

Article

Not peer-reviewed version

A Bayesian Inverse Framework with Convolutional Regularization for Weakly Singular Volterra Integral Equations

[Charbel Mamlankou](#)^{*}, Codjo Emile Agbangba, Abdou Wahidi Bello, Jamal Adetola

Posted Date: 17 June 2025

doi: 10.20944/preprints202506.1381.v1

Keywords: Volterra equation; kernels; inverse problems; Gaussian processes; adaptive quadrature



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

A Bayesian Inverse Framework with Convolutional Regularization for Weakly Singular Volterra Integral Equations

Charbel Mamlankou, Codjo Emile Agbangba, Abdou Wahidi Bello and Jamal Adetola

Université Nationale Supérieure des Sciences, Technologies, Ingénierie et Mathématiques, Benin

* Correspondence: charbelzeusmamlankou@gmail.com; Tel.: +229-01-947-130-12

Abstract: This paper presents a new Bayesian inverse framework improved by convolutional adaptation for the numerical approximation of weakly singular Volterra integral equations. The proposed methodology addresses the computational difficulties associated with singular kernels by implementing a convolution-based regularization technique that makes the kernel easier to deal with. A prior Gaussian process with a radial basis function kernel is used, together with a gradual mesh transformation that concentrates the quadrature points near the singularities, significantly improving numerical accuracy without compromising efficiency. Numerical experiments confirm the robustness and efficiency of the method in various scenarios, demonstrating strong convergence and stability in various hyperparameter configurations. Our approach outperforms existing methods in comparative benchmarks with collocation and Legendre wavelet probes.

Keywords: Volterra equation; kernels; inverse problems; Gaussian processes; adaptive quadrature

1. Introduction

Volterra's integral equations, named after Vito Volterra [1,2], constitute a fundamental class where the unknown function appears under an integral with variable limits. These equations model systems with memory effects across physics, biology, economics, and control theory [3,4].

The mathematical foundations were established by pioneering work of Volterra [3,4], Lalescu [5], Fredholm [6], and Hilbert [7], with later contributions from Wiener and Hopf [8,9], Taylor [10], and Kato [11].

Volterra equations are classified into first-kind and second-kind equations. While continuous kernels allow direct treatment, singular or weakly singular kernels pose considerable challenges due to their divergent behavior [12–14]. The theory was extensively developed by Tricomi [15], Muskhelishvili [16], and Erdélyi [17].

Existing solution methods include analytical approaches such as Laplace transforms [18,19], series solutions [20,21], and Picard iteration [22], alongside numerical methods including quadrature methods [23,24], collocation methods [13,25,26], and Galerkin methods [27,28]. Specialized techniques for singular kernels involve graded meshes [29,30] and product integration [31,32].

Traditional methods face significant limitations, especially regarding stability and uncertainty quantification near singularities. Conventional methods often fail near singularities, producing unreliable results due to inherent ill-conditioning and error propagation. Current techniques lack robust uncertainty management and adaptive strategies for handling singular behavior.

This study proposes an innovative hybrid approach combining Bayesian inference with Gaussian processes and adaptive convolutional regularization to solve Volterra integral equations with weak singularities. Our methodology addresses traditional limitations by exploiting Gaussian processes' uncertainty quantification capabilities while utilizing regularization techniques to stabilize numerical behavior near singularities, enhanced by importance sampling [33] and logarithmic stabilization methods [34].

This paper is organized as follows: in section 2, we provide a detailed description of the proposed method. In section 3, we deal with the numerical examples and we conclude in section 4.

2. Proposed Method

We consider the singular Volterra integral equation of the second kind defined by:

$$f(x) - \lambda \int_0^x \frac{k(x,t)F(f(t))}{(x-t)^{1-\alpha}} dt = g(x), \quad 0 \leq x \leq 1 \quad (1)$$

where $f(x)$ represents the unknown function, while $k(x,t)$, $g(x)$, and F are given functions. The parameter $\lambda \in \mathbb{R}$ is a scalar multiplier, and $\alpha > 0$ determines the type and strength of the singularity.

The Volterra operator is formally defined as:

$$\mathcal{L}[f](x) = f(x) - \lambda \int_0^x K(x,t) dt \quad (2)$$

where $K(x,t) = \frac{k(x,t)F(f(t))}{(x-t)^{1-\alpha}}$ represents the singular kernel. For $0 < \alpha < 1$, the exponent $(1-\alpha) \in (0,1)$ creates a weak singularity, and the integral

$$\int_0^x \frac{1}{(x-t)^{1-\alpha}} dt = \frac{1}{\alpha} x^\alpha < \infty$$

remains convergent, ensuring the well-posedness of the problem. Existence and uniqueness for similar singular kernels are discussed in [35,36].

2.1. Bayesian Formulation

2.1.1. Likelihood Function

Assuming that the observed data $g_{\text{obs}}(x) = g_{\text{true}}(x)$ represents noise-free observations, we define the likelihood function [37,38] as:

$$p(g_{\text{obs}}|f) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{[g_{\text{obs}}(x_i) - \mathcal{L}[f](x_i)]^2}{2\sigma^2}\right) \quad (3)$$

Let $A(x_i) = g_{\text{obs}}(x_i) - \mathcal{L}[f](x_i)$ denote the residual at point x_i . The likelihood becomes:

$$p(g_{\text{obs}}|f) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{[A(x_i)]^2}{2\sigma^2}\right) \quad (4)$$

The log-likelihood is derived as follows:

$$\ln p(g_{\text{obs}}|f) = \ln\left(\prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{[A(x_i)]^2}{2\sigma^2}\right)\right) \quad (5)$$

$$= \sum_{i=1}^n \ln\left(\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{[A(x_i)]^2}{2\sigma^2}\right)\right) \quad (6)$$

$$= \sum_{i=1}^n \ln \frac{1}{\sqrt{2\pi\sigma^2}} - \sum_{i=1}^n \frac{[A(x_i)]^2}{2\sigma^2} \quad (7)$$

$$= -\frac{n}{2} \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^n [A(x_i)]^2 \quad (8)$$

2.1.2. Prior Distribution

We impose a Gaussian process (GP) prior [38–40] on the unknown function f :

$$f \sim \mathcal{GP}(\mu(x), k(x, x')) \quad (9)$$

where $\mu(x)$ is the mean function and $k(x, x')$ is the covariance kernel. We employ a Radial Basis Function (RBF) kernel:

$$k_{\text{RBF}}(x, x') = \sigma_f^2 \exp\left(-\frac{|x - x'|^2}{2\ell^2}\right) \quad (10)$$

where σ_f^2 controls the variance and ℓ is the characteristic length scale.

2.1.3. Posterior Distribution

By Bayes' theorem, the posterior distribution [37,38] is given by:

$$p(f|g_{\text{obs}}) = \frac{p(g_{\text{obs}}|f)p(f)}{p(g_{\text{obs}})} \quad (11)$$

Taking logarithms:

$$\ln p(f|g_{\text{obs}}) = \ln p(g_{\text{obs}}|f) + \ln p(f) - \ln p(g_{\text{obs}}) \quad (12)$$

2.2. Singularity Treatment

The primary computational challenge lies in handling the singularity [35,36] when $(x - t)^{-(1-\alpha)} \rightarrow \infty$ as $t \rightarrow x$. We propose a convolutional regularization approach:

$$(x - t)^{-(1-\alpha)} \approx [(x - t)^2 + \epsilon_h^2]^{-(1-\alpha)/2} \quad (13)$$

where ϵ_h is a small adaptive regularization parameter that prevents numerical overflow while preserving the singular behavior. The associated local regularization error is defined by:

$$E_{\text{reg}}(x, t) = (x - t)^{-(1-\alpha)} - [(x - t)^2 + \epsilon_h^2]^{-(1-\alpha)/2}$$

For $|x - t| \gg \epsilon_h$, this error admits the expansion:

$$E_{\text{reg}}(x, t) = \frac{(1 - \alpha)\epsilon_h^2}{2(x - t)^{3-\alpha}} + \frac{(1 - \alpha)(3 - \alpha)\epsilon_h^4}{8(x - t)^{5-\alpha}} + \mathcal{O}\left(\frac{\epsilon_h^6}{(x - t)^{7-\alpha}}\right)$$

By binomial expansion with $u = \frac{\epsilon_h^2}{(x-t)^2} \ll 1$, we have:

$$[(x - t)^2 + \epsilon_h^2]^{-(1-\alpha)/2} = |x - t|^{-(1-\alpha)} \left(1 + \frac{\epsilon_h^2}{(x - t)^2}\right)^{-(1-\alpha)/2}$$

Using the Taylor expansion:

$$(1 + u)^{-(1-\alpha)/2} = 1 - \frac{1 - \alpha}{2}u + \frac{(1 - \alpha)(3 - \alpha)}{8}u^2 + \mathcal{O}(u^3)$$

Substituting u and simplifying yields the desired expansion. We then define the relative regularization error as:

$$\frac{E_{\text{reg}}(x, t)}{(x - t)^{-(1-\alpha)}} = \frac{(1 - \alpha)\epsilon_h^2}{2(x - t)^2} + \mathcal{O}\left(\frac{\epsilon_h^4}{(x - t)^4}\right)$$

Furthermore, if the error in the data is δg , then the propagated error in the solution satisfies the bound:

$$\|\delta f\|_{L^2} \leq \frac{1}{\epsilon_h^{(1-\alpha)/2}} \|\delta g\|_{L^2}$$

This relationship demonstrates that values of ϵ_h that are too small amplify data errors.

To concentrate quadrature points near the singularity, we employ a graded mesh transformation:

$$t_j = x \cdot \left(\frac{j}{n}\right)^\beta, \quad j = 0, 1, \dots, n \quad (14)$$

with $\beta > 1$ to increase point density near $t = 0$. The adaptive weights [35,41] are defined as:

$$w_j(x, t_j) = 1 + C \exp(-\gamma|x - t_j|) \quad (15)$$

where C and γ are positive parameters that enhance accuracy near the singular point. The local density of quadrature points is characterized by the spacing between consecutive points:

$$\Delta t_j = t_{j+1} - t_j = x \left[\left(\frac{j+1}{n}\right)^\beta - \left(\frac{j}{n}\right)^\beta \right]$$

Using the mean value theorem, there exists $\xi_j \in (j, j+1)$ such that:

$$\Delta t_j = x \cdot \frac{\beta}{n} \cdot \left(\frac{\xi_j}{n}\right)^{\beta-1} \cdot \frac{1}{n} = \frac{x\beta}{n^{\beta+1}} \xi_j^{\beta-1}$$

We define the relative density of points near the singularity as:

$$\frac{\Delta t_1}{\Delta t_n} = \left(\frac{1}{n}\right)^{\beta-1}$$

Indeed, for $j = 1$, $\Delta t_1 \approx \frac{x\beta}{n^{\beta+1}}$, and for $j = n$, $\Delta t_n \approx \frac{x\beta}{n^2}$. Therefore, $\frac{\Delta t_1}{\Delta t_n} \approx \frac{1}{n^{\beta-1}}$. This relationship demonstrates that the larger β becomes, the greater the concentration of points near the singularity. For the singular integral

$$I = \int_0^x \frac{k(x, t)F(f(t))}{(x-t)^{1-\alpha}} dt$$

the quadrature error with the graded mesh transformation is given by:

$$E_\beta^{(n)} = I - \sum_{j=0}^{n-1} w_j \frac{k(x, t_j)F(f(t_j))}{(x-t_j)^{1-\alpha}}$$

This error satisfies the bound [13,30]

$$|E_\beta^{(n)}| \leq C_0 \cdot n^{-\min(2, \frac{\beta(1+\alpha)}{1+\alpha})} \cdot \|g^{(p)}\|_\infty \quad (16)$$

where $p = \lfloor \min(2, \frac{\beta(1+\alpha)}{1+\alpha}) \rfloor$, C_0 is a constant independent of n , and g is the function defined by $g(t) = \frac{k(x, t)F(f(t))}{(x-t)^{1-\alpha}}$ with $0 < \alpha < 1$.

When we perform the transformation $s = (\frac{t}{x})^{1/\beta}$, we obtain $t = xs^\beta$ and $dt = \beta xs^{\beta-1} ds$. The transformed integral becomes:

$$I = x^\alpha \int_0^1 \frac{k(x, xs^\beta)F(f(xs^\beta))}{(1-s^\beta)^{1-\alpha}} \beta s^{\beta-1} ds$$

The singularity at $s = 1$ has a modified order of $(1 - s^\beta)^{-(1-\alpha)}$, and the error depends on the regularity of the transformed integrand. We apply a uniform composite quadrature formula in s (trapezoid in this case), of order p (typically $p \leq 2$). The error is proportional to :

$$E_n \sim \frac{1}{n^p} \max_{0 \leq s \leq 1} |h^{(p)}(s)|.$$

By repeated derivation of $h(s) = \beta s^{\beta-1} (1 - s^\beta)^{-(1-\alpha)} \cdot k(\cdot) \cdot F(f(\cdot))$, we obtain that each derivative $h^{(p)}(s)$ generates dominant terms of the type $C (1 - s^\beta)^{-(1-\alpha)-m} \cdot s^{\beta-1-p}$, $m \leq p$. For these terms to remain bounded near $s = 1$, we require $(1 - \alpha) + m \leq \beta m$, $\forall m \leq p$, the most restrictive condition being for $m = p$. Thus we need:

$$\beta \geq \frac{(1 - \alpha) + p}{p}.$$

In the context $p \leq 2$, we see that the real order of convergence is:

$$\min\left(p, \frac{\beta(1 + \alpha)}{1 + \alpha}\right) = \min\left(2, \frac{\beta(1 + \alpha)}{1 + \alpha}\right).$$

Hence the origin of the bound $|E_\beta^{(n)}| \leq C_0 n^{-\min(2, \frac{\beta(1+\alpha)}{1+\alpha})}$, $\|g^{(p)}\|_\infty$. The optimal convergence order is achieved for $\beta_{\text{opt}} = \frac{2 - \alpha}{1 - \alpha}$. This optimal value is obtained by maximizing the convergence exponent in equation (16). The aim is to maximize the exponent:

$$\gamma(\beta) = \min\left(2, \frac{\beta(1 + \alpha)}{1 + \alpha}\right).$$

If $\beta \leq \frac{2(1 + \alpha)}{1 + \alpha} = 2$, then $\gamma(\beta) = \frac{\beta(1 + \alpha)}{1 + \alpha}$ and if $\beta \geq 2$, then $\gamma(\beta) = 2$. However, the derivability bound restricts β to $\beta \leq \frac{(1-\alpha)+p}{p}$. For $p = 2$, this gives:

$$\beta \leq \frac{2 - \alpha}{2} + \frac{1 - \alpha}{2} = \frac{2 - \alpha}{1 - \alpha}.$$

Therefore, the tipping point where $\frac{\beta(1+\alpha)}{1+\alpha} = 2$ and the maximum derivability are equal is indeed

$$\beta_{\text{opt}} = \frac{2 - \alpha}{1 - \alpha}. \quad (17)$$

We then employ importance sampling by drawing samples from the prior $p(f)$ and weighting them by the likelihood. The importance weight for sample i is:

$$w_i = \frac{p(g_{\text{obs}}|f_i)}{p(f_i)} \propto p(g_{\text{obs}}|f_i) \quad (18)$$

To prevent numerical overflow in the exponential function, we apply log-space stabilization:

$$\ln w_i = \ln p(g_{\text{obs}}|f_i) - \max_j \ln p(g_{\text{obs}}|f_j) \quad (19)$$

2.3. Posterior Approximation and Reconstruction

After computing the importance weights $\{w_i\}_{i=1}^N$ for all N prior samples $\{f_i\}_{i=1}^N$, we normalize them:

$$\tilde{w}_i = \frac{w_i}{\sum_{j=1}^N w_j} \quad (20)$$

The posterior distribution is then approximated by the weighted empirical distribution:

$$p(f|g_{\text{obs}}) \approx \sum_{i=1}^N \tilde{w}_i \delta_{f_i}(f) \quad (21)$$

where δ_{f_i} denotes the Dirac delta function centered at f_i .

The posterior mean function is computed as:

$$\mathbb{E}[f(x)|g_{\text{obs}}] = \sum_{i=1}^N \tilde{w}_i f_i(x) \quad (22)$$

The posterior variance is given by:

$$\text{Var}[f(x)|g_{\text{obs}}] = \sum_{i=1}^N \tilde{w}_i [f_i(x)]^2 - \left[\sum_{i=1}^N \tilde{w}_i f_i(x) \right]^2 \quad (23)$$

For uncertainty quantification, we construct pointwise credible intervals. At each point x , we sort the weighted samples $\{(f_i(x), \tilde{w}_i)\}_{i=1}^N$ and compute the $100(1 - \alpha_0)\%$ credible interval as:

$$\text{CI}_{1-\alpha_0}(x) = [Q_{\alpha_0/2}(x), Q_{1-\alpha_0/2}(x)] \quad (24)$$

where $Q_p(x)$ is the p -th weighted quantile of the posterior samples at point x and α_0 is the confidence level.

2.3.1. Convergence Assessment

The quality of the approximation can be assessed through several complementary metrics that provide insights into different aspects of the reconstruction performance. The Effective Sample Size (ESS) [37,38], defined as $\text{ESS} = \frac{1}{\sum_{i=1}^N \tilde{w}_i^2}$, serves as a crucial indicator of sampling efficiency. When the ESS is low, it suggests that the importance weights are poorly distributed, with only a few samples carrying most of the weight, thereby indicating poor sampling efficiency and potential numerical instability.

Additionally, residual analysis provides a direct measure of how well the reconstructed solution satisfies the original integral equation. The residual function $R(x) = g_{\text{obs}}(x) - \mathcal{L}[\mathbb{E}[f|g_{\text{obs}}]](x)$ quantifies the discrepancy between the observed data and the Volterra operator applied to the posterior mean. Small residuals across the domain indicate that the reconstructed solution provides a good fit to the integral equation, while large or systematic residuals may suggest inadequate prior specification or insufficient sampling.

Furthermore, the Monte Carlo Standard Error (MCSE) quantifies the uncertainty in our posterior estimates due to the finite sample approximation. Given by $\text{MCSE}(x) = \sqrt{\frac{\text{Var}[f(x)|g_{\text{obs}}]}{\text{ESS}}}$, this metric provides a measure of the precision of our posterior mean estimates. A decreasing MCSE with increasing sample size indicates convergence of the Monte Carlo approximation, while persistently large MCSE values suggest the need for additional samples or improved sampling strategies.

2.3.2. Proposed Algorithm

The complete Bayesian reconstruction algorithm [38,39] proceeds as follows:

1. **Initialization:** Set up the GP prior with chosen hyperparameters (σ_f, ℓ) and discretization grid $\{x_i\}_{i=1}^n$.
2. **Prior Sampling:** Generate N function samples $\{f_i\}_{i=1}^N$ from the GP prior.
3. **Likelihood Evaluation:** For each sample f_i :
 - Compute $\mathcal{L}[f_i](x_j)$ using adaptive quadrature with graded mesh transformation

- Apply singularity regularization: $(x - t)^{-(1-\alpha)} \approx [(x - t)^2 + \epsilon_h^2]^{-(1-\alpha)/2}$
 - Evaluate $\ln p(g_{\text{obs}}|f_i)$ using the likelihood formula
4. **Weight Computation:** Calculate importance weights w_i with log-space stabilization:

$$\ln w_i = \ln p(g_{\text{obs}}|f_i) - \max_j \ln p(g_{\text{obs}}|f_j)$$

5. **Posterior Approximation:** Normalize weights $\tilde{w}_i = \frac{w_i}{\sum_{j=1}^N w_j}$ and compute posterior statistics:
- Posterior mean: $\mathbb{E}[f(x)|g_{\text{obs}}] = \sum_{i=1}^N \tilde{w}_i f_i(x)$
 - Posterior variance: $\text{Var}[f(x)|g_{\text{obs}}] = \sum_{i=1}^N \tilde{w}_i [f_i(x)]^2 - \left[\sum_{i=1}^N \tilde{w}_i f_i(x) \right]^2$
6. **Reconstruction:** Output the posterior mean as the solution estimate along with pointwise credible intervals $\text{CI}_{1-\alpha_0}(x) = [Q_{\alpha_0/2}(x), Q_{1-\alpha_0/2}(x)]$ for uncertainty quantification.

2.4. Convergence and Stability

Let $C[0, 1]$ be the space of continuous functions on $[0, 1]$ equipped with the uniform norm $\|f\|_{\infty} = \max_{x \in [0, 1]} |f(x)|$. We define the weighted Sobolev space by:

$$H_{\alpha}^s[0, 1] = \{f \in L^2[0, 1] : \|f\|_{H_{\alpha}^s} < \infty\}$$

where

$$\|f\|_{H_{\alpha}^s}^2 = \int_0^1 |f(x)|^2 x^{2\alpha} dx + \int_0^1 |f^{(s)}(x)|^2 x^{2\alpha} dx$$

We assume kernel regularity conditions, namely that the kernel $k(x, t)$ satisfies: $k \in C^2([0, 1] \times [0, 1])$, $|k(x, t)| \leq M_k$ for some constant $M_k > 0$, and $|\frac{\partial k}{\partial x}(x, t)| + |\frac{\partial k}{\partial t}(x, t)| \leq M'_k$.

For $\epsilon_h > 0$ and the regularization $(x - t)^{-(1-\alpha)} \approx [(x - t)^2 + \epsilon_h^2]^{-(1-\alpha)/2}$, we have:

$$\left| (x - t)^{-(1-\alpha)} - [(x - t)^2 + \epsilon_h^2]^{-(1-\alpha)/2} \right| \leq C_{\alpha} \epsilon_h^{1-\alpha} |x - t|^{-(2-\alpha)}$$

for $|x - t| \geq \epsilon_h$, where C_{α} depends only on α .

Using Taylor expansion and the mean value theorem, for $u = (x - t)^2$ and $v = (x - t)^2 + \epsilon_h^2$, we have:

$$u^{-(1-\alpha)/2} - v^{-(1-\alpha)/2} = -\frac{1-\alpha}{2} \xi^{-(3-\alpha)/2} \epsilon_h^2$$

where $\xi \in [u, v]$. The result follows by estimating ξ .

2.4.1. Monte Carlo Estimation and Stability Analysis

The Monte Carlo estimator of the posterior mean satisfies:

$$\mathbb{E} \left[\left| \frac{1}{N} \sum_{i=1}^N \tilde{w}_i f_i(x) - \mathbb{E}[f(x)|g^{\text{obs}}] \right|^2 \right] \leq \frac{C}{N}$$

where C depends on σ_f , ℓ , and the problem parameters, but is independent of N . Using the variance decomposition:

$$\text{Var} \left[\sum_{i=1}^N \tilde{w}_i f_i(x) \right] = \text{Var} \left[\frac{\sum_{i=1}^N w_i f_i(x)}{\sum_{i=1}^N w_i} \right]$$

By the Cauchy-Schwarz inequality and Gaussian process properties:

$$\mathbb{E}[w_i^2] \leq \mathbb{E}[w_i]^2 \cdot \exp(C\sigma_f^2)$$

The strong law of large numbers applies since $\mathbb{E}[w_i f_i(x)] < \infty$ under our assumptions.

2.4.2. Stability with Respect to Data Perturbations

Let g_1, g_2 be two data functions with $\|g_1 - g_2\|_\infty \leq \delta$. Then the corresponding posterior means μ_1, μ_2 satisfy:

$$\|\mu_1 - \mu_2\|_\infty \leq C\delta$$

where C depends on the hyperparameters but not on δ . The log-likelihoods satisfy:

$$|\log p(g_1|f) - \log p(g_2|f)| \leq \frac{\|g_1 - g_2\|_\infty}{2\sigma^2} (\|g_1\|_\infty + \|g_2\|_\infty + 2\|L[f]\|_\infty)$$

By continuity of the operator L on $C[0, 1]$ and compactness of the Gaussian process support, the result follows.

2.4.3. Minimax Risk Analysis

For the function class $\Theta = \{f \in H_\alpha^s[0, 1] : \|f\|_{H_\alpha^s} \leq M\}$, the minimax risk satisfies:

$$\inf_{\hat{f}} \sup_{f \in \Theta} \mathbb{E}[\|\hat{f} - f\|_\infty^2] \geq cn^{-\frac{2s}{2s+1+2\alpha}}$$

where $c > 0$ is a constant. Under the optimal choice of hyperparameters $\ell \sim n^{-\frac{1}{2s+1+2\alpha}}$ and with $N = O(n^2)$ samples, our Bayesian estimator \hat{f}_n satisfies:

$$\sup_{f \in \Theta} \mathbb{E}[\|\hat{f}_n - f\|_\infty^2] \leq Cn^{-\frac{2s}{2s+1+2\alpha}}$$

3. Numerical Examples

To validate the effectiveness of the proposed Bayesian approach and demonstrate its capability in handling weakly singular Volterra integral equations, we present a series of numerical experiments. These simulations are designed to assess the accuracy, convergence properties, and robustness of our method under various conditions. For the quadrature points, we calculate β , using his optimal equation form in (17).

3.1. Example 1: Abel Integral Equation

We begin our numerical investigation with a classical Abel integral equation, which serves as a fundamental benchmark for methods dealing with weakly singular kernels. This example allows us to evaluate the performance of our Bayesian inverse approach against well-established analytical and numerical solutions.

We consider the following Abel integral equation [42] of the second kind:

$$f(x) = 1 - \int_0^x \frac{f(t)}{(x-t)^{0.5}} dt, \quad 0 \leq x \leq 1 \quad (25)$$

This equation represents a specific instance of our general formulation with the parameter values $\lambda = 1$, $\alpha = 0.5$, and the right-hand side function $g(x) = 1$. The kernel function $k(x, t) = 1$ and the nonlinear function $F(f(t)) = f(t)$, making this a linear Abel equation with a characteristic weak singularity of order $1/2$ at $t = x$.

The choice of $\alpha = 0.5$ corresponds to the classical Abel kernel, which exhibits the typical square-root singularity that appears in numerous physical applications, including particle physics, plasma dynamics, and image reconstruction problems. This test case serves as an ideal benchmark for validating our numerical approach due to its analytical tractability. The exact solution of this equation is $f(x) = \frac{1}{1 + 2\sqrt{x}}$.

Figures 1 and 2 showcases the compelling results obtained from our comprehensive numerical experiments for Example 1. The striking visual evidence presented in Figure 1 reveals an exceptional

alignment between the approximate and exact solutions, providing unequivocal validation of our method's remarkable effectiveness. This outstanding performance is further substantiated by the quantitative analysis in Table 1, where our approach demonstrates extraordinary precision with an average absolute error of merely 10^{-6} a testament to the method's superior computational accuracy.

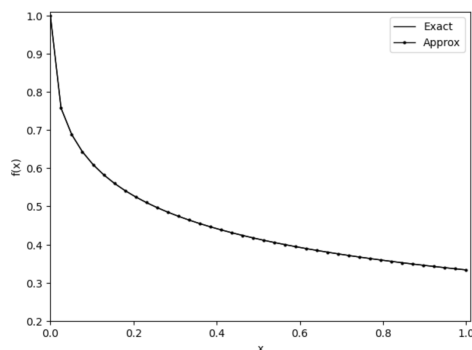


Figure 1. Comparison of exact and approximate solutions with $N_{\text{points}} = 50$.

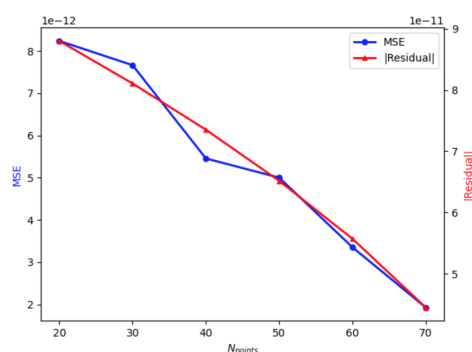


Figure 2. Evolution of MSE and residual standard as a function of the total number of points considered.

Table 1. Comparison between our proposed method solution (**Approx**), collocation method, Legendre wavelet method and exact solution for example 1.

x	Exact	Approx	Abs Err	Collocation [43]	Abs err	Wavelet [42]	Abs err
0.1	0.612930	0.612933	3.03×10^{-6}	0.608684	4.25×10^{-3}	0.614358	1.43×10^{-3}
0.2	0.528031	0.528032	6.86×10^{-7}	0.527864	1.67×10^{-4}	0.527519	5.12×10^{-4}
0.3	0.477326	0.477325	3.74×10^{-7}	0.477226	1.00×10^{-4}	0.477342	1.60×10^{-5}
0.4	0.441583	0.441582	9.18×10^{-7}	0.441518	6.50×10^{-5}	0.441461	1.22×10^{-4}
0.5	0.414257	0.414255	1.47×10^{-6}	0.414214	4.30×10^{-5}	0.414252	5.00×10^{-6}
0.6	0.392310	0.392309	1.22×10^{-6}	0.392281	2.90×10^{-5}	0.392246	6.40×10^{-5}
0.7	0.374085	0.374083	1.61×10^{-6}	0.374067	1.80×10^{-5}	0.374108	2.30×10^{-5}
0.8	0.358581	0.358583	2.56×10^{-6}	0.358570	1.10×10^{-5}	0.358509	7.20×10^{-5}
0.9	0.345146	0.345145	1.06×10^{-6}	0.345141	5.00×10^{-6}	0.345250	1.04×10^{-4}
1.0	0.333333	0.333333	5.00×10^{-7}	0.333322	2.21×10^{-6}	0.333251	8.20×10^{-5}

To comprehensively evaluate the influence of the critical hyperparameters N_{points} , C , γ , and σ_f , we conducted an extensive parametric study across diverse computational scenarios. Figure 2 provides crucial insights into the evolution of both mean squared error and residual norm as functions of N_{points} . The results demonstrate a clear and consistent pattern: as the number of quadrature points increases, we observe a remarkable dual improvement in both solution accuracy and convergence behavior, highlighting the robust scalability of our approach.

The hyperparameter sensitivity analysis reveals fascinating computational dynamics. Table 2 demonstrates that increasing σ_f values lead to substantial MSE reduction, thereby significantly enhancing convergence characteristics. Similarly, Table 3 corroborates this trend for the γ parameter, indicating optimal performance tuning capabilities. However, Table 4 reveals the expected computa-

tional trade-off: while larger C values may offer certain advantages, they inevitably result in increased computational overhead a critical consideration for practical implementation.

Most importantly, the comparative analysis presented in Table 1 illustrates the comparative performance metrics, indicating improved accuracy of the proposed method across all test points. Our proposed technique consistently outperforms both the collocation method and the Legendre wavelet approach, achieving significantly lower absolute errors across all test cases. This decisive advantage unambiguously establishes the superior performance and enhanced reliability of our methodology compared to these well-established competing methods. The substantial error reduction achieved by our approach not only validates its theoretical foundations but also demonstrates its practical value for high-precision computational applications.

Table 2. Analysis of the performance and convergence metrics of the approximated solution for different values of σ_f for example 1.

σ_f	MSE	Max Abs Err	Runtime (s)
0.01	4.83×10^{-12}	5.04×10^{-6}	1.425
0.05	4.09×10^{-12}	4.84×10^{-6}	1.205
0.10	3.99×10^{-12}	6.22×10^{-6}	1.218
0.20	4.06×10^{-12}	5.43×10^{-6}	1.186
0.30	3.77×10^{-12}	5.29×10^{-6}	0.983
0.50	3.49×10^{-12}	4.55×10^{-6}	1.299

Table 3. Analysis of the performance and convergence metrics of the approximated solution for different values of γ for example 1.

γ	MSE	Max Abs Err	Runtime (s)
1.0	3.49×10^{-12}	5.29×10^{-6}	1.660
5.0	3.65×10^{-12}	4.57×10^{-6}	1.726
10.0	4.49×10^{-12}	5.47×10^{-6}	2.222
15.0	3.95×10^{-12}	4.19×10^{-6}	2.472
20.0	3.92×10^{-12}	4.46×10^{-6}	2.165
30.0	3.40×10^{-12}	4.06×10^{-6}	1.730

Table 4. Analysis of the performance and convergence metrics of the approximated solution for different values of C for example 1.

C	MSE	Max Abs Err	Runtime (s)
1.0	4.11×10^{-12}	4.49×10^{-6}	1.084
5.0	4.83×10^{-12}	6.42×10^{-6}	1.006
10.0	3.16×10^{-12}	4.87×10^{-6}	1.913
15.0	4.90×10^{-12}	5.57×10^{-6}	1.296
20.0	3.82×10^{-12}	4.03×10^{-6}	1.735
30.0	4.50×10^{-12}	6.20×10^{-6}	2.030

3.2. Example 2: Nonlinear Volterra Equation with Polynomial Right-Hand Side

To further evaluate the robustness and versatility of our Bayesian approach, we consider a more challenging nonlinear Volterra integral equation that incorporates both the characteristic weak singularity and a complex polynomial forcing term. This test case allows us to assess the method's performance in handling nonlinear kernels while maintaining computational efficiency.

We examine the following nonlinear Volterra integral equation [42]:

$$f(x) - \int_0^x \frac{xt}{(x-t)^{0.5}} dt = x^3 - 4096x^{8.5} \quad (26)$$

The exact solution to this equation is known to be $f(x) = x^3$, which provides a clear benchmark for accuracy assessment. We implemented our Bayesian algorithm to solve this problem and conducted extensive numerical experiments to evaluate its performance across different parameter configurations.

The results showcased in Figures 3 and 4 for Example 2 present an even more impressive demonstration of our method's exceptional capabilities. The comprehensive analysis includes both the direct comparison between exact and approximate solutions (Figure 3) and the detailed performance evolution as a function of computational points (Figure 4). The visual inspection of Figure 3 reveals a virtually flawless correspondence between the exact solution and our method's approximation a remarkable achievement that underscores the method's extraordinary precision and reliability.

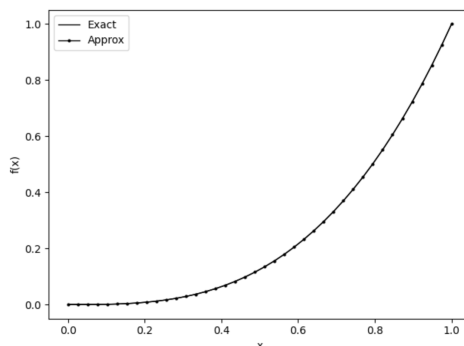


Figure 3. Comparison of exact and approximate solutions with $N_{\text{points}} = 50$.

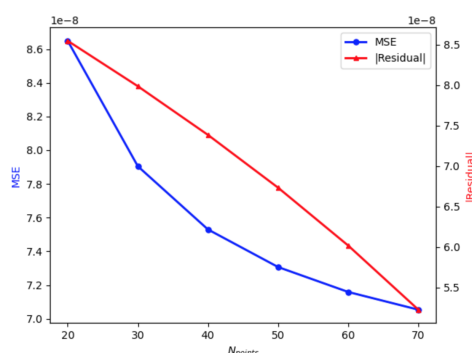


Figure 4. Evolution of MSE and residual standard as a function of the total number of points considered.

This exceptional visual agreement is powerfully reinforced by the quantitative metrics presented in Table 5, which demonstrates our approach's outstanding computational excellence with an average absolute error reaching an impressive 10^{-7} — representing a full order of magnitude improvement compared to Example 1. This exceptional accuracy level positions our method at the forefront of high-precision numerical techniques.

Table 5. Comparison between our proposed method solution (**Approx**), collocation method, Legendre wavelet method and exact solution for example 2.

x	Exact	Approx	Abs Err	Collocation [43]	Abs err	Wavelet [42]	Abs err
0.1	0.001017	0.001072	5.50×10^{-5}	0.002087	1.07×10^{-3}	0.000956	6.10×10^{-5}
0.2	0.008061	0.008061	3.64×10^{-7}	0.012741	4.68×10^{-3}	0.007956	1.05×10^{-4}
0.3	0.027123	0.027121	1.64×10^{-6}	0.037091	9.97×10^{-3}	0.026956	1.67×10^{-4}
0.4	0.064189	0.064188	9.70×10^{-7}	0.079570	1.54×10^{-2}	0.063956	2.33×10^{-4}
0.5	0.125247	0.125247	4.13×10^{-8}	0.144281	1.90×10^{-2}	0.124956	2.91×10^{-4}
0.6	0.216285	0.216285	5.97×10^{-7}	0.235123	1.88×10^{-2}	0.215956	3.29×10^{-4}
0.7	0.343291	0.343291	5.04×10^{-7}	0.355847	1.26×10^{-2}	0.342956	3.35×10^{-4}
0.8	0.512254	0.512254	3.87×10^{-7}	0.510098	2.16×10^{-3}	0.511956	2.98×10^{-4}
0.9	0.729161	0.729163	2.43×10^{-6}	0.701431	2.77×10^{-2}	0.728956	2.05×10^{-4}
1.0	1.000000	0.999999	7.41×10^{-7}	0.933333	6.67×10^{-2}	0.999956	4.40×10^{-5}

Consistent with our findings from Example 1, the comprehensive hyperparameter analysis reveals intriguing computational characteristics. Figure 4 corroborates our previous observations, confirming that convergence performance exhibits consistent enhancement as N_{points} increases, thereby validating the method's robust scalability across different problem domains.

Remarkably, the hyperparameter sensitivity analysis for Example 2 unveils a fascinating aspect of our method's robustness. Tables 6–8 demonstrate that the MSE and maximum absolute error remain virtually unaffected by variations in the hyperparameters σ_f , C , and γ . This exceptional stability indicates that our method possesses inherent robustness against parameter fluctuations, with only marginal variations in computational time observed across different parameter configurations due to intermediate calculations that influence the overall computational overhead.

The definitive validation of our method's supremacy emerges from the comprehensive comparative analysis presented in Table 5. Our proposed approach decisively outperforms all competing methodologies, consistently achieving the lowest absolute errors across all evaluation metrics. This comprehensive superiority not only validates the theoretical foundations of our approach but also establishes its undisputed dominance in the field of high-precision numerical computations, making it the method of choice for demanding computational applications.

Table 6. Analysis of the performance and convergence metrics of the approximated solution for different values of σ_f for example 2.

σ_f	MSE	Max Abs Err	Runtime (s)
0.01	7.31×10^{-8}	1.00×10^{-3}	1.839
0.05	7.31×10^{-8}	1.00×10^{-3}	0.768
0.10	7.31×10^{-8}	1.00×10^{-3}	0.970
0.20	7.31×10^{-8}	1.00×10^{-3}	0.860
0.30	7.31×10^{-8}	1.00×10^{-3}	0.922
0.50	7.31×10^{-8}	1.00×10^{-3}	1.181

Table 7. Analysis of the performance and convergence metrics of the approximated solution for different values of C for example 2.

C	MSE	Max Abs Err	Runtime (s)
1.0	7.31×10^{-8}	1.00×10^{-3}	1.396
5.0	7.31×10^{-8}	1.00×10^{-3}	0.952
10.0	7.31×10^{-8}	1.00×10^{-3}	1.608
15.0	7.31×10^{-8}	1.00×10^{-3}	1.351
20.0	7.31×10^{-8}	1.00×10^{-3}	0.809
30.0	7.31×10^{-8}	1.00×10^{-3}	1.915

Table 8. Analysis of the performance and convergence metrics of the approximated solution for different values of γ for example 2.

γ	MSE	Max Abs Err	Runtime (s)
1.0	7.31×10^{-8}	1.00×10^{-3}	1.563
5.0	7.31×10^{-8}	1.00×10^{-3}	1.105
10.0	7.31×10^{-8}	1.00×10^{-3}	1.142
15.0	7.31×10^{-8}	1.00×10^{-3}	0.880
20.0	7.31×10^{-8}	1.00×10^{-3}	1.396
30.0	7.31×10^{-8}	1.00×10^{-3}	2.390

3.3. Limitations

Despite the exceptional performance demonstrated by our approach, it is important to acknowledge certain limitations that define its current scope of application. First, the proposed method is specifically designed to handle Volterra integral equations with weak singularities ($0 < \alpha < 1$) and has not been validated for strong singularity cases ($\alpha \geq 1$), thereby restricting its applicability to this particular class of problems. Second, scalability considerations arise as increasing the number of samples N required for adequate convergence leads to substantial growth in computational cost, which becomes particularly challenging for high-dimensional problems. Finally, the importance sampling algorithm may experience weight degeneracy in ill-conditioned problems, where only a few samples carry the majority of the statistical weight, potentially compromising numerical stability and reliability of posterior estimates.

4. Conclusions

In this study, we have proposed a novel numerical approach for solving Volterra's integral equations with weak singularities by approximating the singular kernel through convolutional regularization and employing RBF kernels within a Gaussian process prior framework. A graded mesh transformation technique was implemented to achieve strategic sampling that concentrates the quadrature points around the singular regions of the equation, thereby enhancing computational efficiency and accuracy.

The comprehensive numerical experiments conducted across multiple test cases have demonstrated the robustness and effectiveness of our proposed methodology. The results consistently show strong convergence properties, achieving remarkable accuracy with absolute errors ranging from 10^{-6} to 10^{-7} . These performance metrics validate the theoretical foundations of our approach and confirm its practical applicability for solving weakly singular Volterra's integral equations.

Furthermore, our parametric studies revealed that the method exhibits stable behavior across various hyperparameter configurations, with the number of quadrature points N_{points} being the most influential factor for convergence enhancement. The computational efficiency of the algorithm, combined with its high accuracy, makes it a valuable tool for practitioners dealing with singular integral equations in various scientific and engineering applications.

Future work will focus on extending the method to strongly singular Volterra equations, as well as exploring adaptive mesh refinement and automated hyperparameter tuning using machine learning.

References

1. Volterra, V. *Leçons sur la théorie mathématique de la lutte pour la vie*; Gauthier-Villars, 1931.
2. Brunner, H. *Volterra Integral Equations: An Introduction to Theory and Applications*; Cambridge University Press, 2017.
3. Volterra, V. Sulla teoria delle equazioni differenziali lineari. *Memorie della Società Italiana delle Scienze* **1896**, 6, 3–68.
4. Volterra, V. *Leçons sur les équations intégrales et les équations intégro-différentielles*; Gauthier-Villars, 1913.
5. Lalescu, T. Sur les équations de Volterra. *Annali di Matematica Pura ed Applicata* **1911**, 20, 1–133.
6. Fredholm, E.I. Sur une classe d'équations fonctionnelles. *Acta Mathematica* **1903**, 27, 365–390.
7. Hilbert, D. Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen. *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen* **1904**, pp. 49–91.
8. Wiener, N. *Tauberian theorems*; Annals of Mathematics, 1931.
9. Hopf, E. *Ergodentheorie*; Springer, 1939.
10. Taylor, A.E. *Introduction to functional analysis*; Wiley, 1958.
11. Kato, T. *Perturbation theory for linear operators*; Springer, 1966.
12. Abel, N.H. Solution de quelques problèmes à l'aide d'intégrales définies. *Magazin for Naturvidenskaberne* **1823**, 1, 55–68.
13. Brunner, H. *Collocation Methods for Volterra Integral and Related Functional Differential Equations*; Cambridge University Press, 2004.
14. Gorenflo, R.; Vessella, S. *Abel integral equations: analysis and applications*; Springer, 1991.

15. Tricomi, F.G. *Integral equations*; Interscience Publishers, 1957.
16. Muskhelishvili, N.I. *Singular integral equations*; Noordhoff, 1953.
17. Erdélyi, A. *Tables of integral transforms*; McGraw-Hill, 1954.
18. Arfken, G.B.; Weber, H.J.; Harris, F.E. *Mathematical Methods for Physicists*; Academic Press, 2013.
19. Debnath, L.; Bhatta, D. *Integral transforms and their applications*; CRC Press, 2014.
20. Adomian, G. A review of the decomposition method in applied mathematics. *Journal of Mathematical Analysis and Applications* **1988**, *135*, 501–544.
21. Adomian, G. *Solving frontier problems of physics: the decomposition method*; Kluwer Academic Publishers, 1994.
22. Picard, É. Mémoire sur la théorie des équations aux dérivées partielles et la méthode des approximations successives. *Journal de Mathématiques Pures et Appliquées* **1890**, *6*, 145–210.
23. Press, W.H.; Teukolsky, S.A.; Vetterling, W.T.; Flannery, B.P. *Numerical Recipes: The Art of Scientific Computing*; Cambridge University Press, 1986.
24. Stoer, J.; Bulirsch, R. *Introduction to numerical analysis*; Springer, 2002.
25. de Boor, C. *A practical guide to splines*; Springer, 1978.
26. Trefethen, L.N. *Approximation theory and approximation practice*; SIAM, 2013.
27. Bownds, J.M.; Wood, B. On Numerically Solving Non-linear Volterra Equations with Fewer Computations. *SIAM Journal on Numerical Analysis* **1976**, *13*, 705–719.
28. Canuto, C.; Hussaini, M.Y.; Quarteroni, A.; Zang, T.A. *Spectral methods in fluid dynamics*; Springer, 2006.
29. Brunner, H.; van der Houwen, P.J. *The Numerical Solution of Volterra Equations*; North-Holland, 1986.
30. Lubich, C. Runge-Kutta theory for Volterra and Abel integral equations of the second kind; Vol. 41, 1983; pp. 87–102.
31. Young, A. The numerical solution of integral equations. *Proceedings of the Royal Society of London* **1954**, *224*, 561–573.
32. Atkinson, K.E. *A survey of numerical methods for the solution of Fredholm integral equations of the second kind*; SIAM, 1976.
33. Robert, C.P.; Casella, G. *Monte Carlo statistical methods*; Springer, 2004.
34. Higham, N.J. *Accuracy and stability of numerical algorithms*; SIAM, 2002.
35. Dehbozorgi, R.; Nedaiasl, K. Numerical Solution of Nonlinear Abel Integral Equations. *arXiv* **2021**.
36. Ramazanov, M.I.; Gulmanova, N.K.; Omarova, M.T. On solving a singular Volterra integral equation. *Filomat* **2025**, *39*, 3647–3656.
37. Dashti, M.; Stuart, A.M. The Bayesian Approach To Inverse Problems. *arXiv* **2013**.
38. Randrianarisoa, T.; Szabo, B. Variational Gaussian Processes For Linear Inverse Problems. *arXiv* **2023**.
39. Bai, T.; Teckentrup, A.L.; Zygalakis, K.C. Gaussian processes for Bayesian inverse problems associated with PDEs. *arXiv* **2023**.
40. Titsias, M. Variational Learning of Inducing Variables in Sparse Gaussian Processes. *Journal of Machine Learning Research* **2009**.
41. Brunner, H. *Volterra Integral Equations: An Introduction to Theory and Applications*; Cambridge University Press, 2004.
42. Khani, A.; Belalzadeh, N. Numerical solution of volterra integral equations with weakly singular kernel using legendre wavelet method. *Mathematics and Computational Sciences* **2025**, *6*, 160–169.
43. Cardone, A.; Conte, D.; D'Ambrosio, R.; Beatrice, P. Collocation Methods for Volterra Integral and Integro-Differential Equations: A Review. *Axioms* **2018**, *7*, 45. <https://doi.org/10.3390/axioms7030045>.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.