \mathbb{N}$ such that the unknown number of Maxwell elements n^* satisfies $n^* \leq N$. Thus, the set of possible numbers of Maxwell elements is represented by $I := \{1, \dots, N\}$. The material parameters $x = \big(\mu, \mu_1, \tau_1, \dots, \mu_N, \tau_N, 0, 0, \dots\big)$ are then computed by a minimization algorithm that works in the same way as the minimization process $R : I \subset \mathbb{N} \to \mathcal{D}(F_n) \subset \ell^2(\mathbb{N})$, $n \mapsto x$, described in Section 3.2, with the difference that the number of Maxwell elements is now known and given by N. In this way we obtain a minimizer x^*.

After the minimization process we apply an algorithm to

$$x^* = (\mu^*, \mu_1^*, \tau_1^*, \dots, \mu_N^*, \tau_N^*, 0, 0, \dots)$$

that clusters the relaxation times $(\tau_1^*, \dots, \tau_N^*)$ according to the decade condition. The cluster algorithm consists of two parts:

1. minimization with N Maxwell elements, output: x^*, and
2. clustering of relaxation times $(\tau_1^*, \dots, \tau_N^*)$ which reduces N Maxwell elements to $n^* \leq N$ elements.

Let us consider the clustering step 2 in more detail. For a decade $[10^k, 10^{k+1}]$ with some $k \in \mathbb{N}_0$, we arrange the Maxwell elements by defining index sets

$$J_k := \{j \in \{1, \dots, N\} : \tau_j \in [10^k, 10^{k+1}]\},$$

where $k \in \mathbb{N}_0$ denotes the decade in which the relaxation time is classified. For the choice of decades, it is useful to take into account the physical conditions of the experiment and the material under consideration. The maximum number of Maxwell elements N should not be larger than the number of available decades. Accordingly, it makes sense to set N equal to the number of decades.

As the next step, the pairs (μ_j, τ_j), $j \in J_k$, must be assigned to a Maxwell element. Collecting the relaxation times and thus the Maxwell elements in index sets, the number of nonempty index sets yields the current number n of elements in the material. To cut down multiple Maxwell elements to a single one, we use the following approximate calculation for the new material parameters $(\tilde{\mu}_k, \tilde{\tau}_k)$ with $\tilde{\tau}_k \in [10^k, 10^{k+1}]$ from the already reconstructed

ETNA
Kent State University and
Johann Radon Institute (RICAM)

material parameters:

$$\tilde{\mu}_k := \sum_{j \in J_k} \mu_j^*, \quad \tilde{\tau}_k := \sum_{j \in J_k} \frac{\mu_j^*}{\tilde{\mu}_k} \tau_j^*. \tag{5.1}$$

In this way, we get the following algorithm:

ALGORITHM 5.1.
Input: N, $\sigma \in L^2([0,T])$
1) *Compute* x^* *by the minimization procedure* $R(N) = x^*$.
2) *Determine the index sets* $J_k \neq \emptyset$ *for* $k \in \mathbb{N}_0$ *with* $J_k := \{j \in \{1,\dots,N\} : \tau_j^* \in [10^k, 10^{k+1}]\}$ *and set* $\tilde{n}$ *as the number of nonempty index sets.*
3) *For each* $k \in \mathbb{N}_0$ *with* $J_k \neq \emptyset$ *compute*

$$\tilde{\mu}_k := \sum_{j \in J_k} \mu_j^*, \qquad \tilde{\tau}_k := \sum_{j \in J_k} \frac{\mu_j^*}{\tilde{\mu}_k} \tau_j^*$$

and set $\tilde{x} := (\mu^*, \tilde{\mu}_1, \tilde{\tau}_1, \dots, \tilde{\mu}_{\tilde{n}}, \tilde{\tau}_{\tilde{n}}, 0, 0, \dots)$.
Output: $(\tilde{n}, \tilde{x})$

In the cluster algorithm, the set of possible numbers of Maxwell elements is fixed by the maximum number N and is not a free parameter to be computed additionally and allowing a specific choice adapted to the given experiment. Furthermore, the cluster algorithm requires the classification of relaxation times in decades. This assumption is not necessary for the Bayes algorithm. On the other hand, the Bayes algorithm allows a simple control of the preferred number of Maxwell elements via the success probability q in the prior. Thus, the Bayes algorithm has a more significant regularizing effect, since its penalty term is physically meaningful and derived by statistical inversion theory. This will be evident in the following reconstruction results that use perturbed data.

5.2. Reconstructions from exact data. We generate simulated data for given material parameters that serve as the basis of our experiments. In Figure 5.1 strains for two different displacement rates $\dot{\varepsilon}_u$ and the maximum strain $\bar{\varepsilon} = 20\%$ are depicted. The fast displacement rate is 10 mm/s, with the maximum strain reached after 2 seconds. At the slow displacement rate of 1 mm/s, the strain of 20% is achieved only after 20 seconds. Here, we set $T = 100$ seconds. For the first experiment we choose a material that is characterized by $n^* = 3$ Maxwell elements and parameters x^* as given in Table 5.1.

TABLE 5.1
Exact material parameters x^* *in experiment 1 to simulate data.*

j	0	1	2	3
μ_j^* [MPa]	10	4	7	1
τ_j^* [s]	–	0.2	3.7	25

The stress $\sigma(t)$ can be computed exactly using (2.6) and is illustrated in Figure 5.1. The goal is to reconstruct the material parameters from Table 5.1 from given data $\sigma(t)$. All tables present the stiffnesses in megapascals [MPa] and the relaxation times in seconds [s].

First, we consider the reconstruction corresponding to the displacement rate of $\dot{\varepsilon}_u = 10$ mm/s. The possible number of Maxwell elements is determined by $I := \{1,\dots,5\}$. Furthermore, we assume, for the relaxation times, $\tau_j \geq \gamma := 0.01$. Table 5.2 shows the

ETNA
Kent State University and
Johann Radon Institute (RICAM)

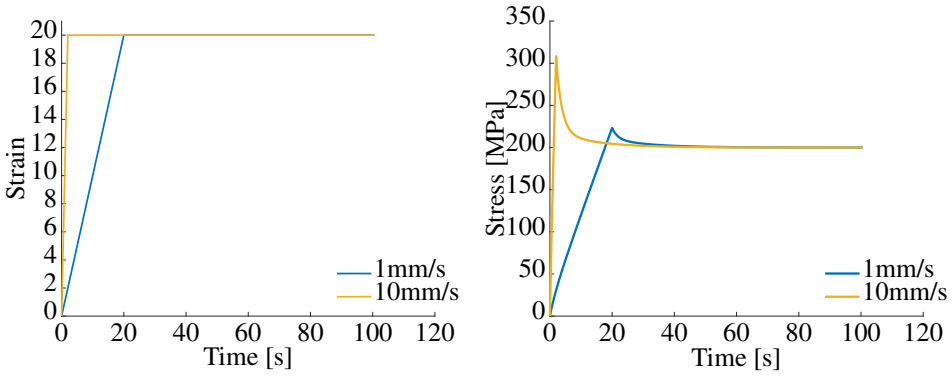

FIG. 5.1. *Strain and stress data for different displacement rates* $\dot{\varepsilon}_u$.

results of the cluster algorithm. While the minimization algorithm determines stiffnesses μ_j and relaxation times τ_j for the given maximum number of Maxwell elements (here $N = 5$), the cluster algorithm is able to combine them accordingly and results in the correct number $n^* = 3$.

TABLE 5.2
Reconstructed material parameters before and after clustering.

j	0	1	2	3	4	5
Reconstructed values						
μ_j	10.000	4.000	3.685	1.621	1.694	1.000
τ_j	–	0.200	3.695	3.706	3.706	25.000
After clustering						
μ_j	10.000	4.000	7.000	1.000		
τ_j	–	0.200	3.700	25.000		

While there is only one Maxwell element with $\tau \in [0, 1)$ and $\tau \in [10, 100)$ after the minimization process, we have three elements in the decade $[1, 10)$. The cluster algorithm then combines these elements. Note that the parameters of Table 5.1 are reconstructed exactly. The relaxation time $\tau_2 = 3.695$ is closest to the correct value of 3.7 before clustering and $\mu_2 \gg \mu_4 > \mu_3$ holds. The weighting (5.1) then results in these excellent results.

We compare the results of the cluster algorithm with the developed Bayes approach. As mentioned, an advantage of the Bayes algorithm is that it does not require a priori information about the different decades of relaxation times. We choose $q = 0.1$ and $M := \max I = 5$. Then, similar to the cluster algorithm, we can reconstruct the parameters exactly. The results are listed in Table 5.3.

TABLE 5.3
Reconstructed material parameters by Algorithm 3.1.

j	0	1	2	3
μ_j	10.000	4.000	7.000	1.000
τ_j	–	0.200	3.700	25.000

ETNA
Kent State University and
Johann Radon Institute (RICAM)

5.3. Reconstructions from noisy data. We continue by performing reconstructions from noise-contaminated data. We add normally distributed noise to the discretized stress σ such that $\|\sigma - \sigma^\delta\| < \delta$ holds with a noise level $\delta > 0$. Furthermore, by

$$\delta_{\text{rel}} := \frac{\|\sigma - \sigma^\delta\|}{\|\sigma^\delta\|},$$

we denote the relative data error. In the following experiments, the noise level is such that $\delta_{\text{rel}} \approx 1\%$ holds.

In Figure 5.2 both strain and stress are plotted. The red line represents the strain curve with a displacement rate of $\dot{\varepsilon}_u = 10$ mm/s and a maximum strain of $\bar{\varepsilon} = 20\%$. The black line is the associated stress with the material parameters from Table 5.1. For reconstruction, we will use only the noisy stress, which is represented by the blue line.

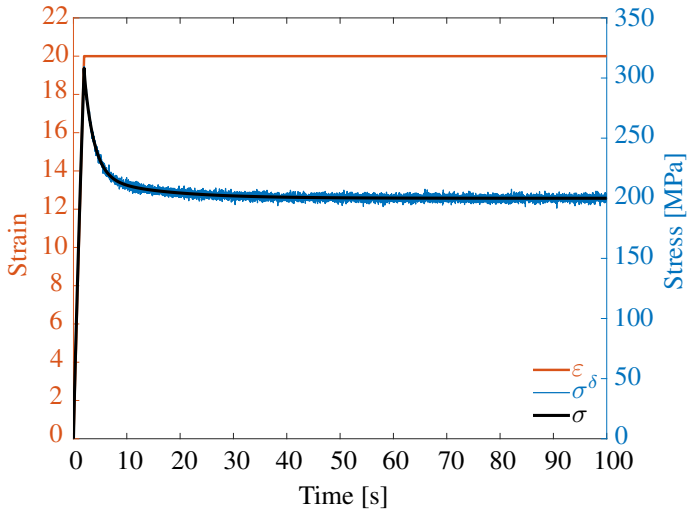

FIG. 5.2. *Strain and stress versus time curves with and without noise for a displacement rate of $\dot{\varepsilon}_u = 10$ mm/s.*

In Table 5.4 we see the output of the clustering algorithm. Here, first, the functional

$$T_N(x) := \|\sigma^\delta - F_N(x)\|^2$$

is minimized with $N = 5$. We realize that the reconstruction results differ significantly from the exact values. However, it is obvious that the reconstruction of (μ_1, τ_1) is the one that is most strongly affected by noise, while the other parameters are calculated quite stably. This confirms the considerations of the authors in [9], where it was shown that the reconstruction of small relaxation times is extremely ill-conditioned and this behavior worsens as $\tau \to 0$. Since stiffness and relaxation time are always to be considered in pairs, the deviations in τ_1 also affect the corresponding stiffness μ_1. Since each Maxwell element provides a certain contribution to the total stress, too small values of one parameter are compensated by higher values of the other parameters to match the total stress.

Again, we compare the reconstructions with those from the Bayes method. The results are presented in Table 5.5. Here, in (3.4) we choose $q = 0.1$ and $\alpha = 1$. We can see that the results are significantly better than the outcome of the cluster algorithm. The basic stiffness and parameters of the second and third Maxwell elements are reconstructed equally well.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

TABLE 5.4
Reconstructed parameters by the cluster algorithm for the data from Figure 5.2.

j	0	1	2	3
μ_j	9.997	57.376	7.028	0.995
τ_j	–	0.013	3.703	25.356

However, here we have good approximations of the exact parameters $(4, 0.2)$ by the values $(4.4238, 0.1617)$ for the first Maxwell element as well.

TABLE 5.5
Reconstructed parameters by Algorithm 3.1 for data from Figure 5.2.

j	0	1	2	3
μ_j	9.9997	4.424	7.046	0.998
τ_j	–	0.162	3.693	25.307

The question arises whether this result can also be achieved by simply minimizing the residual with different numbers n. To check this, we implement a third reconstruction method. We set $I = \{n_1, \dots, n_m\}$ and minimize subsequently the residuals

$$\|\sigma^\delta - F(n_j, x)\|^2, \quad j = 1, \dots, m, \tag{5.2}$$

to determine the parameters x. Finally, we choose (n, x) with the minimal residual. Since the penalty term, given by the prior, is omitted, the entire statistical inversion aspect is neglected.

As we can see in Table 5.6, this reconstruction is very inaccurate. We obtain the maximum possible number of Maxwell elements, however, with $\tau_1 = \tau_2$. This shows that, without additional a priori information on different decades for τ_j or statistical prior, a reliable reconstruction seems impossible.

TABLE 5.6
Reconstructed parameters by minimizing the residual for the data from Figure 5.2.

j	0	1	2	3	4	5
μ_j	9.9997	4.366	0.100	0.876	7.030	0.996
τ_j	–	0.100	0.100	0.339	3.701	25.347

We want to perform another experiment applying the three methods. To this end, we use the same setting with a maximum strain of $\bar{\varepsilon} = 20\%$, a displacement rate of $\dot{\varepsilon}_u = 10$ mm/s, and a relative noise level of $\delta_{\text{rel}} \approx 1\%$. However, we change the material parameters, as listed in Table 5.7. Here, the material under consideration has only two Maxwell elements.

The results of the different algorithms are shown in Table 5.8. As in the last experiment, the basic stiffness as well as the parameters of the Maxwell element with longer relaxation time are reconstructed by all algorithms with a reasonable accuracy. However, the cluster algorithm as well as the minimization of the residual results in too many Maxwell elements. The cluster algorithm leads to another Maxwell element with very large relaxation time, but very low stiffness, that has little influence on the total stress. However, again, the first Maxwell element $(\mu_1^*, \tau_1^*) = (8, 0.5)$ is reconstructed accurately. Minimizing the residual causes four Maxwell elements where there should be only the first one. Thus, a total of five elements are

ETNA
Kent State University and
Johann Radon Institute (RICAM)

TABLE 5.7
Second set of material parameters x^* to simulate data.

j	0	1	2
μ_j^* [MPa]	5	8	0.5
τ_j^* [s]	–	0.8	50

reconstructed. Only the Bayes algorithm succeeds in producing a suitable reconstruction of the parameters.

TABLE 5.8
Reconstructed parameters by the three algorithms to the exact parameters from Table 5.7.

Cluster algorithm

j	0	1	2	3
μ_j	4.998	29.740	0.502	8.053×10^{-10}
τ_j	–	0.215	50.309	576.988

Bayes algorithm

j	0	1	2
μ_j	4.998	7.923	0.502
τ_j	–	0.502	50.241

Residual minimization (5.2)

j	0	1	2	3	4	5
μ_j	4.998	0.130	0.130	4.161	3.698	0.502
τ_j	–	0.100	0.100	0.811	0.812	50.270

5.4. Analysis of different displacement rates $\dot{\varepsilon}_u$. In the following, we consider different displacement rates and how they affect the outcome of the reconstructions. For this purpose, we return to our first experiment corresponding to the parameters from Table 5.1. The experimental setup remains the same: the sample is stretched to a maximum of $\bar{\varepsilon} = 20\%$, and the additive noise has a relative noise level of $\delta_{\text{rel}} = 1\%$. We already know the result of this experiment for a displacement rate of $\dot{\varepsilon}_u = 10$ mm/s from Table 5.5. Table 5.9 shows the reconstructions for the Bayes algorithm for this experiment with a displacement rate of $\dot{\varepsilon}_u = 1$ mm/s, probability $q = 0.1$, and regularization parameter $\alpha = 1$.

TABLE 5.9
Reconstructed parameters using Algorithm 3.1 with $q = 0.1$ for $\dot{\varepsilon}_u = 1$ mm/s.

j	0	1	2
μ_j	9.997	7.450	1.007
τ_j	–	3.576	25.100

As we can see, the Bayes algorithm cannot reconstruct all Maxwell elements reliably, because the first Maxwell element is missing. This is why we want to check whether the Bayes algorithm is able to find this Maxwell element if we weaken the strong weighting to a small

ETNA
Kent State University and
Johann Radon Institute (RICAM)

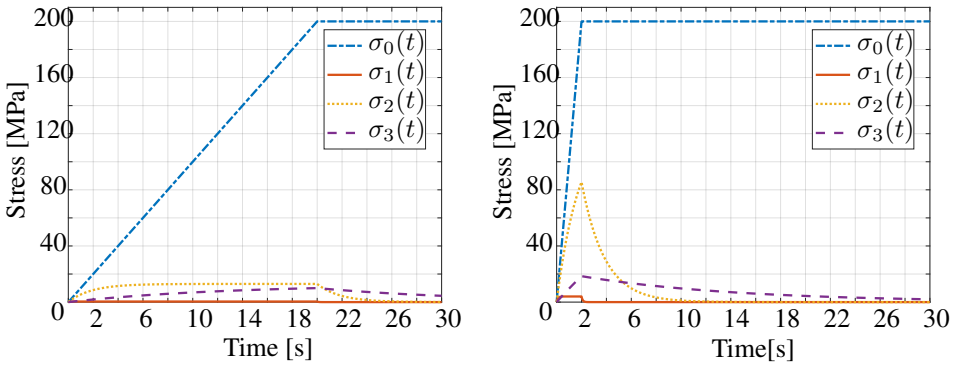

FIG. 5.3. *Individual stress components for a displacement rates of* 1 *mm/s (left) and* 10 *mm/s (right).*

number of Maxwell elements. For this purpose, Table 5.10 shows the reconstruction of the Bayes algorithm with $q = 0.3$.

TABLE 5.10
Reconstructed parameters by Algorithm 3.1 with $q = 0.3$ *for* $\dot{\varepsilon}_u = 1$ *mm/s.*

j	0	1	2	3
μ_j	9.996	7.411	0.299	0.796
τ_j	–	3.500	15.369	27.100

In fact, three Maxwell elements are reconstructed this time, but the first Maxwell element $(\mu_1^*, \tau_1^*) = (4, 0.2)$ is not determined. Moreover, the third Maxwell element $(\mu_3^*, \tau_3^*) = (1, 25)$ is split into two Maxwell elements $(0.2988, 15.3686)$ and $(0.7955, 27.1003)$.

The displacement rate seems somehow to affect the reconstruction outcome. To this end, we consider the different stress–time curves generated by a displacement rate of 1 mm/s and 10 mm/s (Figure 5.3). For each of these curves, the contribution of each element is demonstrated separately. That is, σ_0 is the stress generated by the single spring, while $\sigma_{1,2,3}$ are the stresses of the three Maxwell elements. The sum of these individual stresses yields the total stress (2.6). The maximum strain is 20%, which is achieved at a displacement rate of 1 mm/s and 10 mm/s after 20 seconds and 2 seconds, left and right panels, respectively. This represents a significant difference for the individual stresses of the Maxwell elements. In both cases, the maximum stress of the single spring is 200 MPa, but at different time instants. However, the maximum stress values of the Maxwell elements are much lower at the slower displacement rate of 1 mm/s than at a higher rate. In both cases, the maximum is achieved at $t = \bar{\varepsilon}/\dot{\varepsilon}$.

TABLE 5.11
Maximum values of individual stress components with slow and fast displacement rates at time $t = \bar{\varepsilon}/\dot{\varepsilon}$.

	$j =$	1	2	3
$\dot{\varepsilon}_u = 1$	$\sigma_j(\bar{\varepsilon}/\dot{\varepsilon})$	0.4	12.94	9.9
$\dot{\varepsilon}_u = 10$	$\sigma_j(\bar{\varepsilon}/\dot{\varepsilon})$	4	85.57	18.48

Table 5.11 lists the different values. For a slower displacement rate, the stress values of

ETNA
Kent State University and
Johann Radon Institute (RICAM)

the Maxwell elements are nonzero over a longer period of time, recognizable by the second Maxwell element in Figure 5.3. However, since the values, especially for the first Maxwell element, are much smaller than at a higher displacement rate, these small values are much more affected by noise. Therefore, it is advisable to use higher shift rates. For this reason, we confine ourselves to a displacement rate of 10 mm/s in the next examples.

5.5. Effect of q on the reconstruction results. As already deduced from Table 5.10, increasing q causes the algorithm to favor a higher number of Maxwell elements, as intended by the prior (cf. Section 3.1). In this section we investigate the sensitivity of the results regarding q.

We consider again the second experiment as already described in Section 5.3, with the material parameters as listed in Table 5.7. We already know that the Bayes algorithm for $q = 0.1$ reconstructs the parameters very well. We now change the value of q to $q = 0.9$. The prior $\rho_0(n)$ thus favors a higher number of Maxwell elements, that is, the penalty term $\phi_0(n)$ is monotonically decreasing and attains its maximum for $n = 1$.

TABLE 5.12
Reconstructed parameters of the second experiment by Algorithm 3.1 with $q = 0.9$.

j	0	1	2	3	4	5
μ_j	4.998	0.130	0.130	4.161	3.698	0.502
τ_j	–	0.100	0.100	0.811	0.812	50.270

Table 5.12 shows the reconstructed parameters for this experiment. As we can see, the Bayes algorithm reconstructs five elements instead of two. This is consistent with the prior. A test series using different values of success probability q increasing with step size 0.1 reveals that only for $q = 0.4$ is the correct number of Maxwell elements reconstructed; see Table 5.13. Further tests prove that for $0 < q \leq 0.4$ the material parameters are reliably reconstructed. This includes the parameters that we obtained for $q = 0.1$ (cf. Table 5.8).

TABLE 5.13
Reconstructed parameters of the second experiment by Algorithm 3.1 with $q = 0.4$.

j	0	1	2
μ_j	4.998	7.923	0.502
τ_j	–	0.807	50.241

A third trial uses the same setting with $\bar{\varepsilon} = 20\%$, $\dot{\varepsilon}_u = 10$ mm/s, and $\delta_{\text{rel}} \approx 1\%$. The parameters are modified according to Table 5.14. The experiment is extended to $T = 10\,000$ seconds, since the largest relaxation time is $\tau_5 = 1200$ s. It is worth mentioning that in the previous experiments we used a temporal sampling rate of $\Delta t = 0.01$ seconds. That is, $t_i = i\,\Delta t$ for $i = 0, \ldots, m$. In the third experiment, we choose the same $\Delta t = 0.1$ s to obtain a sampling rate that is small enough to catch the influence of the first Maxwell element having a relaxation time of $\tau_1 = 0.8$ s. Moreover, we choose a probability of $q = 0.1$, which favors a small number of elements. Table 5.15 presents the reconstruction result.

Although the algorithm prefers small numbers of Maxwell elements, all elements and the corresponding parameters are reconstructed sufficiently accurately. This results suggests the use of a small q, like, e.g., $q = 0.1$, if the number of Maxwell elements is unknown, since the algorithm is able to deliver the correct number as output.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

TABLE 5.14
Exact parameters x^* for the third experiment.

j	0	1	2	3	4	5
μ_j^*	10	8	7	1	4	0.5
τ_j^*	–	0.8	3.7	25	500	1200

TABLE 5.15
Reconstructed parameters of the third experiment by the Bayes algorithm with $q = 0.1$.

j	0	1	2	3	4	5
μ_j	10.000	8.427	7.011	0.893	4.077	0.419
τ_j	–	0.763	3.846	27.47	505.8	1282.5

5.6. Regularization by an additional penalty term. As outlined in Section 3, penalty terms are necessary for a stable reconstruction of the parameters. Regularization was also applied to the clustering algorithm in the presence of noise in the data; cf. [45]. The numerical results emphasize that the penalty term leads to more accurate results than the cluster algorithm without penalty. The only data sets that are not recovered accurately are those with a displacement rate of $\dot{\varepsilon}_u = 1$ mm/s. Although this can be avoided by a corresponding experimental setup, we investigate whether an additional penalty term leads to better reconstruction results. As penalty term, we use

$$\Omega_1(x) = \tfrac{1}{2}\|x\|^2. \tag{5.3}$$

For a fixed number of Maxwell elements n, this leads to a classical Tikhonov–Phillips regularization of the form

$$\min_{x\in\mathcal{D}(F_n)} T_{\alpha,n}(x) := \min_{x\in\mathcal{D}(F_n)} \{\tfrac{1}{2}\|F_n(x) - \sigma^\delta\|^2 + \alpha\Omega_1(x)\};$$

see, e.g., [13, 27, 36, 47]. The parameter $\alpha > 0$ acts as a regularization parameter and balances the influence of the data term and the penalty term on the minimizer.

Table 5.16 shows the result of the Bayes algorithm with penalty (5.3) and regularization parameter $\alpha = 0.5$. We see that the algorithm now reconstructs three Maxwell elements. However, the values of the first and second Maxwell elements are identical, and the correct first Maxwell element $(\mu_1^*, \tau_1^*) = (4, 0.2)$ is not recovered at all. As expected, the penalty term here ensures that large values are penalized. This also explains the reduction of the stiffness value μ_3. As a remedy, we apply a further penalty term of the form

$$\Omega_2(x) = \tfrac{1}{2}\tau_1^2 \tag{5.4}$$

to penalize large values τ_1. The regularization parameter is increased to $\alpha = 100$.

TABLE 5.16
Reconstructed parameters of the first experiment for a displacement rate $\dot{\varepsilon}_u = 1$ mm/s by the Bayes algorithm with additional penalty term (5.3) and $\alpha = 0.5$.

j	0	1	2	3
μ_j	10.000	3.765	3.765	1.117
τ_j	–	3.376	3.376	23.238

ETNA
Kent State University and
Johann Radon Institute (RICAM)

Table 5.17 shows the corresponding outcome of this experiment. We see a significant improvement compared to the results of Tables 5.9 and 5.16. All three Maxwell elements are reconstructed accurately, and the third Maxwell element is no longer negatively affected by Ω_2, though it does only take τ_1 into account. We admit that the first Maxwell element still shows some error sensitivity, even though the reconstruction means a significant improvement.

TABLE 5.17
Reconstructed parameters of the first experiment for a displacement rate $\dot{\varepsilon}_u = 1$ mm/s by the Bayes algorithm with additional penalty term (5.4) and $\alpha = 100$.

j	0	1	2	3
μ_j	9.997	2.015	7.327	1.001
τ_j	–	0.100	3.621	25.190

We conclude that the recommendation to use a higher displacement rate, such as $\dot{\varepsilon}_u = 10$ mm/s, remains valid even with an additional penalty term.

6. Conclusion. In this paper, we have considered the inverse problem of identifying material parameters in a viscoelastic structure using a generalized Maxwell model. One major challenge is the fact that the number of Maxwell elements in this model, and thus the number of material parameters, is unknown. Based on statistical inversion using a binomial prior, we developed a novel reconstruction method which is able to compute the number of elements along with the corresponding material parameters in a stable way.

Since the forward operator acts on $\mathbb{N} \times \ell^2(\mathbb{N})$, it was as a further novelty necessary to extend the existing regularization theory to semigroups, where we equipped $\mathbb{N}$ with the discrete topology.

The influence of the success probability q has been studied and the method is compared to the cluster algorithm through the use of extensive numerical tests. While the cluster algorithm is very sensitive with respect to noise in the data, the Bayes algorithm proves to be stable, where in case of low displacement rates additional penalty terms improve the reconstruction results. Using statistical inversion theory yields a penalty term that is tailored to the problem and does not demand any a priori information on relaxation times and number of Maxwell elements.

Controlling the success probability q allows the reconstruction to be adapted according to any prior information on different structures. In general, it is advisable to choose a small value of q in order to favor a small number of Maxwell elements, whereas larger n can also be reconstructed reliably. The numerical evaluation proved the stability of the new method for different parameter settings as well as its superiority with respect to the cluster algorithm.

Acknowledgements. The authors would like to thank the anonymous reviewers for their valuable comments that helped to significantly improve the original manuscript.

REFERENCES

[1] B. BABAEI, A. DAVARIAN, K. M. PRYSE, E. L. ELSON, AND G. M. GENIN, *Efficient and optimized identification of generalized Maxwell viscoelastic relaxation spectra*, J. Mech. Behavior Biomed. Mat., 55 (2016), pp. 32–41.

[2] F. BINDER, F. SCHÖPFER, AND T. SCHUSTER, *Defect localization in fibre-reinforced composites by computing external volume forces from surface sensor measurements*, Inverse Problems, 31 (2015), Art. No. 025006, 12 pages.

[3] A. BONFANTI, J. L. KAPLAN, G. CHARRAS, AND A. KABLA, *Fractional viscoelastic models for power-law materials*, Soft Matter, 16 (2020), pp. 6002–6020.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

[4] L. BORCEA, *Electrical impedance tomography*, Inverse Problems, 18 (2002), pp. R99–R136.
[5] T. F. COLEMAN AND Y. LI, *On the convergence of interior-reflective Newton methods for nonlinear minimization subject to bounds*, Math. Programming 67 (1994), pp. 189–224.
[6] ———, *An interior trust region approach for nonlinear minimization subject to bounds*, SIAM J. Optim., 6 (1996), pp. 418–445.
[7] D. L. COLTON AND R. KRESS, *Inverse Acoustic and Electromagnetic Scattering Theory*, 2nd ed., Springer, Berlin, 1998.
[8] L. R. CROFT, P. W. GOODWILL, AND S. M. CONOLLY, *Relaxation in X-space magnetic particle imaging*, IEEE Trans. Med. Imag., 31 (2012), pp. 2335–2342.
[9] S. DIEBELS, T. SCHEFFER, T. SCHUSTER, AND A. WEWIOR, *Identifying elastic and viscoelastic material parameters by means of a Tikhonov regularization*, Math. Probl. Eng., 2018 (2018).
[10] M. ELLER, R. GRIESMAIER, AND A. RIEDER, *Tangential cone condition for the full waveform forward operator in the viscoelastic regime: the nonlocal case*, SIAM J. Appl. Math., 84 (2024), pp. 412–432.
[11] I. EMRI AND N. W. TSCHOEGL, *An iterative computer algorithm for generating line spectra from linear viscoelastic response functions*, Int. J. Polym. Mat. Polym. Biomat., 40 (1998), pp. 55–79.
[12] H. W. ENGL, M. HANKE, AND A. NEUBAUER, *Regularization of Inverse Problems*, Kluwer, Dordrecht, 1996.
[13] H. W. ENGL, K. KUNISCH, AND A. NEUBAUER, *Convergence rates for Tikhonov regularisation of nonlinear ill-posed problems*, Inverse Problems, 5 (1989), pp. 523–540.
[14] M. FERNANDA, P. COSTA, AND C. RIBEIRO, *Parameter estimation of viscoelastic materials: a test case with different optimization strategies*, in Numerical Analysis and Applied Mathematics ICNAAM 2011, 1389, AIP Conference Proceedings, American Institute of Physics, Melville, 2011, pp. 771–774.
[15] V. GIURGIUTIU, *Structural Health Monitoring with Piezoelectric Wafer Active Sensors*, Academic Press, Oxford, 2007.
[16] F. GOLDSCHMIDT AND S. DIEBELS, *Modeling the moisture and temperature dependent material behavior of adhesive bonds*, PAMM, 15 (2015), pp. 295–296.
[17] ———, *Modelling and numerical investigations of the mechanical behavior of polyurethane under the influence of moisture*, Arch. Appl. Mech., 85 (2015), pp. 1035–1042.
[18] M. HANKE, A. NEUBAUER, AND O. SCHERZER, *A convergence analysis of the Landweber iteration for nonlinear ill-posed problems*, Numer. Math., 72 (1995), pp. 21–37.
[19] S. HARTMANN AND P. NEFF, *Polyconvexity of generalized polynomial-type hyperelastic strain energy functions for near-incompressibility*, Internat. J. Solids Structures, 40 (2003), pp. 2767–2791.
[20] B. HOFMANN, B. KALTENBACHER, C. PÖSCHL, AND O. SCHERZER, *A convergence rates result for Tikhonov regularization in Banach spaces with non-smooth operators*, Inverse Problems, 23 (2007), pp. 987–1010.
[21] V. INGLE, S. KOGON, AND D. MANOLAKIS, *Statistical and Adaptive Signal Processing*, Artech, Boston, 2005.
[22] K. ITO AND B. JIN, *Inverse Problems*, World Scientific, Hackensack, 2015.
[23] M. JAKSZTO, *Another proof that L^p-bounded pointwise convergence implies weak convergence*, Real Anal. Exchange, 36 (2010/11), pp. 479–482.
[24] M. JOHLITZ AND A. LION, *Chemo-thermomechanical ageing of elastomers based on multiphase continuum mechanics*, Contin. Mech. Thermodyn., 25 (2013), pp. 605–624.
[25] M. JOHLITZ, H. STEEB, S. DIEBELS, A. CHATZOURIDOU, J. BATAL, AND W. POSSART, *Experimental and theoretical investigation of nonlinear viscoelastic polyurethane systems*, J. Mat. Sci., 42 (2007), pp. 9894–9904.
[26] B. KALTENBACHER, *Some Newton-type methods for the regularization of nonlinear ill-posed problems*, Inverse Problems, 13 (1997), pp. 729–753.
[27] B. KALTENBACHER, A. NEUBAUER, AND O. SCHERZER, *Iterative Regularization Methods for Nonlinear Ill-Posed Problems*, de Gruyter, Berlin, 2008.
[28] B. KALTENBACHER, F. SCHÖPFER, AND T. SCHUSTER, *Iterative methods for nonlinear ill-posed problems in Banach spaces: convergence and applications to parameter identification problems*, Inverse Problems, 25 (2009), Art. No. 065003, 19 pages.
[29] A. KIRSCH, *An Introduction to the Mathematical Theory of Inverse Problems*, Springer, New York, 2011.
[30] R. KLEIN, T. SCHUSTER, AND A. WALD, *Sequential subspace optimization for recovering stored energy functions in hyperelastic materials from time-dependent data*, in Time-Dependent Problems in Imaging and Parameter Identification, B. Kaltenbacher, T. Schuster, and A. Wald, eds., Springer, Cham, 2021, pp. 165–190.
[31] M. V. KLIBANOV AND V. G. ROMANOV, *Reconstruction procedures for two inverse scattering problems without the phase information*, SIAM J. Appl. Math., 76 (2016), pp. 178–196.
[32] A. K. LOUIS, *Approximate inverse for linear and some nonlinear problems*, Inverse Problems, 11 (1995), pp. 1211–1223. Corrigendum in Inverse Problems, 12 (1996), pp. 175-190.
[33] ———, *Inverse und Schlecht Gestellte Probleme*, B. G. Teubner, Stuttgart, 1989.
[34] MATHWORKS, *MATLAB - Mathworks - MATLAB & Simulink*, www.mathworks.com, 2016.
[35] R. E. MEGGINSON, *An Introduction to Banach Space Theory*, Springer, New York, 1998.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

[36] A. NEUBAUER, *Tikhonov regularization of nonlinear ill-posed problems in Hilbert scales*, Appl. Anal., 46 (1992), pp. 59–72.
[37] S. PARK AND R. SCHAPERY, *Methods of interconversion between linear viscoelastic material functions. Part I—A numerical method based on Prony series*, Int. J. Solids. Struct., 36 (1999), pp. 1653–1675.
[38] G. PLONKA AND M. TASCHE, *Prony methods for recovery of structured functions*, GAMM-Mitt., 37 (2014), pp. 239–258.
[39] S. REESE AND S. GOVINDJEE, *A theory of finite viscoelasticity and numerical aspects*, Int. J. Solids. Struct., 35 (1998), pp. 3455–3482.
[40] A. RIEDER, *Keine Probleme mit Inversen Problemen*, Vieweg, Braunschweig, 2003.
[41] ———, *An all-at-once approach to full waveform inversion in the viscoelastic regime*, Math. Methods Appl. Sci., 44 (2021), pp. 6376–6388.
[42] B. ROOT, *High-resolution array with Prony, MUSIC, and ESPRIT algorithms*, Tech. Report, Naval Research Lab, Washington D.C., 1992.
[43] D. ROTHERMEL AND T. SCHUSTER, *Solving an inverse heat convection problem with an implicit forward operator by using a projected quasi-Newton method*, Inverse Problems, 37 (2021), Art. No. 045014, 36 pages.
[44] D. ROTHERMEL, T. SCHUSTER, R. SCHORR, AND M. PEGLOW, *Determination of the temperature-dependent thermal material properties in the cooling process of steel plates*, Math. Probl. Eng., 2021 (2021), Art. No. 6653388, 13 pages.
[45] R. ROTHERMEL, W. PANFILENKO, P. SHARMA, A. WALD, T. SCHUSTER, A. JUNG, AND S. DIEBELS, *A method for determining the parameters in a rheological model for viscoelastic materials by minimizing Tikhonov functionals*, Appl. Math. Sci. Eng., 30 (2022), pp. 141–165.
[46] T. SCHEFFER, H. SEIBERT, AND S. DIEBELS, *Optimisation of a pretreatment method to reach the basic elasticity of filled rubber materials*, Arch. Appl. Mech., 83 (2013), pp. 1659–1678.
[47] O. SCHERZER, *The use of Morozov's discrepancy principle for Tikhonov regularization for solving nonlinear ill-posed problems*, Computing, 51 (1993), pp. 45–60.
[48] T. SCHUSTER, B. KALTENBACHER, B. HOFMANN, AND K. S. KAZIMIERSKI, *Regularization Methods in Banach Spaces*, de Gruyter, Berlin, 2012.
[49] J. SEYDEL AND T. SCHUSTER, *Identifying the stored energy of a hyperelastic structure by using an attenuated Landweber method*, Inverse Problems, 33 (2017), Art. No. 124004, 31 pages.
[50] ———, *On the linearization of identifying the stored energy function of a hyperelastic material from full knowledge of the displacement field*, Math. Methods Appl. Sci., 40 (2017), pp. 183–204.
[51] P. SHARMA, A. SAMBALE, M. STOMMEL, M. MAISL, H. G. HERRMANN, AND S. DIEBELS, *Moisture transport in PA6 and its influence on the mechanical properties*, Contin. Mech. Thermodyn., 32 (2020), pp. 307–325.
[52] K. SHUKLA, J. CHAN, AND M. V. DE HOOP, *A high order discontinuous Galerkin method for the symmetric form of the anisotropic viscoelastic wave equation*, Comput. Math. Appl, 99 (2021), pp. 113–132.
[53] N. W. TSCHOEGL, *The Phenomenological Theory of Linear Viscoelastic Behavior*, Springer, Berlin, 1989.
[54] R. V. MISES, *On Saint Venant's principle*, Bull. Amer. Math. Soc., 51 (1945), pp. 555–562.
[55] A. WALD AND T. SCHUSTER, *Tomographic terahertz imaging using sequential subspace optimization*, in New Trends in Parameter Identification for Mathematical Models, B. Hofmann, A. Leitão, and J. P. Zubelli, eds., Springer, Cham, 2018, pp. 261–290.
[56] A. S. WINEMAN AND K. R. RAJAGOPAL, *Mechanical Response of Polymers*, Cambridge University Press, Cambridge, 2000.
[57] A. WÖSTEHOFF AND T. SCHUSTER, *Uniqueness and stability result for Cauchy's equation of motion for a certain class of hyperelastic materials*, Appl. Anal., 94 (2015), pp. 1561–1593.