

MATHEMATICAL ANALYSIS OF SOME ITERATIVE METHODS FOR THE RECONSTRUCTION OF MEMORY KERNELS*

MARTIN HANKE[†]

Abstract. We analyze three iterative methods that have been proposed in the computational physics community for the reconstruction of memory kernels in a stochastic delay differential equation known as the generalized Langevin equation. These methods use the autocorrelation function of the solution of this equation as input data. Although they have been demonstrated to be useful, a straightforward Laplace analysis does not support their conjectured convergence. We provide more detailed arguments to explain the good performance of these methods in practice. In the second part of this paper we investigate the solution of the generalized Langevin equation with a perturbed memory kernel. We establish sufficient conditions including error bounds such that the stochastic process corresponding to the perturbed problem converges to the unperturbed process in the mean square sense.

Key words. generalized Langevin equation, Laplace transform, strong convergence

AMS subject classifications. 60G15, 65R32, 65C20

1. Introduction. The generalized Langevin equation is a stochastic delay differential equation which serves, for example, as a model for the motion of a macroparticle in a solvent. Let m be the mass and $V = V(t) \in \mathbb{R}^3$ the velocity of the macroparticle at time t . Then the corresponding equation assumes the form

$$(1.1) \quad m \dot{V}(t) = - \int_{-\infty}^t k(t - \tau) V(\tau) d\tau + F_R(t),$$

where k is the so-called memory kernel and F_R is a random force describing the force resulting on the macromolecule due to local interactions in between the molecules of the solvent. This random force is typically taken to be a stationary centered Gaussian process whose autocorrelation function

$$(1.2) \quad \mathbb{E} (F_R(\tau + t) F_R(\tau)^*) = \frac{1}{\beta} k(t), \quad t \geq 0,$$

is connected to the memory kernel in order to satisfy the so-called fluctuation dissipation theorem (cf. Pottier [13]), i.e., to conserve the energy of the total system; accordingly, the scalar $\beta = (k_B T)^{-1}$ depends on the temperature T , and k_B is the Boltzmann constant.

In (1.1) we have formulated the generalized Langevin equation with an infinite horizon in the past to emphasize that we are restricting ourselves to the stationary solution of this problem corresponding to a physical system in thermodynamical equilibrium; see Section 5 for a more detailed discussion of this setting. The stationary solution V of (1.1) is again a centered Gaussian process whose (rescaled) velocity autocorrelation function

$$(1.3) \quad r(t) = \beta m \mathbb{E} (V(\tau + t) V(\tau)^*), \quad t \geq 0,$$

satisfies the deterministic delay differential equation

$$(1.4) \quad m \dot{r}(t) = - \int_0^t k(t - \tau) r(\tau) d\tau, \quad r(0) = I.$$

*Received February 26, 2021. Accepted May 31, 2021. Published online on July 5, 2021. Recommended by Sergei Pereverzyev. The research leading to this work has been done within the Collaborative Research Center TRR 146; corresponding funding by the DFG is gratefully acknowledged.

[†]Institut für Mathematik, Johannes Gutenberg-Universität Mainz, 55099 Mainz, Germany
 (hanke@mathematik.uni-mainz.de).

In computational physics the generalized Langevin equation is employed as a means to avoid the explicit treatment of the solvent molecules in a numerical simulation of the system; using (1.1), they only appear implicitly through the memory kernel and the corresponding random force. Accordingly, to set up this effective model, one first has to determine a suitable memory term. In principle, this can be done analytically using appropriate projection techniques (the Mori-Zwanzig formalism, cf., e.g., Chen, Li, and Liu [1]), but for practical purposes this is usually far too complicated. Instead, one can preliminary run all atom molecular dynamics (MD) simulations—or even use experimental data—to compute the velocity autocorrelation (1.3) and then search for a suitable memory kernel of the effective model (1.1), where the goodness of fit is somehow determined by the error in (1.3). This is the inverse problem that we are dealing with in the following.

The contributions of this paper are as follows. In Section 2 we recapitulate this inverse problem from a mathematical perspective and provide some necessary theoretical background. Then we consider in Section 3 two iterative methods for solving this problem that have recently been introduced in the computational physics community. We provide an analysis of these schemes to explain convergence (and nonconvergence) phenomena that have been observed in these works. A variant of these two methods with somewhat different theoretical properties is briefly considered in Section 4 for the sake of completeness. As a related but independent topic we elaborate in Section 5 on admissible topologies for approximating the memory kernel, such that the resulting stationary processes converge in the mean square sense to the reference process. We conclude with some final remarks in Section 6.

2. The setting of the inverse problem. In the remainder of this paper we consider the generalized Langevin equation (1.1) in d -dimensional space, i.e., $V = V(t) \in \mathbb{R}^d$. Further details about this stochastic integro-differential equation are postponed to Section 5.

Without loss of generality, we restrict ourselves to dimensionless variables, i.e., we let $m = \beta = 1$, and throughout we make the assumption that the memory kernel is a continuous and absolutely integrable matrix-valued function $k : \mathbb{R}_0^+ \rightarrow \mathbb{R}^{d \times d}$. This implies that its Laplace transform $\widehat{k} = \widehat{k}(s)$ is an analytic function of $s \in \mathbb{C}$, $\text{Re } s > 0$, with values in $\mathbb{C}^{d \times d}$, which extends continuously to the imaginary axis $s = i\xi$, $\xi \in \mathbb{R}$. Since k reappears as autocorrelation function of the random force, cf. (1.2), it must be a function of positive type (aka nonnegative definite) by Bochner’s theorem, cf., e.g., Cramér and Leadbetter [2]. This means that the Hermitian matrix

$$(2.1) \quad \widehat{k}(i\xi) + \widehat{k}(i\xi)^*$$

is positive semidefinite for every $\xi \in \mathbb{R}$. In this paper we make the stronger assumption that (2.1) is positive definite for every $\xi \in \mathbb{R}$, in which case the function k is said to be of *strict positive type*, cf. Gripenberg, London, and Staffans [3].

Under these assumptions the delay differential equation (1.4) takes the form

$$(2.2) \quad \dot{r}(t) = - \int_0^t k(t - \tau)r(\tau) \, d\tau, \quad r(0) = I,$$

and it admits a unique solution $r \in C^1(\mathbb{R}^+, \mathbb{R}^{d \times d}) \cap L^1(\mathbb{R}_0^+, \mathbb{R}^{d \times d})$ which satisfies

$$(2.3) \quad \int_0^t k(t - \tau)r(\tau) \, d\tau = \int_0^t r(t - \tau)k(\tau) \, d\tau$$

and depends continuously on k (see [3] and Lemma 5.1 below). This solution is known as the differential resolvent associated with k . The inverse problem of interest in this paper consists in finding the memory kernel k associated with a given r .

This inverse problem can be formulated as a linear problem by using (2.3) to rewrite (2.2) as a Volterra integral equation

$$(2.4) \quad \int_0^t r(t-\tau)k(\tau) \, d\tau = -\dot{r}(t), \quad t > 0,$$

for k . In general this requires (i) the numerical differentiation of the given data to compute the right-hand side of (2.4) and (ii) the subsequent solution of a first-kind Volterra integral equation; since $r(0) = I$, the Volterra integral operator in (2.4) is 1-smoothing as defined in Lamm [7], and hence, solving (2.4) is a moderately ill-posed problem. However, the integral operator is also known approximately only because of the given uncertainties in the measured differential resolvent.

Instead of solving the first-kind integral equation (2.4) one can differentiate this equation to obtain the second-kind integral equation

$$(2.5) \quad k(t) + \int_0^t \dot{r}(t-\tau)k(\tau) \, d\tau = -\ddot{r}(t), \quad t > 0.$$

The solution of this equation is well-posed, but first- and second-order derivatives of the data are needed to set up the system. In the physical application mentioned above, this is no major problem when the data are given by MD simulations, because the second derivative of r is the autocorrelation function of the total force acting on the macroparticle, i.e., the right-hand side of (1.1), while

$$\dot{r}(t) = \frac{d}{dt} \mathbb{E} (V(\tau+t)V(\tau)^*) = \mathbb{E} (\dot{V}(\tau+t)V(\tau)^*), \quad t \geq 0,$$

is given by the correlation between the total force and the velocity; both are, in principle, available in MD simulations. We refer to Linz [9] for an overview of numerical algorithms for solving these two integral equations numerically.

Alternatively, one can formulate the inverse problem as a nonlinear inverse problem by defining the forward operator $A : k \mapsto r$ that maps any absolutely integrable memory kernel of strict positive type onto its associated differential resolvent, i.e., the solution of the delay differential equation (2.2). Then one can apply various iterative methods for solving the nonlinear problem

$$A(k) = r.$$

Typically, each iterative step of these methods requires an application of the forward map A , i.e., a solution of (2.2) for a preliminary approximation of k , which is a well-posed task provided that k satisfies the above assumptions. An example of such an iterative scheme is the IMRV iteration introduced in [5], which will be investigated in Section 3.

Finally, we mention that the memory kernel is uniquely determined by the data r . For if two continuous memory kernels k_1 and k_2 give rise to the same differential resolvent, then (2.4) implies that

$$\int_0^t r(t-\tau)(k_1 - k_2)(\tau) \, d\tau = 0, \quad 0 \leq t < \infty,$$

and it follows, e.g., from [9, Theorem 5.1] that the only continuous solution of this integral equation is the zero solution, hence $k_1 = k_2$.

3. Iterative methods for reconstructing the memory kernel. In this section we analyze two iterative methods for solving the inverse problem which have been suggested in the computational physics literature.

We start with the IMRV iteration [5] mentioned above.¹ It takes

$$(3.1) \quad k_0 = -\ddot{r}$$

as initial guess for the memory kernel and then continues recursively by computing

$$(3.2) \quad k_{n+1} = k_n + \ddot{r}_n - \ddot{r}, \quad n = 0, 1, 2, \dots$$

Here, $r_n = A(k_n)$ denotes the differential resolvent associated with k_n , i.e.,

$$(3.3) \quad \dot{r}_n(t) = - \int_0^t k_n(t - \tau) r_n(\tau) d\tau, \quad r_n(0) = I.$$

The IMRV scheme can be implemented in a *stochastic* or a *deterministic mode*: The stochastic mode—suggested in [5]—determines r_n by solving the generalized Langevin equation (1.1) with the memory kernel k_n instead of k and by evaluating the resulting autocorrelation function of the velocity. In the deterministic mode one either solves the above delay differential equation for r_n or the inhomogeneous equation

$$(3.4) \quad \ddot{r}_n(t) = - \int_0^t \dot{r}_n(t - \tau) k_n(\tau) d\tau - k_n(t), \quad \dot{r}_n(0) = 0,$$

for \dot{r}_n . The latter is obtained from (3.3) by first using (2.3) and then differentiating. Again we refer to [9] for numerical methods for solving (3.3) or (3.4).

In [11] Meyer, Pelagejcev, and Schilling proposed another iterative method for determining memory kernels in a more general nonequilibrium physical context. Restricted to the equilibrium case (1.1) and rewritten in our notation here, the method employs the very same initial guess (3.1) and proceeds by updating

$$(3.5) \quad k_{n+1} = k_n + \ddot{y}_n - \ddot{r}, \quad n = 0, 1, 2, \dots,$$

where y_n is given by

$$\dot{y}_n(t) = - \int_0^t r(t - \tau) k_n(\tau) d\tau, \quad y_n(0) = I,$$

and hence, \ddot{y}_n satisfies

$$(3.6) \quad \ddot{y}_n(t) = - \int_0^t \dot{r}(t - \tau) k_n(\tau) d\tau - k_n(t), \quad t > 0.$$

This can be viewed as a linearized version of the IMRV method because the function \dot{r}_n of (3.4), which depends nonlinearly on k_n , is replaced by a linear approximation. The computation of (3.6) can only be realized in the deterministic mode.

We mention that it follows from (3.6) that the update $\ddot{y}_n - \ddot{r}$ of (3.5) is the residual of the second-kind Volterra integral equation (2.5), which arises when k is replaced by k_n . This reveals that (3.5) is the standard Picard fixed point iteration (*aka* Richardson iteration, cf., e.g., Saad [14]) applied to the linear formulation (2.5) of the inverse problem.

¹IMRV stands for **I**terative **M**emory **R**econstruction using the **V**elocity autocorrelation data r as input.

It has to be emphasized that (in the above form) both iterative methods (3.2) and (3.5) are not fail-proof because some iterates k_n may happen not to be positive real, i.e., not to make sense physically.

Before we return to this aspect it is instructive to investigate the two methods in the Laplace domain. For this we denote by \widehat{k}_n , \widehat{r} , and \widehat{r}_n , the Laplace transforms of k_n , r , and r_n , respectively, and observe that the convolution equation (2.2) corresponds to

$$(3.7) \quad I - s\widehat{r}(s) = \widehat{k}(s)\widehat{r}(s), \quad \operatorname{Re} s \geq 0,$$

in the Laplace domain, i.e.,

$$(3.8) \quad \widehat{r}(s) = (sI + \widehat{k}(s))^{-1}.$$

Further, the iteration (3.2) corresponds to

$$\widehat{k}_{n+1}(s) = \widehat{k}_n(s) + s^2(\widehat{r}_n(s) - \widehat{r}(s)) = \widehat{k}_n(s) + s^2\widehat{r}_n(s)(\widehat{k}(s) - \widehat{k}_n(s))\widehat{r}(s),$$

and the Laplace transform \widehat{e}_n of the iteration error $e_n = k_n - k$ thus satisfies

$$(3.9) \quad \widehat{e}_{n+1}(s) = \widehat{e}_n(s) - s^2\widehat{r}_n(s)\widehat{e}_n(s)\widehat{r}(s).$$

In the scalar case ($d = 1$) this yields, for k_n close to k ,

$$\widehat{e}_{n+1}(s) \approx (1 - (s\widehat{r}(s))^2)\widehat{e}_n(s),$$

and therefore a necessary condition for local convergence of the IMRV iteration is that

$$(3.10) \quad |1 - (1 + \widehat{k}(s)/s)^{-2}| < 1$$

for all values of $s = i\xi$, $\xi \in \mathbb{R}$, along the imaginary axis.

The linearized iterative scheme (3.5), on the other hand, corresponds to

$$\widehat{k}_{n+1}(s) = \widehat{k}_n(s) - s\widehat{r}(s)\widehat{k}_n(s) + sI - s^2\widehat{r}(s) = \widehat{k}_n(s) - s\widehat{r}(s)\widehat{k}_n(s) + s\widehat{r}(s)\widehat{k}(s),$$

which corresponds to the recursion

$$(3.11) \quad \widehat{e}_{n+1}(s) = (I - s\widehat{r}(s))\widehat{e}_n(s)$$

for the corresponding iteration error; the scalar convergence criterion for the linearized scheme therefore amounts to

$$(3.12) \quad |1 - (1 + \widehat{k}(s)/s)^{-1}| < 1, \quad s = i\xi, \quad \xi \in \mathbb{R},$$

which is equivalent to

$$\operatorname{Im}(\widehat{k}(i\xi)/\xi) > -\frac{1}{2}, \quad \xi \in \mathbb{R}.$$

Figure 3.1 displays the resulting domains that constrain the values of $\widehat{k}(i\xi)/\xi$ to satisfy the convergence criterion (3.10) for the IMRV scheme and (3.12) for its linearized counterpart, respectively. Note that the origin corresponds to the most favorable value of $\widehat{k}(i\xi)/\xi$, for which the convergence of both schemes is superlinear.

EXAMPLE 3.1. A standard example of a memory kernel is given by

$$k(t) = e^{-t}, \quad t \geq 0.$$

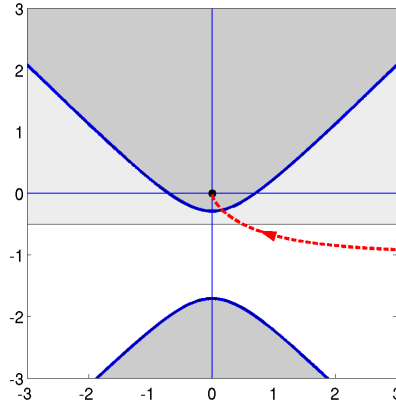


FIG. 3.1. Convergence domain in the complex plane for $\widehat{k}(i\xi)/\xi$, $\xi \in \mathbb{R}$, of the IMRV method (3.2) (darker gray) and for the linearized scheme (3.5) (lighter gray).

In this case

$$(3.13) \quad \widehat{k}(s) = \frac{1}{s+1}$$

and

$$\frac{\widehat{k}(i\xi)}{\xi} = \frac{1}{\xi + \xi^3} - i \frac{1}{1 + \xi^2},$$

and the dashed curve in Figure 3.1 shows the corresponding values of $\widehat{k}(i\xi)/\xi$ in the complex plane when ξ varies in \mathbb{R}^+ , the arrow indicating the orientation for increasing frequencies. (For negative values of ξ the sign of the real part is changing.) It can be seen that the points on the curve belong to the convergence domain for large frequencies ξ only.

Despite the shortcomings of a possible break down and the fact that the above example reveals that the two iterative methods cannot converge in general, they do perform reasonably well in practice when suitably adapted (see below). This is due to the fact that the iteration error decays rapidly for high frequencies, i.e., for large absolute values of ξ .

THEOREM 3.2. *Assume that the continuous memory kernel $k : \mathbb{R}_0^+ \rightarrow \mathbb{R}^{d \times d}$ is of strict positive type, is locally absolutely continuous, and that k and its derivative \dot{k} are absolutely integrable. Then the errors $e_n = k_n - k$ of the IMRV iterates (3.2) and of the linearized scheme (3.5) satisfy*

$$(3.14) \quad \widehat{e}_n(i\xi) = O(|\xi|^{-2n-3}), \quad \xi \in \mathbb{R}, |\xi| \rightarrow \infty,$$

for every $n \in \mathbb{N}_0$.

Proof. The assumption that k and its time derivative \dot{k} belong to $L^1(\mathbb{R}^+, \mathbb{R}^{d \times d})$ implies that, for $\xi \in \mathbb{R}$, we have

$$(3.15) \quad \widehat{k}(i\xi) = \frac{1}{i\xi} k(0) + \frac{1}{i\xi} \int_0^\infty e^{-i\xi t} \dot{k}(t) dt = O(1/|\xi|), \quad |\xi| \rightarrow \infty.$$

Since $k_0 = -\ddot{r}$ according to (3.1) and since $r(0) = I$ and $\dot{r}(0) = 0$, it follows from (3.7) that

$$\widehat{k}_0(s) = sI - s^2 \widehat{r}(s) = s \widehat{k}(s) \widehat{r}(s)$$

and

$$\widehat{e}_0(s) = s\widehat{k}(s)\widehat{r}(s) - \widehat{k}(s) = \widehat{k}(s)(s\widehat{r}(s) - I) = -(\widehat{k}(s))^2\widehat{r}(s).$$

As

$$(3.16) \quad \widehat{r}(i\xi) = O(1/|\xi|), \quad |\xi| \rightarrow \infty,$$

by virtue of (3.8), (3.14) holds true for $n = 0$ because of (3.15).

Now we proceed by induction. Consider the IMRV method first, and assume that the assertion (3.14) is true for some $n \in \mathbb{N}_0$. From (3.15) and the definition of \widehat{r}_n it then follows, cf. (3.8), that

$$i\xi \widehat{r}_n(i\xi) = \left(I + \frac{\widehat{k}_n(i\xi)}{i\xi} \right)^{-1} = \left(I + \frac{\widehat{k}(i\xi)}{i\xi} + \frac{\widehat{e}_n(i\xi)}{i\xi} \right)^{-1} = I + O(1/\xi^2)$$

as $|\xi| \rightarrow \infty$, $\xi \in \mathbb{R}$, and the same estimate is true with \widehat{r}_n replaced by \widehat{r} . Inserting this into (3.9) and using the induction hypothesis (3.14), we thus conclude that

$$\widehat{e}_{n+1}(i\xi) = \widehat{e}_n(i\xi) - (I + O(1/\xi^2)) \widehat{e}_n(i\xi) (I + O(1/\xi^2)) = O(|\xi|^{-2n-5})$$

for $|\xi| \rightarrow \infty$, as was to be shown.

For the iterative scheme (3.5) the argument is somewhat simpler because

$$I - i\xi \widehat{r}(i\xi) = \widehat{k}(i\xi) \widehat{r}(i\xi) = O(\xi^{-2}), \quad \xi \in \mathbb{R}, \quad |\xi| \rightarrow \infty,$$

according to (3.15) and (3.16). The induction step thus follows immediately from (3.11). \square

Note the analogy between this result and known strategies for approximating the memory kernel with Padé approximations of \widehat{k} , cf., e.g., [8, 10].

Translating Theorem 3.2 into the time domain, it says that for both iterative schemes² the iteration error e_n has $2n + 1$ continuous derivatives and there holds

$$e_n(0) = e'_n(0) = \dots = e_n^{(2n+1)}(0) = 0.$$

This means that the function value of k_n and its first $2n + 1$ derivatives at the origin coincide with those of the true memory kernel—provided that the latter exist. As a result, the approximate memories k_n adapt well to the true memory near zero, and the interval where this fit is good is likely to increase with n ; in fact, as stated in [5], the errors in the data fit “*did not disappear but were only shifted to larger times t* ”. A related statement concerning the linear iteration (3.5) can be found near the end of Section 2.2 in [12]. As a remedy it is suggested in [5] to restrict the update of the memory kernel to a sliding time window.

Because of our assumption that k is continuous, it follows from (2.5) that r has two continuous derivatives. This means that any lack of stability in solving the memory kernel from (2.5) or via the iterative schemes (3.2) or (3.5) is connected to the numerical computation of \ddot{r} , \ddot{r}_n , or \ddot{y}_n , respectively.

We finally mention that both iterative schemes can be modified by incorporating a relaxation parameter $\alpha > 0$ and updating

$$k_{n+1} = k_n + \alpha(\ddot{r}_n - \ddot{r}) \quad \text{or} \quad k_{n+1} = k_n + \alpha(\ddot{y}_n - \ddot{r}),$$

²Unless some iterate k_j , $j = 1, \dots, n$, fails to be strict positive real in the IMRV case.

respectively. For the linear scheme it is easy to see that, in the scalar case, the domain of convergence increases slightly for $0 < \alpha < 1$ and that convergence is assured if

$$\operatorname{Im}(\widehat{k}(i\xi)/\xi) > \frac{\alpha}{2} - 1, \quad \text{for all } \xi \in \mathbb{R}.$$

However, the Taylor approximation property of Theorem 3.2 is lost whenever $\alpha \neq 1$. Qualitatively the same holds true for the IMRV scheme. Therefore the introduction of such a relaxation parameter does not seem to make much sense.

4. A modified IMRV scheme. Another variant of the IMRV method has been used in [6] because it led to better approximations of the memory kernel than the original IMRV scheme when applied to a more complex physical application. This method employs first-order derivatives instead of second-order ones and updates the kernel according to

$$(4.1) \quad k_{n+1} = k_n + \alpha(\dot{r}_n - \dot{r}), \quad n = 0, 1, 2, \dots,$$

where r_n has the same meaning as in (3.3) and $\alpha > 0$ is a relaxation parameter.

Translated into the Laplace domain, the iteration (4.1) corresponds to

$$\widehat{k}_{n+1}(s) = \widehat{k}_n(s) + \alpha s(\widehat{r}_n(s) - \widehat{r}(s)) = \widehat{k}_n(s) + \alpha s\widehat{r}_n(s)(\widehat{k}(s) - \widehat{k}_n(s))\widehat{r}(s),$$

and hence, the iteration error obeys the recursion

$$(4.2) \quad \widehat{e}_{n+1}(s) = \widehat{e}_n(s) - \alpha s\widehat{r}_n(s)\widehat{e}_n(s)\widehat{r}(s).$$

Proceeding as in the previous section it turns out that the Laplace transforms \widehat{k}_n of k_n do not provide approximations with increasing accuracy at infinity. In fact, by making the preassumption that all iterates k_n stay in $L^1(\mathbb{R}^+, \mathbb{R}^{d \times d})$, it then follows from (3.8) that for $s = i\xi$, with $\xi \in \mathbb{R}$, it holds

$$\widehat{r}_n(i\xi) = \frac{1}{i\xi}I + o(1/|\xi|), \quad |\xi| \rightarrow \infty,$$

and similar for $\widehat{r}(i\xi)$. This shows that

$$\widehat{e}_{n+1}(i\xi) \approx \widehat{e}_n(i\xi) - \frac{\alpha}{i\xi}\widehat{e}_n(i\xi), \quad |\xi| \rightarrow \infty,$$

to leading order, so that

$$(4.3) \quad \|\widehat{e}_{n+1}(i\xi)\|^2 \approx (1 + \alpha^2/\xi^2) \|\widehat{e}_n(i\xi)\|^2 \quad \text{for } |\xi| \text{ large, } \xi \in \mathbb{R},$$

where $\|\cdot\|$ denotes the spectral norm in $\mathbb{R}^{d \times d}$. Since the leading factor on the right-hand side of (4.3) is greater than one, it follows that the error is stagnating or slightly increasing for large absolute frequencies. This shows that a result similar to Theorem 3.2 (under whatsoever assumptions on k) cannot hold true for the modified IMRV scheme.

On the other hand, it has to be emphasized that, if

$$k_0 = -\ddot{r}$$

as before, then $k_0(0) = -\ddot{r}(0) = k(0)$ because of (2.5), and then it follows from (4.1) and the definition (3.3) of the differential resolvents that

$$k_{n+1}(0) = k_n(0) = \dots = k(0)$$

for all $n \in \mathbb{N}_0$. Proceeding in the same way we conclude from (4.1) that

$$\dot{k}_{n+1}(0) = \dot{k}_n(0) + \alpha(\ddot{r}_n(0) - \ddot{r}(0)) = \dot{k}_n(0) + \alpha(k(0) - k_n(0)) = \dot{k}_n(0)$$

for all $n \in \mathbb{N}_0$ by virtue of (3.4), provided that k is continuously differentiable. Differentiating (2.5) once more we conclude that

$$\dot{k}(0) = -\ddot{r}(0) - \dot{r}(0)k(0) = -\ddot{r}(0) = \dot{k}_0(0),$$

proving that

$$\dot{k}_n(0) = \dot{k}(0) \quad \text{for all } n \in \mathbb{N}_0.$$

Repeating this argument one step further (sufficient smoothness assumed), we obtain that

$$\ddot{k}_n(0) = \dots = \ddot{k}_0(0) = -\ddot{r}(0) = \ddot{k}(0) - (k(0))^2$$

for all $n \in \mathbb{N}_0$, i.e., the curvature of all approximate memory kernels has the same (wrong) value—this gives evidence of the stagnation expected from the Laplace analysis.

We finally mention that, in the scalar case, a result similar to (4.3) also holds true for $s \rightarrow 0$, $s = i\xi$, with $\xi \in \mathbb{R}$: Since

$$\widehat{r}_n(i\xi) = \widehat{k}_n(0)^{-1} + O(|\xi|), \quad |\xi| \rightarrow 0,$$

where

$$\widehat{k}_n(0) = \int_0^\infty k_n(t) dt$$

is real and positive, it follows from (4.2) that

$$\widehat{e}_{n+1}(i\xi) \approx (1 - i\alpha\kappa\xi)\widehat{e}_n(i\xi)$$

for $|\xi| \approx 0$ and $k_n \approx k$, where

$$\kappa = \left(\int_0^\infty k(t) dt \right)^{-2} > 0.$$

This means that

$$|\widehat{e}_{n+1}(i\xi)|^2 \approx (1 + \alpha^2\kappa^2\xi^2)|\widehat{e}_n(i\xi)|^2 \quad \text{for } |\xi| \text{ small, } \xi \in \mathbb{R},$$

and for $k_n \approx k$, so that the error is stagnating or slightly increasing also for frequencies close to zero in the scalar case.

EXAMPLE 4.1. To illustrate this result consider again the exponential memory kernel from Example 3.1. In this example the recursion (4.2) simplifies for $k_n \approx k$ to

$$(4.4) \quad \widehat{e}_{n+1}(s) \approx (1 - \alpha s(\widehat{r}(s))^2)\widehat{e}_n(s).$$

Figure 4.1 displays the values of the factor

$$(4.5) \quad q_\alpha(s) = 1 - \alpha s(\widehat{r}(s))^2 = 1 - \alpha s \left(\frac{s+1}{s^2+s+1} \right)^2$$

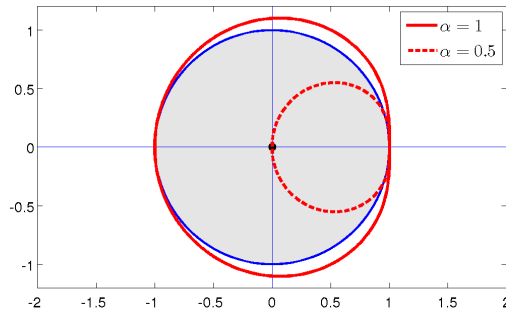


FIG. 4.1. Values of $q_\alpha(i\xi)$ of (4.5) in the complex plane as ξ varies in \mathbb{R} .

in front of $\widehat{e}_n(s)$ on the right-hand side of (4.4) along the imaginary axis, i.e., for $s = i\xi$, $\xi \in \mathbb{R}$. The absolute value of this factor—for $s \in i\mathbb{R}$ —needs to be below one for local convergence of the iteration (4.1).

The plot shows the corresponding values of q_α for the two relaxation parameters $\alpha = 1$ (solid line) and $\alpha = 0.5$ (dashed line). For $|\xi| \rightarrow 0$ and $|\xi| \rightarrow \infty$ these function values do indeed approach one from outside the unit disk. For $\alpha = 1$ and $\xi = \pm 1$ we have $q_1(\pm i) = -1$, and $|q_1(i\xi)| \geq 1$ for all other values of $\xi \in \mathbb{R}$. On the other hand, $q_{0.5}(i\xi)$ is mostly inside the unit disk; only for $|\xi|$ small and $|\xi|$ large this fails to hold true.

Despite the lack of convergence for very small and large absolute frequencies, the example above gives reason to hope for a decent performance of this modification of the IMRV scheme with underrelaxation, i.e., with a positive relaxation parameter $\alpha \ll 1$.

5. Mean square convergence. In the second part of this paper we determine topologies for approximating the memory kernel to guarantee mean square convergence of the generalized Langevin dynamics as the error in the memory kernel tends to zero. Before doing so we first recall the concise specification of the corresponding stochastic differential equation and its solution.

Let the continuous memory kernel $k : \mathbb{R}_0^+ \rightarrow \mathbb{R}^{d \times d}$ be absolutely integrable and of strict positive type, and let $F : \mathbb{R} \rightarrow \mathbb{R}^{d \times d}$ be a centered stationary Gaussian process on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with

$$\mathbb{E}(F(\tau + t)F(\tau)^*) = k(t), \quad t \geq 0.$$

Then a stationary solution of the corresponding generalized Langevin equation

$$(5.1) \quad \dot{V}(t) = - \int_{-\infty}^t k(t - \tau)V(\tau) d\tau + F(t),$$

is a centered Gaussian stochastic process $V : \mathbb{R} \rightarrow \mathbb{R}^d \times \mathbb{R}^d$ on the same probability space such that the joint process (F, V) is stationary. It is easy to verify that this process is given by

$$(5.2) \quad V(t) = \int_{-\infty}^t r(t - \tau)F(\tau) d\tau,$$

where r is the differential resolvent associated with k and V satisfies

$$\mathbb{E} (V(\tau + t)V(\tau)^*) = r(t), \quad t \geq 0,$$

and

$$\mathbb{E} (F(\tau + t)V(\tau)^*) = \int_{-\infty}^0 k(t - s)r(s) ds, \quad s \in \mathbb{R},$$

where we employ the extension $r(t) = r(-t)^*$ of the autocorrelation function to negative values of the time t .

Note that in the physics community the generalized Langevin equation is often considered to be a causal process, i.e., a process starting at $t = 0$ without any history and that the random force F is independent of previous velocities. This causal form of the generalized Langevin equation also has a stationary solution V with the very same autocorrelation function. However, Hohenegger and McKinley [4] showed that the joint (causal) process (F, V) fails to be stationary.

To formulate the main result of this section we start from two different memory kernels $k_j, j = 1, 2$, with the above properties and their associated differential resolvents

$$(5.3) \quad \dot{r}_j = - \int_0^t k_j(t - \tau)r_j(\tau) d\tau, \quad r_j(0) = I.$$

We think of k_1 to be the true memory kernel and of k_2 to be its approximation.

To simplify the exposition we employ the binary operation $*$ to denote the convolution

$$(\varphi * \psi)(t) = \int_0^t \varphi(t - \tau)\psi(\tau) d\tau, \quad 0 \leq t < \infty,$$

of two functions from $L^1(\mathbb{R}^+, \mathbb{R}^{d \times d})$ or of a function from $L^1(\mathbb{R}^+, \mathbb{R})$ and a function from $L^1(\mathbb{R}^+, \mathbb{R}^{d \times d})$. In both function spaces we denote the associated norm by $\|\cdot\|_{L^1(\mathbb{R}^+)}$, where in the matrix valued case, the corresponding integrand is the spectral norm of the associated matrices. As in Section 3 we write $\widehat{\varphi}$ for the Laplace transform of a given function φ from $L^1(\mathbb{R}^+, \mathbb{R})$ or from $L^1(\mathbb{R}^+, \mathbb{R}^{d \times d})$, respectively.

The following result is well-known; the proof is a combination of results and arguments from [3], which we include here for the ease of the reader.

LEMMA 5.1. *Let $k_1, k_2 \in L^1(\mathbb{R}^+, \mathbb{R}^{d \times d})$ be of strict positive type. Then r_1 and r_2 belong to $L^1(\mathbb{R}^+, \mathbb{R}^{d \times d})$, and there exist positive numbers $\delta = \delta(k_1)$ and $\ell = \ell(k_1)$ such that*

$$\|r_2 - r_1\|_{L^1(\mathbb{R}^+)} \leq \ell \|k_2 - k_1\|_{L^1(\mathbb{R}^+)},$$

provided that $\|k_2 - k_1\|_{L^1(\mathbb{R}^+)} \leq \delta$.

Proof. Let $f(t) = e^{-t}, t \geq 0$, and

$$(5.4) \quad \varphi_j = f * k_j - fI, \quad j = 1, 2.$$

Obviously, $\varphi_j \in L^1(\mathbb{R}^+, \mathbb{R}^{d \times d})$ and

$$I + \widehat{\varphi}_j(s) = (1 - \widehat{f}(s))I + \widehat{f}(s)\widehat{k}_j(s) = \frac{1}{s + 1} (sI + \widehat{k}_j(s))$$

by virtue of (3.13). For arbitrary $x \in \mathbb{C}^d, x \neq 0$, the function

$$u(s) = 2 \operatorname{Re} \left(x^* (sI + \widehat{k}_j(s))x \right) = 2(\operatorname{Re} s)x^*x + x^*(\widehat{k}_j(s) + \widehat{k}_j(s)^*)x$$

is a scalar harmonic function in the open right-half complex plane. Since k_j is strict positive real, u has positive boundary values on the imaginary axis, thus proving that $u(s) > 0$ for all $s \in \mathbb{C}$ with $\operatorname{Re} s \geq 0$ by the maximum principle for harmonic functions. This shows that $sI + \widehat{k}_j(s)$ and $I + \widehat{\varphi}_j(s)$ are invertible for every s with $\operatorname{Re} s \geq 0$.

The Paley-Wiener theory thus implies that the second-kind convolution equation

$$(5.5) \quad q_j + \varphi_j * q_j = \varphi_j = q_j + q_j * \varphi_j$$

has a unique solution $q_j \in L^1(\mathbb{R}^+, \mathbb{R}^{d \times d})$, cf. [3, Section 2.4], the so-called resolvent associated with φ_j . It is easy to verify that

$$r_j = fI - f * q_j \in L^1(\mathbb{R}^+, \mathbb{R}^{d \times d})$$

is the differential resolvent associated with k_j , i.e., that it solves (5.3), and hence

$$(5.6) \quad \|r_2 - r_1\|_{L^1(\mathbb{R}^+)} \leq \|f\|_{L^1(\mathbb{R}^+)} \|q_2 - q_1\|_{L^1(\mathbb{R}^+)} = \|q_2 - q_1\|_{L^1(\mathbb{R}^+)}.$$

From (5.5) it follows that

$$\begin{aligned} q_2 + ((\varphi_2 - \varphi_1) - q_1 * (\varphi_2 - \varphi_1)) * q_2 &= q_2 + (\varphi_2 - q_1 - q_1 * \varphi_2) * q_2 \\ &= \varphi_2 - q_1 * (q_2 + \varphi_2 * q_2) = \varphi_2 - q_1 * \varphi_2, \end{aligned}$$

and therefore

$$\begin{aligned} \|q_2\|_{L^1(\mathbb{R}^+)} - \|\varphi_2 - \varphi_1\|_{L^1(\mathbb{R}^+)} (1 + \|q_1\|_{L^1(\mathbb{R}^+)}) &\|q_2\|_{L^1(\mathbb{R}^+)} \\ &\leq \|\varphi_2\|_{L^1(\mathbb{R}^+)} (1 + \|q_1\|_{L^1(\mathbb{R}^+)}) \\ &\leq (\|\varphi_1\|_{L^1(\mathbb{R}^+)} + \|\varphi_2 - \varphi_1\|_{L^1(\mathbb{R}^+)}) (1 + \|q_1\|_{L^1(\mathbb{R}^+)}). \end{aligned}$$

Note that (5.4) implies that

$$\|\varphi_2 - \varphi_1\|_{L^1(\mathbb{R}^+)} \leq \|f\|_{L^1(\mathbb{R}^+)} \|k_2 - k_1\|_{L^1(\mathbb{R}^+)} = \|k_2 - k_1\|_{L^1(\mathbb{R}^+)},$$

so that

$$(5.7) \quad \|q_2\|_{L^1(\mathbb{R}^+)} \leq \frac{(\|\varphi_1\|_{L^1(\mathbb{R}^+)} + \|k_2 - k_1\|_{L^1(\mathbb{R}^+)}) (1 + \|q_1\|_{L^1(\mathbb{R}^+)})}{1 - \|k_2 - k_1\|_{L^1(\mathbb{R}^+)} (1 + \|q_1\|_{L^1(\mathbb{R}^+)}} ,$$

provided that $\|k_2 - k_1\|_{L^1(\mathbb{R}^+)}$ is so small that the denominator on the right-hand side of (5.7) is positive.

Another use of (5.5) establishes the second-kind Volterra integral equation

$$q_2 - q_1 + \varphi_1 * (q_2 - q_1) = \psi$$

for $q_2 - q_1$ with right-hand side

$$\psi = \varphi_2 - \varphi_1 - (\varphi_2 - \varphi_1) * q_2.$$

The solution of this integral equation can be written in terms of the associated resolvent q_1 , cf., [3, Section 2.3], i.e.,

$$q_2 - q_1 = \psi - q_1 * \psi,$$

which yields the estimate

$$\begin{aligned}
 \|q_2 - q_1\|_{L^1(\mathbb{R}^+)} &\leq (1 + \|q_1\|_{L^1(\mathbb{R}^+)})\|\psi\|_{L^1(\mathbb{R}^+)} \\
 &\leq (1 + \|q_1\|_{L^1(\mathbb{R}^+)})\|(1 + \|q_2\|_{L^1(\mathbb{R}^+)})\|\varphi_2 - \varphi_1\|_{L^1(\mathbb{R}^+)} \\
 &\leq (1 + \|q_1\|_{L^1(\mathbb{R}^+)})\|(1 + \|q_2\|_{L^1(\mathbb{R}^+)})\|k_2 - k_1\|_{L^1(\mathbb{R}^+)}.
 \end{aligned}$$

The assertion thus follows from (5.6) and (5.7). \square

Recall that any centered stationary Gaussian process $F : \mathbb{R} \rightarrow \mathbb{R}^{d \times d}$ over a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ satisfying

$$\mathbb{E}(F(\tau + t)F(\tau)^*) = k(t), \quad t \geq 0,$$

for some $k \in L^1(\mathbb{R}^+, \mathbb{R}^{d \times d})$ has a spectral representation

$$(5.8) \quad F(t) = \int_{-\infty}^{\infty} e^{i\xi t} z(\xi) dW(\xi),$$

where

$$(5.9) \quad z(\xi) = \frac{1}{\sqrt{2\pi}} (\widehat{k}(i\xi) + \widehat{k}(i\xi)^*)^{1/2}, \quad \xi \in \mathbb{R},$$

and where W is an associated two-sided d -dimensional Brownian motion over the same probability space; see, e.g., [2]. Note that the matrix square root (5.9) is well-defined because every autocorrelation function is of positive type.

In the sequel we denote by

$$\|F\|_{L^2(\Omega)} = (\mathbb{E}(F^*F))^{1/2}$$

the square root of the expected Euclidean norm squared of a d -dimensional real random variable F on Ω .

THEOREM 5.2. *Let $k_1, k_2 : \mathbb{R}_0^+ \rightarrow \mathbb{R}^{d \times d}$ be continuous and absolutely integrable functions of strict positive type, and assume $\widehat{k}_2 - \widehat{k}_1$ to be absolutely integrable over the imaginary axis. Denote by (F_2, V_2) the stationary solution of (5.1) with memory kernel $k = k_2$ on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Then the generalized Langevin equation (5.1) with $k = k_1$ has a particular stationary solution (F_1, V_1) on the same probability space with*

$$(5.10) \quad \sup_{t \in \mathbb{R}} \|F_2(t) - F_1(t)\|_{L^2(\Omega)} \leq 4\sqrt{d/\pi^3} \|\widehat{k}_2 - \widehat{k}_1\|_{L^1(i\mathbb{R})}^{1/2}.$$

Moreover, there exist positive numbers c and δ , depending only on k_1 , such that

$$(5.11) \quad \|V_2(t) - V_1(t)\|_{L^2(\Omega)} \leq c\|k_2 - k_1\|_{L^1(\mathbb{R}^+)} + c\|\widehat{k}_2 - \widehat{k}_1\|_{L^1(i\mathbb{R})}^{1/2}$$

for every $t \in \mathbb{R}$, provided that $\|k_2 - k_1\|_{L^1(\mathbb{R}^+)} \leq \delta$.

Proof. Denote by z_j the matrix square root (5.9) for $k = k_j, j = 1, 2$. Let

$$F_2(t) = \int_{-\infty}^{\infty} e^{i\xi t} z_2(\xi) dW(\xi)$$

be the spectral representation of F_2 , cf. (5.8), and define

$$(5.12) \quad F_1(t) = \int_{-\infty}^{\infty} e^{i\xi t} z_1(\xi) dW(\xi)$$

for $t \in \mathbb{R}$, using the very same Brownian motion. Then F_1 is a stationary centered Gaussian process on $(\Omega, \mathcal{F}, \mathbb{P})$ with

$$\mathbb{E} (F_1(\tau + t)F_1(\tau)^*) = k_1(t), \quad t \geq 0.$$

We make use of an estimate by Vainikko (see [15, p. 91]) for the stability of the square root of self-adjoint nonnegative definite operators (with respect to the spectral norm $\| \cdot \|$), namely

$$\sqrt{2\pi} \|z_2(\xi) - z_1(\xi)\| \leq \frac{4}{\pi} \|(\widehat{k}_2(i\xi) + \widehat{k}_2(i\xi)^*) - (\widehat{k}_1(i\xi) + \widehat{k}_1(i\xi)^*)\|^{1/2}.$$

Given our assumption that $\widehat{k}_2 - \widehat{k}_1$ is absolutely integrable over the imaginary axis, the Itô calculus thus yields

$$\begin{aligned} \|F_2(t) - F_1(t)\|_{L^2(\Omega)}^2 &= \text{trace} \left(\int_{-\infty}^{\infty} (z_2(\xi) - z_1(\xi))(z_2(\xi) - z_1(\xi))^* d\xi \right) \\ &\leq d \frac{8}{\pi^3} \int_{-\infty}^{\infty} \|\widehat{k}_2(i\xi) + \widehat{k}_2(i\xi)^* - \widehat{k}_1(i\xi) - \widehat{k}_1(i\xi)^*\| d\xi \\ &\leq d \frac{16}{\pi^3} \int_{-\infty}^{\infty} \|\widehat{k}_2(i\xi) - \widehat{k}_1(i\xi)\| d\xi \end{aligned}$$

for every $t \in \mathbb{R}$, and this implies (5.10).

Let r_j be the differential resolvent (5.3) of k_j , $j = 1, 2$. Then $r_j \in L^1(\mathbb{R}^+, \mathbb{R}^{d \times d})$ by Lemma 5.1, and, according to (5.2),

$$(5.13) \quad V_j(t) = \int_{-\infty}^t r_j(t - \tau)F_j(\tau) d\tau, \quad j = 1, 2,$$

is the velocity component of the stationary solution (F_j, V_j) of (5.1) with $k = k_j$ driven by F_j . Since

$$\|F_1(t)\|_{L^2(\Omega)}^2 = \text{trace} \left(\mathbb{E}(F_1(t)F_1(t)^*) \right) = \text{trace}(k_1(0))$$

for every $t \in \mathbb{R}$, we conclude from (5.10) and (5.13) that

$$\begin{aligned} &\|V_2(t) - V_1(t)\|_{L^2(\Omega)} \\ &\leq \|r_2 - r_1\|_{L^1(\mathbb{R}^+)} \left(\text{trace}(k_1(0)) \right)^{1/2} + \|r_2\|_{L^1(\mathbb{R}^+)} \sup_{t \in \mathbb{R}} \|F_2(t) - F_1(t)\|_{L^2(\Omega)} \\ &\leq \|r_2 - r_1\|_{L^1(\mathbb{R}^+)} \left(\text{trace}(k_1(0)) \right)^{1/2} + 4\sqrt{d/\pi^3} \|r_2\|_{L^1(\mathbb{R}^+)} \|\widehat{k}_2 - \widehat{k}_1\|_{L^1(i\mathbb{R})}^{1/2}. \end{aligned}$$

The desired inequality (5.11) thus follows from Lemma 5.1. \square

Theorem 5.2 provides a mean pathwise error bound for the two processes (F_1, V_1) and (F_2, V_2) , and this requires a proper “alignment” of the driving random forces F_2 and F_1 , which is achieved via (5.12). The spirit of this result is similar to the notion of strong convergence of numerical algorithms for stochastic differential equations. The error bounds in the theorem state that the expected distance between the processes F_2 and F_1 as well as V_2 and V_1 depends on the error in the memory kernel, measured with respect to $L^1(\mathbb{R}^+)$ and also in terms of the

L^1 -norm of its Laplace transform along the imaginary axis; the latter, however, enters via its square root into these bounds.

Note that the assumption of Theorem 5.2 that $\widehat{k}_2 - \widehat{k}_1$ is absolutely integrable over the imaginary axis means that $k_2 - k_1$ has a certain smoothness (for example, two absolutely integrable derivatives are sufficient) and that $k_2(0) = k_1(0)$; compare (3.15).

6. Conclusions. In this work we have analyzed three iterative methods that have recently been introduced in the computational physics community to retrieve the unknown memory kernel of a generalized Langevin equation from autocorrelation data of the resulting process. Although all three methods are, in principle, doomed to diverge eventually according to our analysis, they do have some appealing theoretical properties which justify their consideration in practical applications. Concerning their performance in these applications and a discussion of their respective merits, we refer to [5, 6, 11, 12].

The implementations which have been published in these references incorporate additional post-processing steps to stabilize the iteration; these features add further nonlinearities and increase their complexity, making it difficult to provide a conclusive theoretical analysis. Our theoretical results are based on the assumption that the autocorrelation data and their derivatives are given exactly. In the applications mentioned before these data are extracted from numerical simulations, and hence, they are discrete and fail to be exact. The numerical results in these works indicate a certain robustness of the considered iterative methods with respect to discretization and measurement errors.

As an independent contribution we have seen in Section 5 that, if the primary purpose of retrieving the memory kernel is to simulate the given process, then the memory should be approximated in the norm of L^1 —both in time and along the imaginary axis in the Laplace domain. In fact, the approximation in the Laplace domain is the more restrictive one, and it enters with a sublinear Hölder-type estimate into the error bound.

Acknowledgements. The author likes to thank Janka Bauer, Gerhard Jung, Peter Kloeden, and Friederike Schmid for helpful discussions.

REFERENCES

- [1] M. CHEN, X. LI, AND C. LIU, *Computation of the memory functions in the generalized Langevin models for collective dynamics of macromolecules*, J. Chem. Phys., 141 (2014), Art. 064112, 12 pages.
- [2] H. CRAMÉR AND M.R. LEADBETTER, *Stationary and Related Stochastic Processes*, Wiley, New York, 1967.
- [3] G. GRIPENBERG, S.-O. LONDEN, AND O. STAFFANS, *Volterra Integral and Functional Equations*, Cambridge University Press, Cambridge, 1990.
- [4] C. HOHENEGGER AND S.A. MCKINLEY, *Reconstructing complex fluid properties from the behavior of fluctuating immersed particles*, SIAM J. Appl. Math., 78 (2018), pp. 2200–2226.
- [5] G. JUNG, M. HANKE, AND F. SCHMID, *Iterative reconstruction of memory kernels*, J. Chem. Theory Comput., 13 (2017), pp. 2481–2488.
- [6] ———, *Generalized Langevin dynamics: construction and numerical integration of non-Markovian particle-based models*, Soft Matter, 14 (2018), pp. 9368–9382.
- [7] P.K. LAMM, *A survey of regularization methods for first-kind Volterra equations*, in *Surveys on Solution Methods for Inverse Problems*, D. Colton, H. W. Engl, A. K. Louis, J. R. McLaughlin, and W. Rundell, eds., Springer, Vienna, 2000, pp. 53–82.
- [8] H. LEI, N.A. BAKER, AND X. LI, *Data-driven parameterization of the generalized Langevin equation*, Proc. Natl. Acad. Sci. USA, 113 (2016), pp. 14183–14188.
- [9] P. LINZ, *Analytical and Numerical Methods for Volterra Equations*, SIAM, Philadelphia, 1985.
- [10] L. MA, X. LI, AND C. LIU, *The derivation and approximation of coarse-grained dynamics from Langevin dynamics*, J. Chem. Phys., 145 (2016), Art. 204117, 13 pages.
- [11] H. MEYER, P. PELAGEJCEV, AND T. SCHILLING, *Non-Markovian out-of-equilibrium dynamics: A general numerical procedure to construct time-dependent memory kernels for coarse-grained observables*, Europhys. Lett., 128 (2019), Art. 40001, 7 pages.

- [12] H. MEYER, S. WOLF, G. STOCK, AND T. SCHILLING, *A numerical procedure to evaluate memory effects in non-equilibrium coarse-grained models*, *Adv. Theory Simul*, 4 (2020), Art. 2000197, 8 pages.
- [13] N. POTTIER, *Nonequilibrium Statistical Processes*, Oxford University Press, Oxford, 2010.
- [14] Y. SAAD, *Iterative Methods for Sparse Linear Systems*, 2nd ed., SIAM, Philadelphia, 2003.
- [15] G.M. VAINIKKO AND A.Y. VERETENNIKOV, *Iteration Procedures in Ill-Posed Problems*, Nauka, Moskau, 1986. (In Russian.)