Bayesian identification of sound sources with the Helmholtz equation

Sebastian Engel · Dominik Hafemeyer · Christian Münch · Daniel Schaden

Received: date / Accepted: date - The correct dates will be entered by the editor

Abstract In this work we discuss the problem of identifying sound sources from pressure measurements with a Bayesian approach. The acoustics are modelled by the Helmholtz equation and the goal is to get information about number, strength and position of the sound sources, under the assumption that measurements of the acoustic pressure are noisy. We propose a problem specific prior distribution of the number, the amplitudes and positions of the sound sources and algorithms to compute an approximation of the associated posterior. We also discuss a finite element discretization of the Helmholtz equation for the practical computation and prove convergence rates of the discretized posterior to the true posterior. The theoretical results are illustrated by numerical experiments, which indicate that the proven rates are actually sharp.

Keywords Uncertainty Quantification · Bayesian Inverse Problem · Sound Source Identification · Finite Elements · Non-Gaussian, Sequential Monte Carlo

1 Introduction

Throughout the paper, unless noted otherwise, all functions and measures are complex valued. Let $D \subseteq \mathbb{R}^d$, $d = 2, 3$, be an bounded convex polygonal/polyhedral domain. Furthermore let the boundary be decomposed in the form $\partial D =: \Gamma = \Gamma_N \cup \Gamma_Z$, with $\mathcal{H}^{d-1}(\Gamma_Z) > 0$, where $\Gamma_N := \bigcup_{j=1}^J \Gamma_j$ can be written as the union of some faces of $\Gamma$ and $\Gamma_Z := \Gamma \setminus \Gamma_N$. $\mathcal{H}^{d-1}$ denotes here the $d-1$-dimensional Hausdorff measure. We model the acoustic pressure $y_u$ with angular frequency $\xi$ and speed of sound $c$ as the solution of

$$\begin{cases} -\Delta y_u - \left(\frac{\xi}{c}\right)^2 y_u = \tau(u) & \text{in } D, \\ \partial_\nu y_u - i \frac{\xi}{c} y_u = 0 & \text{on } \Gamma_Z, \\ \partial_\nu y_u = g & \text{on } \Gamma_N. \end{cases} $$ (1)

The Helmholtz equation describes a stationary wave. It appears in the modelling of acoustics or electromagnetism (see for example [Dautray and Lions, 2012, Chapter 2, §8, Section 7] or [Kosiljakov et al., 1964, Chapters VI, XXV] for a more physics based approach). To be more specific, $g \in \mathcal{H}^{\frac{1}{2}}(\Gamma_N)$ is a complex function which models the amplitude of a primary sound source on a part of the boundary and $\tau(u)$ is a complex linear combination, possibly even a series, of Dirac measures with support in $D$. The Dirac measures model sound sources, which we want to identify. $\rho$ describes the density of the
fluid and $\gamma_\zeta \in \mathbb{C}$ is the wall impedance given by

$$\gamma_\zeta := \beta_\zeta + \frac{\alpha_\zeta}{\zeta}i. \quad (2)$$

The frequency-dependent material constants $\alpha_\zeta > 0$ and $\beta_\zeta > 0$ are related to the viscous and elastic response of the isolating material. However, in what follows, we will assume only that $\beta_\zeta \neq 0$. The boundary condition on $\Gamma_Z$ allows the modelling of an absorbing viscoelastic material covering the boundary walls. The boundary conditions on $\Gamma_N$ models external influence on the acoustic pressure.

Our aim is to deduce the distribution of the number, locations and amplitudes of the sound sources $\tau(u)$ from pressure measurements $y := (y_\alpha(z_j))^m_{j=1} + \eta \in \mathbb{C}^m$ at finitely many locations in the domain. The pressure measurements are taken at distinct points $(z_j)^m_{j=1}$ in $D$. We define the observation operator

$$G : \ell^m_\kappa \to \mathbb{C}^m, \quad u \mapsto (y_\alpha(z_j))^m_{j=1} \quad (3)$$

where $\ell^m_\kappa$ is a suitable subset chosen from a sequence space of amplitudes $\alpha$ and positions $x$. Later in Section 3.2 we analyse under which assumptions and restrictions $G$ is actually well-defined.

The measurement noise is denoted by $\eta \in \mathbb{C}^m$. We analyse the following inverse problem with a Bayesian approach, see for example (Stuart, 2010):

$$y = G(u) + \eta \in \mathbb{C}^m \quad (4)$$

In particular, we give precise meaning to $\eta$, $y$ and $u$ as random variables. We study (4) with a Bayesian approach which means that we are going to propose a problem specific prior, i.e. a distribution, which models prior knowledge of the distribution of the number, amplitudes and positions of the sound sources. We are then going to deduce the posterior distribution of those quantities, which incorporates our prior knowledge and observations from measurements. For the general principle we refer to (Stuart, 2010) or (Sullivan, 2015) for example.

This work is structured as follows. In Section 2 we give a precise meaning to (1) and prove existence and regularity of solutions to (1). We proceed to approximate (1) by finite elements and prove a priori estimates for the error between the exact acoustic pressure and its finite element approximation. Moreover, we show that the corresponding discretization of observation operator $G_h$ is well-defined and meaningful.

In Section 3, we specify the prior and posterior distribution of the number, amplitudes and positions of the unknown sound sources.

Based on the theoretical stochastic framework we propose a sequential Monte Carlo method in Section 4 to solve (1). In particular, we propose an algorithm to approximate the posterior distribution of the number, the amplitudes and the positions of the sources.

Section 5 is reserved for numerical results which illustrate the theoretical framework.

Before we continue we would like to put this work into perspective. In (Bermúdez et al., 2004), (Pieper et al., 2018), deterministic models are considered to recover number, amplitude and positions of the sound sources from measurements of the acoustic pressure. Both papers rely on techniques of optimal control theory. While the cost functional in (Bermúdez et al., 2004) is smooth, a non-smooth and sparsity promoting cost functional is considered in (Pieper et al., 2018). Moreover, (Bermúdez et al., 2004) includes a discussion of finite element approximations.

Concerning Bayesian inverse problem as (4), prior distributions with similar structures as in this work can be found in computational statistics: The focus in (Green, 1995) is to simulate distributions with state spaces of differing dimensions. In particular, (Green, 1995) considers a measure $\pi$ on measurable space $(Q, \mathcal{R})$ where $Q = \bigcup_{k \in \mathbb{N}} Q_k$, $Q_k \subseteq \mathbb{R}^k$, $\mathcal{R} \subseteq \mathbb{N}$. Examples where such distributions can be observed are Bayesian mixture modelling (Richardson and Green, 1997) or non-linear classification and regression (Denison et al., 2002). In our Bayesian inverse problem (4) we will work with a prior distribution which has its support in $\ell^m_\kappa$. Elements in $\ell^m_\kappa$ with finitely many amplitudes $\alpha$ different from zero can be identified with elements of $Q$. However, a model with $Q$ as in (Green, 1995) in our setting would only allow for a finite number of amplitudes $\alpha$ and positions $x$. This is not the case for $\ell^m_\kappa$. Thus we are in a more general setting.

More application-oriented research on the location of sound sources with a Bayesian approach can be found in (Asano et al., 2013). Here, the Gaussian prior and the finite dimensional observation operator are based on a frequency decomposition of the signals. In particular, the authors use a Markov chain Monte Carlo method and compare the simulated results to real experiments. Due to the different structure of the observation operator in (Asano et al., 2013), the model does not consider the Helmholtz equation or its discretization. Further practical results are the subject of (Nakamura et al., 2009), where a similar problem with a dynamically moving robot is studied.
2 The Helmholtz Equation

2.1 Preliminaries

For convenience of the reader we recap some basic results on Sobolev spaces and some auxiliary results. We also introduce the notation for the rest of the work. Most definitions are not given in full generality, but can be easily generalized. Moreover, we define several solution concepts to the Helmholtz equation in this section.

The absolute value of a complex number is denoted by $|\cdot|$. We write $\| \cdot \|$ for the Euclidean norm in $\mathbb{R}^{d}$ and denote by $\text{dist}(x, y)$ the Euclidean distance of $x$ and $y$. We also define the distances

$$\text{dist}(x, D_{1}) := \inf_{y \in D_{1}} \text{dist}(x, y)$$

and

$$\text{dist}(D_{0}, D_{1}) := \inf_{x \in D_{0}, y \in D_{1}} \text{dist}(x, y).$$

**Definition 1** Let $p \in [1, \infty]$ be given. Then $L^{p}(D)$ denotes the $p$-th Lebesgue space with complex values. For $f, h \in L^{p}(D)$ we call $h$ the $l$-th weak derivative of $f$ if we have

$$\int_{D} f \partial_{l} v \, dx = - \int_{D} h v \, dx$$

for all complex valued functions $v \in C_{c}^{\infty}(D)$ with compact support in $D$. We write $\partial_{l} f := h$. Higher derivatives are defined analogously. For a multi-index $\alpha \in \mathbb{N}_{0}^{d}$ we define

$$\partial^{\alpha} f := \partial_{1}^{\alpha_{1}} \cdots \partial_{d}^{\alpha_{d}} f.$$  

We also write $\nabla f^{T} = (\partial_{x_{1}} f, \ldots, \partial_{x_{d}} f)$, provided all weak derivatives of first order exist.

The notion of a weak derivative and the following definitions can obviously be adapted to subdomains of $D$.

**Definition 2** Let $p \in [1, \infty]$ and $k \in \mathbb{N}_{0}$ be given. We define the $k$-th Sobolev space as a generalization of $C^{k}(D)$ for $k \in \mathbb{N}$:

$$W^{k,p}(D) := \left\{ f \in L^{p}(D) : \partial^{\alpha} f \in L^{p}(D) \quad \text{for any} \quad \alpha \in \mathbb{N}_{0}^{d} \quad \text{with} \quad |\alpha| \leq k \right\}.$$  

For $k = 0$ we write $W^{k,p}(D) := L^{p}(D)$. Another special case is $p = 2$, then we write $H^{k}(D) := W^{k,2}(D)$.

We define the natural norm on $W^{k,p}(D)$ by

$$\| f \|_{W^{k,p}(D)}^{p} := \sum_{\alpha \in \mathbb{N}_{0}^{d}, |\alpha| \leq k} \| \partial^{\alpha} f \|_{L^{p}(D)}^{p}.$$  

For the next theorem we need the notion of a Lipschitz domain. A Lipschitz domain is a domain $D$ in $\mathbb{R}^{d}$ such that locally the boundary of $D$ can be written as the graph of a Lipschitz map. A precise definition is given in [Wloka 1987, Definition 2.4]. There is said to have the $N^{0,1}$ property. We emphasize that any domain appearing in this paper, for example $D$ or $D \setminus B_{\kappa}(x)$ for $x \in D, \kappa \geq 0$, are Lipschitz domains.

**Theorem 1 (Sobolev Embedding Theorem)** For each Lipschitz domain $E \subset \mathbb{R}^{d}$ we have the continuous embedding

$$H^{2}(E) \hookrightarrow C(E).$$

More precisely, there exists a $C > 0$ such that

$$\| f \|_{C(E)} \leq C \| f \|_{H^{2}(E)}, \quad \forall f \in H^{2}(E).$$

**Proof** The result can be found in many works including Sobolev spaces. See for example [Wloka 1987, Theorem 6.2] and note that $d \in \{2, 3\}$.

In general, functions in $H^{k}(D)$ are not well-defined on the boundary, but their boundary values can still be defined in the sense of traces.

**Theorem 2 (Trace Theorem)** There exists a linear and continuous map $\gamma_{tr} : H^{k}(D) \to L^{2}(\Gamma)$ such that for any $f \in H^{1}(D) \cap C(D)$ we have $\gamma_{tr} f = f|_{\Gamma}$.

**Proof** Note that $D$ is a Lipschitz domain and thus [Wloka 1987, Theorem 8.7] applies.

**Definition 3** We define

$$H^{2}(\Gamma_{N}) := \left\{ \gamma_{tr} f|_{\Gamma_{N}} : f \in H^{1}(D) \right\}.$$  

In the following, we introduce different solution concepts for the Helmholtz equation.

**Definition 4** The function $y \in C^{2}(D) \cap C(\overline{D})$ is called classical solution to the Helmholtz equation if it satisfies

$$\begin{cases} -\Delta y - \left( \frac{\xi}{\rho} \right)^{2} y = u & \text{in} \ D, \\ \partial_{\nu} y - i \frac{\xi}{\rho} y = f & \text{on} \ \Gamma_{Z}, \\ \partial_{\nu} y = g & \text{on} \ \Gamma_{N}. \end{cases}$$

**Definition 5** The function $y \in H^{1}(D)$ is called weak solution to the Helmholtz equation if it satisfies

$$\int_{D} \nabla y \cdot \nabla v - \left( \frac{\xi}{\rho} \right)^{2} y v \, dx - i \frac{\xi}{\rho} \int_{\Gamma_{Z}} y \bar{v} \, dS(x) + \int_{\Gamma_{N}} g \bar{v} \, dS(x)$$

for all $v \in H^{1}(D)$. 

$$= \int_{D} u \bar{v} \, dx + \int_{\Gamma_{Z}} f \bar{v} \, dS(x) + \int_{\Gamma_{N}} g \bar{v} \, dS(x).$$
Definition 8 (Source set) The space of sequences $y \in L^2(D)$ is called the source set if it satisfies

$$
\int_D y(-\Delta v) - \left(\frac{\xi}{c}\right)^2 yv \, dx
$$

$$
\int_D u\bar{v} \, dx + \int_{\Gamma_Z} f\bar{v} \, ds(x) + \int_{\Gamma_N} g\bar{v} \, ds(x)
$$

for all $v \in H^2(D)$ with $\partial_v v + i\frac{\xi}{c} v = 0$ on $\Gamma_Z$ and $\partial_v v = 0$ on $\Gamma_N$.

Both notions of weak and very weak solutions are derived from partial integration of smooth functions. In case of a very weak solution of the Helmholtz equation with $u = \sum_{\ell=1}^{\infty} \alpha_\ell \delta_{x_\ell}$, the first integral on the right hand side has to be understood as

$$
\int_D u \bar{v} \, dx = \sum_{\ell=1}^{\infty} \alpha_\ell \bar{v}(x_\ell).
$$

This expression is well-defined for $v \in H^2(D)$ since $H^2(D)$ is embedded into $C(\bar{D})$ (see Theorem 1).

The weak solution to the Helmholtz equation has singularities at sources $x_\ell \in D$ if $u = \sum_{\ell=1}^{\infty} \alpha_\ell \delta_{x_\ell}$ and $\alpha_\ell \neq 0$. This makes it complicated to estimate the norm of the very weak solution on the entire domain $D$. Therefore, we restrict the possible source locations $x$ to a source domain and give measurement points $z \in D$ to a measurement domain and demand a positive distance of the domains to each other.

Definition 7 (Source and measurement domain) Let $D' \subseteq D$ be given. For $\kappa > 0$ the source domain $D_\kappa$ is defined as

$$
D_\kappa := \{ x \in D' \mid \text{dist}(x, \Gamma) > \kappa \}.
$$

The measurement domain is defined as

$$
M_\kappa := \{ x \in D \mid \text{dist}(x, D_\kappa) > \kappa \text{ and } \text{dist}(x, \Gamma) > \kappa \}.
$$

With restriction to these domains we can derive estimates of the norms $\| \cdot \|_{H^2(M_\kappa)}$, $\| \cdot \|_{W^2,\infty(M_\kappa)}$ or point evaluations of the form $y_u(z)$ for $z \in M_\kappa$ with sources $x \in D_\kappa$. We further allow for an infinite amount of sources.

Definition 8 (Source set) The space $\ell^1(\mathbb{C} \times \mathbb{R}^d)$ is the space of sequences $(\alpha_\ell, x_\ell)_{\ell=1}^{\infty} \subseteq \mathbb{C} \times \mathbb{R}^d$ such that

$$
\| (\alpha_\ell, x_\ell)_{\ell=1}^{\infty} \|_{\ell^1(\mathbb{C} \times \mathbb{R}^d)} := \sum_{\ell=1}^{\infty} (|\alpha_\ell| + \|x_\ell\|) < \infty.
$$

We denote this Banach space by $\ell^1$. We define the set $\ell^1_\kappa$ as the restriction of $\ell^1$ to the source domain

$$
\ell^1_\kappa := \{ (\alpha_\ell, x_\ell)_{\ell=1}^{\infty} \mid x_\ell \in D_\kappa, \forall \ell \in \mathbb{N} \}.
$$

Moreover, we introduce the mapping $\tau$ which maps intensities and sources to a series of dirac measures

$$
\tau((\alpha_\ell, x_\ell)_{\ell=1}^{\infty}) := \sum_{\ell=1}^{\infty} \alpha_\ell \delta_{x_\ell}.
$$

The general assumption is now that $D_\kappa$ and $\ell^1_\kappa$ are non-empty. The definition of $\ell^1_\kappa$ requires that $0 \in D_\kappa$, otherwise $\ell^1_\kappa$ is empty. This assumption can be bypassed by an appropriate shift of the domain. Thus we assume w.l.o.g. that $0 \in D_\kappa$. We will show error estimates in the $\| \cdot \|_{\ell^1}$.

In this section we study the very weak solution $y_u$ of the Helmholtz equation. We proceed as in (Bermúdez et al. 2004, Section 3.2) to prove existence, uniqueness and norm estimates with respect to the data. Furthermore, we are interested in sensitivity properties of $y_u$ in $u$. Later, we want $u$ to follow some prior probability distribution. This requires that all appearing constants in this section are independent of $u$. After the results for the infinte dimensional problem have been established, we continue to prove similar results for solutions $y_u^h$ of a discretized problem and show convergence rates.

2.2 Homogeneous Neumann Data and a Single Sound Source

Consider the Helmholtz equation (1) with the right hand side $\tau(u) = \delta_x$ for $u = ((1, x), (0, 0), (0, 0), \ldots) \in \ell^1_\kappa$ and $g = 0$ for some $x \in D_\kappa$.

In the following, we will repeat the main steps from (Bermúdez et al. 2004, section 3.2) to show existence and regularity of the very weak solution. We prove both at the same time using a decomposition argument.

Proposition 1 The Helmholtz equation

$$
\begin{cases}
-\Delta G^x - \left(\frac{\xi}{c}\right)^2 G^x = \delta_x & \text{in } D, \\
\partial_x G^x - i\frac{\xi}{c} G^x = 0 & \text{on } \Gamma_Z, \\
\partial_x G^x = 0 & \text{on } \Gamma_N.
\end{cases}
$$

has a unique very weak solution $G^x$ for any $x \in D_\kappa$. It satisfies

$$
\| G^x \|_{H^2(M_\kappa)} \leq C_\kappa, \\
\| G^x \|_{W^2,\infty(M_\kappa)} \leq C_\kappa.
$$

The constant $C_\kappa > 0$ depends on $\kappa$ but not on $x$. 

Proof. We introduce the fundamental solution of the Helmholtz equation in the whole space, \( -\Delta \Phi^x - \left(\frac{\kappa}{x}\right)^2 \Phi^x = \delta_x \) in \( \mathbb{R}^d \).

This equation has to be understood in the sense of (tempered) distributions in \( \mathcal{S}'(\mathbb{R}^d) \), i.e.

\[
\int_{\mathbb{R}^d} -\Phi^x \Delta \tau - \left(\frac{\kappa}{x}\right)^2 \Phi^x \tau \, dx = \tilde{v}(x) \quad \forall \tau \in \mathcal{S}(\mathbb{R}^d).
\]

We introduce the space \( \mathcal{S}(\mathbb{R}^d) \) as in (Zeidler 1990 pp. 1058-1061). For functions \( v : \mathbb{R}^d \to \mathbb{C} \) we consider the seminorms

\[
p_{k,m}(v) := \sup_{x \in \mathbb{R}^d} (\|x\|^k + 1) \sum_{|\alpha| \leq m} |\partial^\alpha v(x)|,
\]

with \( k, m \in \mathbb{N}_0 \). The space \( \mathcal{S}(\mathbb{R}^d) \) is then defined as the space of all functions \( v \in C^\infty(\mathbb{R}^d, \mathbb{C}) \) such that

\[
p_{k,m}(v) < \infty \quad \text{for all} \quad k, m \in \mathbb{N}_0.
\]

With this definition the solution of (4) is given by

\[
\Phi^x(y) := \begin{cases} 
\frac{1}{2\pi} Y_0 \left(\frac{x}{y}\right) & \text{if } d = 2, \\
\exp \left(\frac{-\kappa |x-y|}{4\pi} \right) & \text{if } d = 3.
\end{cases}
\]

Here, \( Y_0 \) denotes the zero-order second-kind Bessel function (Dautray and Lions 2012 Chapter. 2, §8.6 Proposition 27). We also refer to (Bowman 1958 Chapter 1) for some background on Bessel functions.

In particular, there holds \( \Phi^x \in C^\infty(\mathbb{R} \setminus \{x\}) \) and \( \Phi^x|_D \in L^2(D) \).

We introduce \( p^x \) as the weak solution of

\[
\begin{cases} 
-\Delta p^x - \left(\frac{\kappa}{x}\right)^2 p^x = 0 \quad &\text{in} \quad D, \\
\partial_\nu p^x - \frac{\kappa}{\gamma} p^x = -\partial_\nu \Phi^x + \frac{\kappa}{\gamma} \Phi^x \quad &\text{on} \quad \Gamma_Z, \\
\partial_\nu p^x = -\partial_\nu \Phi^x \quad &\text{on} \quad \Gamma_N.
\end{cases}
\]

According to (Bermúdez et al. 2004 Theorem 3.3) and the remarks before and after (Bermúdez et al. 2004 (3.10)) the solution \( p^x \) exists, is unique and \( p^x \in H^2(D) \).

Moreover, it satisfies

\[
\|p^x\|_{H^2(D)} \leq C \left[ \left\|\frac{\kappa}{\gamma} \partial_\nu \Phi^x - \partial_\nu \Phi^x\right\|_{H^{\frac{1}{2}}(\Gamma_Z)} + \left\|\partial_\nu \Phi^x\right\|_{H^{\frac{1}{2}}(\Gamma_N)} \right].
\]

This implies that \( G^x = p^x + \Phi^x|_D \) is the very weak solution to (3). In particular for any \( \kappa > 0 \) we have

\[
\|G^x\|_{H^2(D \setminus B_\kappa(x))} \leq \|p^x\|_{H^2(D \setminus B_\kappa(x))} + \|\Phi^x\|_{H^2(D \setminus B_\kappa(x))}.
\]

Note that

\[
Y_0(x) \sim \log(x), \quad Y_0'(x) \sim 1/x, \quad \text{and} \quad Y_0''(x) \sim 1/x^2
\]

for \( x > 0, x \to 0 \) (Bowman 1958 Chapter 1). Moreover,

\[
\exp \left(\frac{-\kappa x}{4\pi} \right) \sim 1/x, \quad \left(\frac{\exp \left(\frac{-\kappa x}{4\pi} \right)}{4\pi x}\right)' \sim 1/x^2
\]

and

\[
\left(\frac{\exp \left(\frac{-\kappa x}{4\pi} \right)}{4\pi x}\right)'' \sim 1/x^3
\]

for \( x > 0, x \to 0 \). For \( d = 3 \) we conclude \( \partial^k \Phi^x(y) \leq C|x-y|^{-k-1} \) for \( k \in \{0,1,2\} \). Finally, \( |x-y| \geq \kappa \) for \( y \in D \setminus B_\kappa(x) \) yields

\[
\|\Phi^x\|_{H^2(D \setminus B_\kappa(x))} \leq C \kappa^{-d}
\]

for a constant \( C > 0 \) which is independent of \( \kappa \) and \( x \). Because \( \text{dist}(x, \Gamma) > \kappa \), the right hand side in (8) can be bounded similarly using the same arguments. The proof for \( d = 2 \) is analogous. The estimate in the \( W^{2,\infty}(M) \)-norm follows from (Bermúdez et al. 2004 Lemma 3.5) with \( D_0 := M_\kappa \).

Remark 1 We have to ‘cut out’ discs around \( x \in D \) in Proposition 1 so that the expressions in (9) remain finite. In particular, the fundamental solution \( \Phi^x \in \mathbb{R}^d \) in the proof of Proposition 1 has a singularity at \( x \) and \( \Phi^x \notin H^2(D) \).

Proposition 2 Let \( x \in D_\kappa \) be given. Then for any \( y \in D_\kappa \) with \( |x-y| < \frac{1}{2} \kappa \) there holds

\[
\|G^x - G^y\|_{H^2(M_\kappa)} \leq C_\kappa |x-y|,
\]

with a constant \( C_\kappa \) depending on \( \kappa \) but not on \( x \) and \( y \).

Proof. Observe that

\[
\|G^x - G^y\|_{H^2(M_\kappa)} \leq \|p^x - p^y\|_{H^2(D \setminus B_\kappa(x))} + \|\Phi^x - \Phi^y\|_{H^2(D \setminus B_\kappa(x))},
\]

with \( p^x, p^y, \Phi^x, \Phi^y \) from the proof of Proposition 1. Let \( \kappa \in \mathbb{N}_0 \).

Note that for each \( \xi \) in the line \([x,y]\) we have

\[
\kappa < |x-z| \leq |x-\xi| + |z-\xi| < \frac{1}{2} \kappa + |z-\xi|
\]

and thus

\[
\frac{1}{2} \kappa < |z-\xi|.
\]

By the proof of the previous we have for \( 1 \leq k \in \mathbb{N} \) and \( \xi \in [x,y] \)

\[
\nabla^k \Phi^x(z) \leq C \|\xi - z\|^{2-d-k} \leq C 2^{k+d-2} \kappa^{2-d-k}.
\]
For any $1 \leq k \in \mathbb{N}$ we have by the mean value theorem
\[
|\nabla^k \phi^x(z) - \nabla^k \phi^y(z)| \leq C_{k,k} \|x - y\|.
\]
We conclude
\[
\|\Phi - \Phi\|_{H^2(D),B_n(x)} \leq C_k \|x - y\|.
\]
Using the definition of $p^x$ and equation (8) the expression
\[
\|p^x - p^y\|_{H^2(D),B_n(x)}
\]
can be bounded by
\[
\leq C\left[\frac{|\nabla^k \Phi - \nabla^k \Phi|}{|\nabla^k \Phi - \nabla^k \Phi|_{H^2(D),B_n(x)}} + \|\nabla^k \Phi - \nabla^k \Phi\|_{H^2(D),B_n(x)}\right].
\]
Once again using the mean value theorem and bounding derivatives we get the bound
\[
\|p^x - p^y\|_{H^2(D),B_n(x)} \leq C_k \|x - y\|
\]
which proves the claim.

2.3 Inhomogeneous Neumann Data and Multiple Sound Sources

In this section we present an existence and uniqueness result for the very weak solution of the Helmholtz equation (9) for $g \neq 0$. For a given sequence $u = (\alpha_t, x_t)_{t=1}^{\infty} \in \ell^1_k$ we allow the right hand side to be a series of Diracs given by $\tau(u) = \sum_{t=1}^{\infty} \alpha_t \delta_{x_t}$.

**Theorem 3** Let $u \in \ell^1_k$ and $g \in H^2(\Gamma_N)$ be given. Then the problem
\[
\begin{cases}
-\Delta y_u - \left(\frac{\omega}{c}\right)^2 y_u = \nu u & \text{in } D, \\
\partial_{\nu} y_u - \left(\frac{\omega}{c}\right) y_u = 0 & \text{on } \Gamma_Z, \\
\partial_{\nu} y_u = g & \text{on } \Gamma_N
\end{cases}
\]
is well-defined in the very weak sense. It has a unique very weak solution $y_u \in H^2(M_c) \cap L^2(D)$ which satisfies
\[
\|y_u\|_{H^2(M_c)} \leq C_k \left(\|u\|_{\ell^1} + \|g\|_{H^2(\Gamma_N)}\right)
\]
with a constant $C_k > 0$ depending on $\kappa$ but not on $u$.

**Proof** Let $u = (\alpha_t, x_t)_{t=1}^{\infty} \in \ell^1_k$ and $g \in H^2(\Gamma_N)$ be arbitrary. First we show well-definedness of the very weak formulation of (11). For a suitable testfunction $v \in H^2(D)$ this formulation is given by
\[
\int_D y_u(-\Delta v) - \left(\frac{\omega}{c}\right)^2 y_u v \, dx = \int_D \tau(u)v \, dx + \int_{\Gamma_N} g v \, dS(x).
\]
The only critical term is the one containing the series of Diracs. Since $v \in H^2(D) \subset C(\overline{D})$ by Theorem 1 this term can be bounded by
\[
\int_D \tau(u)v \, dx = \sum_{t=1}^{\infty} \alpha_t \bar{v}(x_t) \leq \sum_{t=1}^{\infty} |\alpha_t| \max_{x \in \overline{D}} |\bar{v}(x)| \leq \|u\|_{\ell^1} \|v\|_{C(\overline{D})} < \infty.
\]
This shows well-definedness of the very weak formulation.

The next step is to show existence, uniqueness and the norm bound of the very weak solution. We use linearity of equation (11) to split up the solution in the form $y_u = y_{u,0} + y_{0,g}$ where $y_{u,0}$ and $y_{0,g}$ shall satisfy
\[
\begin{cases}
-\Delta y_{u,0} - \left(\frac{\omega}{c}\right)^2 y_{u,0} = \nu u & \text{in } D, \\
\partial_{\nu} y_{u,0} - \left(\frac{\omega}{c}\right) y_{u,0} = 0 & \text{on } \Gamma_Z, \\
\partial_{\nu} y_{u,0} = g & \text{on } \Gamma_N.
\end{cases}
\]
and
\[
\begin{cases}
-\Delta y_{0,g} - \left(\frac{\omega}{c}\right)^2 y_{0,g} = 0 & \text{in } D, \\
\partial_{\nu} y_{0,g} - \left(\frac{\omega}{c}\right) y_{0,g} = 0 & \text{on } \Gamma_Z, \\
\partial_{\nu} y_{0,g} = g & \text{on } \Gamma_N.
\end{cases}
\]

The unique very weak solution of the first problem is given by $y_{u,0} = \sum_{t=1}^{\infty} \alpha_t G^{x_t}$ according to Proposition 3. We apply the bound in the same proposition and obtain a bound of the very weak solution of the form
\[
\|y_{u,0}\|_{H^2(M_c)} \leq \sum_{t=1}^{\infty} |\alpha_t| |G^{x_t}|_{H^2(M_c)} \leq C_k \sum_{t=1}^{\infty} |\alpha_t| \leq C_k \|u\|_{\ell^1},
\]
where the constant $C_k > 0$ does not depend on $u$.

We apply (Bermúdez et al., 2004, Theorem 3.3) to obtain existence and uniqueness of a weak solution $y_{0,g}$ for the second equation of (12), which satisfies
\[
\|y_{0,g}\|_{H^2(D)} \leq C |g|_{H^\frac{1}{2}(\Gamma_N)}.
\]
We combine the results for $y_{u,0}$ and $y_{0,g}$ to conclude the theorem.

**Theorem 4** Let $g \in H^\frac{1}{2}(\Gamma_N)$ and $u, v \in \ell^1_k$ be given such that $\|u - v\|_{\ell^1} < \frac{1}{2} \kappa$. Let $y_u$ and $y_v$ be the very weak solutions of the Helmholtz equation (9) with the same Neumann data $g$ but with different right hand sides $\tau(u)$ respectively $\tau(v)$. Then the solutions satisfy
\[
\|y_u - y_v\|_{H^2(M_c)} \leq C_k (\|u\|_{\ell^1} + 1) \|u - v\|_{\ell^1},
\]
where $C_k > 0$ is independent of $u$ and $v$. In particular $y(\cdot) : \ell^1_k \to H^2(M_c)$ is a continuous function.
Proof Let \( u = (\alpha_\ell, x_\ell)_{\ell=1}^\infty \in \ell_1^\infty \) and \( v = (\beta_\ell, y_\ell)_{\ell=1}^\infty \in \ell_1^\infty \) be given with \( \|u - v\|_1 < \frac{1}{2}\kappa \). We apply decomposition \(^{[12]}\) to show that the difference of the solutions satisfies
\[
y_u - y_v = y_{u,0} - y_{v,0}.
\]
This allows us to assume \( g = 0 \).

A straightforward computation shows
\[
\|y_u - y_v\|_{H^2(M_h)} \leq \sum_{\ell=1}^\infty \|\alpha_\ell G^{x_\ell} - \beta_\ell G^{y_\ell}\|_{H^2(M_h)}
\]
and
\[
\|\beta_\ell\|_{L^2(M_h)} \leq \sum_{\ell=1}^\infty \|\alpha_\ell - \beta_\ell\|_{L^2(M_h)}
\]

Applying these inequalities and Proposition \(^{[1]}\) together with Proposition \(^{[2]}\) for every \( \ell \in \mathbb{N} \) we further estimate
\[
\|y_u - y_v\|_{H^2(M_h)} \leq C_\kappa \sum_{\ell=1}^\infty \left( |\alpha_\ell - \beta_\ell| + |\alpha_\ell + \frac{1}{2}\kappa| \right) \|x_\ell - y_\ell\|
\]
\[
\leq C_\kappa \max_{\ell, \in \mathbb{N}} \left( 1, |\alpha_\ell| + \frac{1}{2}\kappa \right) \sum_{\ell=1}^\infty (|\alpha_\ell - \beta_\ell| + \|x_\ell - y_\ell\|)
\]
\[
\leq C_\kappa \max_{\ell, \in \mathbb{N}} (|\alpha_\ell| + 1) \|u - v\|_{L^1},
\]
which proves the claim.

2.4 Finite Element Spaces

We follow \(^{[Bermúdez et al., 2004]}\) and discretize the Helmholtz equation \(^{[1]}\) by piecewise linear finite elements. To this we need consider a family of triangulations \( (T_h)_{h>0} \) of \( D \). More precisely, for each \( h > 0 \) the set \( T_h \) consists of closed triangles/tetrahedrons \( T \subset \mathbb{R}^d \) such that
\[
\bar{D} = \bigcup_{T \in T_h} T \quad \text{for any } h > 0.
\]
The corners of the triangles are called nodes of \( D \). We assume that no hanging nodes or hanging edges exist. More precisely, we suppose that for all \( h > 0 \) and any \( T, T' \in T_h \) the intersection \( T \cap T' \) is either empty, a single point, a common edge of \( T \) and \( T' \) or a common facet of \( T \) and \( T' \). The last case is only relevant for \( d = 3 \). The domain \( D \) is polygonal/polyhedral which implies the existence of such a triangulation. For each triangulation \( T_h \) we write \( h_T \) for the diameter of a triangle \( T \in T_h \). The mesh size of \( T_h \) is given by \( h := \max_{T \in T_h} h_T \). We denote by \( \rho_T \) the diameter of the largest ball contained in \( T \in T_h \). We make the following assumption for the remaining part of the paper:
\[
\exists \sigma_1, \sigma_2 > 0 : \frac{h_T}{\rho_T} \leq \sigma_1, \quad \frac{h}{\rho_T} \leq \sigma_2 \quad \forall T \in T_h, \forall h > 0.
\]
This or similar conditions are called shape regularity and quasi-uniformity, as this for example prevents our triangles from becoming too acute or too flat. We note that these conditions imply that there exists a constant \( C > 0 \) such that
\[
\max_{T \in T_h} \rho_T \geq \frac{h}{\sigma_1 \sigma_2}.
\]
We associate with each triangulation \( T_h \) the finite element space \( \tilde{V}_h \) which consists of real valued functions globally continuous in \( D \) and linear on each element \( T \in T_h \):
\[
\tilde{V}_h := \left\{ v_h \in H^1(D, \mathbb{R}) \cap C(\bar{D}, \mathbb{R}) : \right\}
\]
\[
v_h|_T \text{ is affine linear } \forall T \in T_h \}.
\]
The space \( V_h \) is defined by all functions \( v_h \) of the form \( v_h = v_{h,1} + iv_{h,2} \) with \( v_{h,i} \in \tilde{V}_h \).

2.5 Galerkin Approximations

Definition 9 (Discrete solution) The function \( y_h \in V_h \) is called discrete solution, or Galerkin Approximation, to the Helmholtz equation
\[
\begin{cases}
-\Delta y_h - \left( \frac{\omega^2}{c^2} \right) y_h = \tau(u) & \text{in } D, \\
\partial_n y_h - i \frac{\omega}{c} \gamma \nu \cdot y_h = 0 & \text{on } \Gamma_Z, \\
\partial_n y_h = g & \text{on } \Gamma_N,
\end{cases}
\]
if it satisfies
\[
\int_D \nabla y_h \nabla \bar{v}_h - \left( \frac{\omega^2}{c^2} \right) y_h \bar{v}_h \, dx - \frac{i\omega}{\gamma} \int_{\Gamma_Z} y_h \bar{v}_h dS(x)
\]
\[
= \int_D \tau(u) \bar{v}_h \, dx + \int_{\Gamma_N} g \bar{v}_h dS(x)
\]
(14)
for all \( v_h \in V_h \). We prove existence and uniqueness of solutions to (14) in Theorem 5.

The following proposition allows us to work with \( L^2(D) \) functions instead of diracs in the discrete setting. For such a construction also see \(^{[Scott, 1973/74]}\) Theorem 1 and after Lemma 3.
Proposition 3 Let $x \in \mathcal{D}$. There is a $\delta_{x,h} \in V_h$ such that
\[ \int_D \delta_{x,h} \bar{v}_h \, dx = \bar{v}_h(x), \quad \text{for all } v_h \in V_h. \] (15)
Furthermore we have $\|\delta_{x,h}\|_{L^2(D)} \leq C$ where $C$ does not depend on $x$, but on $h$. The mapping $x \mapsto \delta_{x,h}$ from $\mathcal{D}$ to $V_h$ is Lipschitz continuous.

Proof Let $x \in \mathcal{D}$ and define the functional $j_x$ as follows
\[ j_x : V_h \to \mathbb{C}, \quad j_x(v_h) = \bar{v}_h(x). \]
Because $V_h$ is finite dimensional it is a Banach space together with the norm $\|\cdot\|_{L^2(D)}$. Thus the Riesz representation theorem implies the existence of a $\delta_{x,h} \in V_h$ so that
\[ \bar{v}_h(x) = \int_D \delta_{x,h} \bar{v}_h \, dx \quad \forall v_h \in V_h. \]
By construction and the equivalence of norms on $V_h$ we have
\[ \|\delta_{x,h}\|_{L^2(D)} = |\delta_{x,h}(x)| \leq \|\delta_{x,h}\|_{L^\infty(D)} \leq C \|\delta_{x,h}\|_{L^2(D)}. \]
Note that this $C$ depends on $h$, but not on $x$.

For the Lipschitz continuity let $x_1, x_2 \in \mathcal{D}$. We have that $\delta_{x_1,h}, \delta_{x_2,h} \in V_h$ and both are Lipschitz continuous with constants $\|\nabla \delta_{x_1,h}\|_{L^\infty(D)}$ and $\|\nabla \delta_{x_2,h}\|_{L^\infty(D)}$, see for example (Ambrosio et al. 2000) Proposition 2.13. This yields
\[ \left| \int_D (\delta_{x_1,h} - \delta_{x_2,h})(\delta_{x_1,h} - \delta_{x_2,h}) \, dx \right| \leq |\delta_{x_1,h}(x_1) - \delta_{x_1,h}(x_2)| + |\delta_{x_2,h}(x_1) - \delta_{x_2,h}(x_2)| \leq (\|\nabla \delta_{x_1,h}\|_{L^\infty(D)} + \|\nabla \delta_{x_2,h}\|_{L^\infty(D)}) |x_1 - x_2|. \]
Again using the equivalency of norms we can use the bound from before $\|\nabla \delta_{x_1,h}\|_{L^\infty(D)} \leq C \|\delta_{x_1,h}\| \leq C$, where $C$ does not depend on $x$, but on $h$. The same estimate holds for $\delta_{x_2,h}$ concluding the proof.

Proposition 4 There exists a $h_0 > 0$ such that for all $h \in (0, h_0]$, $f \in L^2(D)$ and $g \in H^\frac{1}{2}(\Gamma_N)$ there exists a unique discrete solution $y_{f,h} \in V_h$ of
\[ \int_D \nabla y_{f,h} \nabla \bar{v}_h - \left( \frac{\xi}{c} \right)^2 y_{f,h} \bar{v}_h \, dx + \frac{i \xi \rho}{\gamma \zeta} \int_{\Gamma_N} y_{f,h} \bar{v}_h \, dS = \int_D f \bar{v}_h \, dx + \int_{\Gamma_N} g \bar{v}_h \, dS \] (16)
for all $v_h \in V_h$. Furthermore there holds
\[ \|y_{f,h}\|_{L^2(D)} \leq C \left( \|f\|_{L^2(D)} + \|g\|_{H^\frac{1}{2}(\Gamma_N)} \right) \]
with $C$ independent of $f$ and $g$.

Proof Let $f \in L^2(D), g \in H^\frac{1}{2}(\Gamma_N)$ be given. Then
\[ \begin{cases} -\Delta y - \left( \frac{\xi}{c} \right)^2 y = f \text{ in } D, \\ \partial_y y - i \frac{\xi \rho}{\gamma \zeta} y = 0 \text{ on } \Gamma_z, \\ \partial_y y = g \text{ on } \Gamma_N \end{cases} 
has a unique weak solution $y_f$ by (Bermúdez et al. 2004 Theorem 3.3). It satisfies
\[ \|y_f\|_{H^2(D)} \leq C \left( \|f\|_{L^2(D)} + \|g\|_{H^\frac{1}{2}(\Gamma_N)} \right) \]
with a constant $C$ independent of $f$ or $g$. We introduce the sesquilinear form
\[ a(y_f, y_f, v_h) := \int_D \nabla (y_f - y_{f,h}) \nabla \overline{v}_h - \left( \frac{\xi}{c} \right)^2 (y_f - y_{f,h}) \overline{v}_h \, dx - \frac{i \xi \rho}{\gamma \zeta} \int_{\Gamma_N} (y_f - y_{f,h}) \overline{v}_h \, dS. \]

$\alpha$ satisfies Gårding’s inequality by (Bermúdez et al. 2004 Lemma 3.2), which means that there exist $C_1, C_2 > 0$ such that
\[ |a(q,q) + C_2 \|q\|^2_{L^2(D)} | \geq C_1 \|q\|^2_{H^1(D)} \quad \forall q \in H^1(D). \]

Now one can simply use the same ideas used in the real case in (Schatz 1974).

A direct consequence of Proposition 4 is the existence and uniqueness of a discrete Green’s function $G_h^f \in V_h$.

Corollary 1 There exists a $h_0 > 0$ such that for all $h \in (0, h_0]$ and all $x \in D_h$ the discrete formulation
\[ \int_D \nabla G_{h}^f \nabla \bar{v}_h - \left( \frac{\xi}{c} \right)^2 G_{h}^f \bar{v}_h \, dx - \frac{i \xi \rho}{\gamma \zeta} \int_{\Gamma_N} G_{h}^f \bar{v}_h \, dS \]
\[ = \int_D \delta_{x,h} \bar{v}_h \, dx, \quad \forall v_h \in V_h, \]
has a unique discrete solution $G_{h}^f \in V_h$ with $\|G_{h}^f\|^2_{L^2(D)} \leq C$ independent of $x$.

Remark 2 The existence of $G_{h}^f$ strongly depends on the mesh size $h$. The upper bound $h_0$ for the mesh size depends on the coefficients of Gårding’s inequality used in the existence proof (above equation (7) in (Schatz 1974)). In (Bermúdez et al. 2004 Lemma 3.2) those coefficients are stated explicitly. The bound $h_0$ is proportional to $(1 + \frac{\xi^2}{c^2})^{-1}$ which guarantees existence of $G_{h}^f$. Hence, for a large ratios $\frac{\xi}{c}$ one has to work with fine meshes to compute reasonable solutions.
Theorem 5 Let \( g \in H^{2 \frac{1}{2}}(\Gamma_N) \) and \( u = (\alpha_t, x_t)_{t=1}^N \in \ell_2^2 \) be given. Then there exists a \( h_0 > 0 \) such that for all \( h \in (0, h_0] \) the Helmholtz equation has a unique discrete solution \( u_{h,v} \in V_h \), that is for all \( v \in V_h \) it satisfies

\[
\int_D \nabla u_{h,v} \nabla \nu_h \, dx - \left( \frac{\zeta}{c} \right)^2 u_{h,v} \nu_h \, dx = \int_D \tau(u) \nu_h \, dx + \int_{\Gamma_N} g \nu_h \, dS(x).
\]

Furthermore, the solution is given by

\[
y_{u,h} = \sum_{\ell=1}^{\infty} \alpha_\ell G^x_{h,\ell} + y_{g,h},
\]

where \( y_{g,h} \) is the discrete solution of the Helmholtz equation with zero forcing term and Neumann data \( g \).

Proof By Lemma 1 and Proposition 4 it is easy to see that

\[
y_{u,h} = \sum_{\ell=1}^{\infty} \alpha_\ell G^x_{h,\ell} + y_{g,h}
\]

exists and is well defined as the series converges in \( L^2(D) \) by \( \sum_{\ell=1}^{\infty} |\alpha_\ell| < \infty \) and \( \|GG^x_{h,\ell}\|_{L^2(D)} \leq C \) with a \( C \) independent of \( x \). Linearity of the integral now shows that is indeed a solution to (18).

2.6 Pointwise error estimates

The proofs in this section follow the proofs in [Bermúdez et al., 2004, Section 4]. The authors explicitly state though that they do not analyse the dependence of the appearing constants with respect to \( \|x - z\| \), where \( x \in D_\kappa \) denotes the position of a Dirac and \( z \in M_\kappa \) an arbitrary measurement point. First we restate a slight variant of [Schatz and Wahlbin, 1977, Corollary 5.1].

Proposition 5 Let \( D_1 \subset \subset D_2 \subset \subset D \) be given. Moreover, let \( f \in C(D_2, \mathbb{R}) \cap H^1(D_2, \mathbb{R}) \) and \( f_h \in \tilde{V}_h \) satisfy

\[
(\nabla (f - f_h), \nabla v_h)_{L^2(D_2)} - \left( \frac{\zeta}{c} \right)^2 (f - f_h, v_h)_{L^2(D_2)} = 0
\]

(19)

for all \( v_h \in \tilde{V}_h \) with supp \( v_h \subset D_2 \). Then there exist constants \( C, C' > 0 \) such that if dist \((D_1, \partial D_2) \geq \kappa, C'h \leq \kappa \) and dist \((D_2, \partial D) \geq \kappa \), then for \( 0 \leq l \leq 2 \)

\[
\|f - f_h\|_{L^\infty(D_1)} \leq C \left( \|\ln h\|_{H^2} \|f\|_{W^{2,\infty}(D_2)} + \kappa^{-1} \|f - f_h\|_{L^2(D_2)} \right).
\]

The constants \( C, C' \) do not depend on \( h, f, f_h, D_1 \) and \( D_2 \).

Proof All the assumptions in [Schatz and Wahlbin, 1977, Theorem 5.1, Corollary 5.1] are satisfied by the remarks following [Schatz and Wahlbin, 1977, A.4].

[Schatz and Wahlbin, 1977, Corollary 5.1] requires (19) to hold on \( D \). But inspecting the proof of [Schatz and Wahlbin, 1977, Corollary 5.1] and the application of [Schatz and Wahlbin, 1977, Theorem 5.1] therein shows that (19) is sufficient.

Finally, we choose \( p = 0, q = 1 \) and have \( r = 2 \) in the statement of [Schatz and Wahlbin, 1977, Corollary 5.1].

Proposition 6 Let \( z \in M_\kappa \). Then there exists \( h_0 > 0 \) such that for any \( h \in (0, h_0] \) there holds

\[
|G^x(z) - G^x_h(z)| \leq C_n |\ln h|^{2}, \quad \forall x \in D_\kappa,
\]

where \( h_0 \) and \( C_n \) do not depend on \( x \) or \( z \).

Proof Let \( z \in M_\kappa \) and let \( x \in D_\kappa \). Lemma 1 shows that there exists a \( h_0 \) independent of \( z \) such that \( G^x_h \) exists. We want to apply Proposition 5 to the real and imaginary part of \( G^x_h - G^x_h \in C(D_\kappa \cap H^1(D_\kappa)) \) to obtain

\[
|G^x(z) - G^x_h(z)| \leq \|G^x - G^x_h\|_{L^\infty(B_{\frac{1}{4}\kappa}(z))} \leq C \left( \|\ln h\|_{H^2} \|f\|_{W^{2,\infty}(B_{\frac{1}{4}\kappa}(z))} + \|G^x - G^x_h\|_{L^1(B_{\frac{1}{4}\kappa}(z))} \right).
\]

(20)

Here \( C \) does not depend on the used balls or in the particular \( z \), but only on \( \kappa \). Therefore we choose \( D_1 = B_{\frac{1}{4}\kappa}(z) \) and \( D_2 = B_{\frac{1}{4}\kappa}(z) \) and apply Proposition 1 to obtain

\[
\text{Re}(G^x), \text{Im}(G^x) \in C(B_{\frac{1}{4}\kappa}(z), \mathbb{R}) \cap H^1(B_{\frac{1}{4}\kappa}(z), \mathbb{R}).
\]

Linearity of real part and imaginary part in the weak formulation yields, after short computation, that for any \( v_h \in \tilde{V}_h \) with supp \( v_h \subset D_2 \) there holds

\[
(\nabla(\text{Re}(G^x - G^x_h)), \nabla v_h)_{L^2(D_2)} - \left( \frac{\zeta}{c} \right)^2 (\text{Re}(G^x - G^x_h), v_h)_{L^2(D_2)} = v_h(x) - v_h(x) = 0,
\]

and

\[
(\nabla(\text{Im}(G^x - G^x_h)), \nabla v_h)_{L^2(D_2)} - \left( \frac{\zeta}{c} \right)^2 (\text{Im}(G^x - G^x_h), v_h)_{L^2(D_2)} = 0 - 0 = 0.
\]

This proves (20) according to Proposition 5. Again by Proposition 1 we obtain

\[
\|G^x\|_{W^{2,\infty}(B_{\frac{1}{4}\kappa}(z))} \leq C_n,
\]

(21)
with $C_\kappa$ depending on $\kappa$ but not on $x$ or $z$. This implies
\[ |G^x(z) - G^x_h(z)| \leq C \left( |\ln h| h^2 + \|G^x - G^x_h\|_{L^1(B_{\bar{M}_\kappa}(z))} \right). \]

$C$ does not depend on $x$ or $z.$ $\|G^x - G^x_h\|_{L^1(B_{\bar{M}_\kappa}(z))}$ is estimated as in the proof of (Schatz and Wahlbin, 1977, Theorem 5.1) in the proof of Proposition 7 (Bermúdez et al., 2004, Theorem 4.1), which is based on the proof of (Schatz and Wahlbin, 1977, Theorem 6.1) and (Bermúdez et al., 2004, Lemma 3.5). Tracking the constants in both proofs shows the independence of $x$ and $z.$

**Theorem 6** Let $z \in M_\kappa,$ $u \in \ell^1_\kappa$ and $g \in H^1((\Gamma_N).$
Then there exists a $h_0 > 0$ such that for all $h \in (0, h_0]$ the discrete solution $y_{u,h} \in V_h$ satisfies
\[ |y_{u,z} - y_{u,h}(z)| \leq C_\kappa \ln h h^2 \left( \|u\|_{\ell^1} + \|g\|_{H^2((\Gamma_N)} \right) \]
with a constant $C_\kappa > 0$ independent of $u, h$ and $z.$

**Proof** We are allowed to evaluate $y_u$ and $y_{u,h}$ at $z$ because $y_u \in H^2(M_\kappa) \subset C(M_\kappa)$ and $y_{u,h} \in V_h$. For the error estimate observe
\[
|y_{u,z} - y_{u,h}(z)| \leq \sum_{i=1}^\infty |\alpha_i| |G^{x_\iota_i} - G^{x_\iota_i}_h(z)|
+ |g_{x_\iota_i}(y_u) - g_{x_\iota_i}(y_{u,h})|
\leq \|u\|_{\ell^1} \sup_{x \in \bar{M}_\kappa} |G^x - G^x_h(z)|
+ |g(y_u) - g_{x_\iota_i}(y_{u,h})|.
\]

We apply Lemma 3 to bound the term under the supremum uniformly in $x$ and obtain the desired rate of $|\ln h| h^2$ for the first term. (Bermúdez et al., 2004, Theorem 4.2) gives the estimate for the second term. It is stated in such a way that the minimal mesh size may depend on $g,$ but checking its proof and the reference to (Schatz and Wahlbin, 1977, Theorem 5.1) in the proof of (Bermúdez et al., 2004, Theorem 4.2) shows that it is independent of $g.$

**Proposition 7** Let $h_0 > 0$ from Lemma 1. The mapping $x \mapsto G^x_h$ from $D_u$ to $V_h$ is Lipschitz continuous.

**Proof** Let $h_0 > 0$ be from Lemma 1 and $h \in (0, h_0].$ For $x_1, x_2 \in D_u$ let $G^x_{x_1,h} - G^x_{x_2,h} \in V_h$ be the discrete solution of the Helmholtz equation with right hand side $f = \delta_{x_1,h} - \delta_{x_2,h}$ and Neumann data $g = 0.$ We use Proposition 4 to get
\[ \|G^{x_1} - G^{x_2}_h\|_{L^2(D)} \leq C \|\delta_{x_1,h} - \delta_{x_2,h}\|_{L^2(D)}. \]
with $C$ independent of $x_1$ and $x_2.$ Proposition 3 now shows the result.

**Corollary 2** Let $z \in M_\kappa.$ Then there exists a $h_0 > 0$ such that for $h \in (0, h_0]$ and $g \in H^1((\Gamma_N)$ the solution mapping $y_{u,h}(z) : \ell^1_{\kappa} \to C$ is continuous. More precisely there holds
\[ |y_{u,h}(z) - y_{u,h}(z)| \leq C_\kappa (\|u\|_{\ell^1} + 1) \|u - v\|_{\ell^1} \]
for $u, v \in \ell^1_{\kappa}$ with $\|u - v\|_{\ell^1} < \frac{1}{2} \kappa.$

**Proof** For $x_1, x_2 \in D_u$ we use the equivalence of norms on $V_h$ and Proposition 7 to show
\[ \|G^{x_1}_h - G^{x_2}_h\|_{L^2(D)} \leq C \|G^{x_1} - G^{x_2}_h\|_{C(\bar{M})} \]
\[ \leq C \|G^{x_1} - G^{x_2}_h\|_{L^2(D)} \leq C \|x_1 - x_2\|_{\ell^1_{\kappa}}. \]

Lemma 1 shows that $\|G^{x_1}_h\|_{L^2(D)} \leq C$ with $C$ independent of $x_1.$ Hence a similar computation as in the proof of Corollary 1 replacing the Greens function with its discrete counterpart and replacing $\|\cdot\|_{H^2(M_\kappa)}$ with $\|\cdot\|_{\ell^1_{\kappa}}$ shows the result.

### 3 The Inverse Problem

Up to this point all definitions and all the analysis was carried out in a deterministic setting. Now let us recall problem 7. Our aim is to analyse the following inverse problem with a Bayesian approach:
\[ y = G(u) + \eta. \]

The observed data $y \in C^m$ is fixed. We also recall the definition of the observation operator, namely
\[ G : \ell^1_{\kappa} \to C^m, \quad u := (\alpha_\iota, x_\iota)_{\iota=1}^m \mapsto (y_z)_{\iota=1}^m. \]

$G$ is well defined by Corollary 2 and the embedding $H^2(M_\kappa) \subset C(M_\kappa).$ In the Bayesian setting we assume that $u$ follows some prior distribution $\mu^h,$ which we will introduce in Section 3.1. We then want to incorporate the knowledge from the measured data $y \in C^m$ to obtain the posterior distribution $\mu^\eta.$ Let us assume that the measurement noise $\eta$ is multivariate complex normal distributed such that $\eta$ has a probability density proportional to
\[ p_\eta(z) \propto \exp \left( -\frac{1}{2} \|z\|_{\Sigma}^2 \right), \quad \forall z \in C^m, \]
where $\|\cdot\|_{\Sigma}$ is a norm on $C$ induced by a scalar product.

Here, $\Sigma \in C^{2m \times 2m}$ is a Hermitian, positive definite complex matrix of the form $\Sigma = \begin{pmatrix} \Gamma & \bar{C} \\ \bar{C} & f \end{pmatrix}.$ In particular, $\eta$ is multivariate complex normal distributed with expectation value zero, positive definite variance matrix.
\( \Gamma \in \mathbb{C}^{m \times m} \) and Hermitian relation matrix \( C \in \mathbb{C}^{m \times m} \). Moreover, \( \| z \|^2_\Sigma \) is defined as:

\[
\| z \|^2_\Sigma := (z^T, z^T) \cdot \Sigma^{-1} \cdot (\frac{\tilde{z}}{\tilde{z}}).
\]

Accordingly, the random vector \( \begin{pmatrix} \text{Re}(\eta) \\ \text{Im}(\eta) \end{pmatrix} \) is multivariate normal distributed in \( \mathbb{R}^{2m} \) with mean zero and covariance matrix

\[
\begin{pmatrix}
\frac{1}{2} \text{Re}[(\Gamma + C) + \frac{1}{2} \text{Im}[-(\Gamma + C)]] \\
\frac{1}{2} \text{Im}[(\Gamma + C) + \frac{1}{2} \text{Re}[-(\Gamma + C)]]
\end{pmatrix}.
\]

Conversely, \( \Gamma \) and \( C \) can be obtained from the covariance matrices \( \frac{1}{2} \text{Re}[(\Gamma + C)], \frac{1}{2} \text{Im}[-(\Gamma + C)] \) and \( \frac{1}{2} \text{Re}[-(\Gamma + C)] \) by

\[
\begin{align*}
\Gamma &= \frac{1}{2} \text{Re}[(\Gamma + C)] + \frac{1}{2} \text{Re}[-(\Gamma + C)] \\
&\quad + i \left( \frac{1}{2} \text{Im}[-(\Gamma + C)] - \frac{1}{2} \text{Im}[(\Gamma + C)] \right), \\
C &= \frac{1}{2} \text{Re}[(\Gamma + C)] - \frac{1}{2} \text{Re}[-(\Gamma + C)] \\
&\quad + i \left( \frac{1}{2} \text{Im}[-(\Gamma + C)] + \frac{1}{2} \text{Im}[(\Gamma + C)] \right),
\end{align*}
\]

see [Halliwell 2015 Sections 1-7].

**Example 1** Assume that observation noise is defined as \( \eta := \eta_1 + i \cdot \eta_2 \) with independent \( \eta_j \sim N(0, \Sigma_j) \), and symmetric positive definite covariance matrices \( \Sigma_j \in \mathbb{R}^{m \times m} \) for \( j = 1, 2 \). Then the noise \( \eta \) is complex normal distributed with positive definite covariance matrix \( \Gamma = \Sigma_1 + \Sigma_2 \) and Hermitian relation matrix \( C = \Sigma_1 - \Sigma_2 \). In this case \( \Sigma \) is real valued.

We will shown in Section 3.3 that our particular noise assumption implies that the posterior distribution is given by

\[
\mu^\nu(F) = \frac{1}{A} \int_F \exp \left( -\frac{1}{2} \| y - G(u) \|^2_\Sigma \right) d\mu^0(u) \tag{23}
\]

for \( F \in \mathcal{F} \), where \( \mathcal{F} \) is a Sigma algebra on \( \ell^1_k \). The constant \( A \) is a normalization constant given by

\[
A = \int_{\ell^1_k} \exp \left( -\frac{1}{2} \| y - G(u) \|^2_\Sigma \right) d\mu^0(u).
\]

The term \( \exp \left( -\frac{1}{2} \| y - G(u) \|^2_\Sigma \right) \) can be viewed as penalization for deviating too far away from the observed data. Thus the posterior combines both prior knowledge of the solution and the measurements favouring those \( u \in \ell^1_k \) which closely predict the observation.

Before we continue we recall some basic definitions and tools.

**Definition 10 (Radon-Nikodym Derivative)** Let \( (X, \mathcal{F}, \nu) \) be a measure space. Let \( f : X \to \mathbb{R}_{\geq 0} \) be a \( \nu \)-measurable function. Define the measure \( \mu \) on \( X \) by

\[
\mu(F) = \int_F f(x) d\nu(x), \quad \forall F \in \mathcal{F}.
\]

Then \( f \) is called Radon-Nikodym derivative of \( \mu \) with respect to \( \nu \) and is denoted by

\[
d\mu = f d\nu.
\]

If two measures \( \mu^1 \) and \( \mu^2 \) have a Radon-Nikodym derivative with respect to a third measure \( \nu \), then we are able to define a metric between them, see [Stuart 2010 Section 6.7].

**Definition 11 (Hellinger Distance)** Let \( \mu^1, \mu^2, \nu \) be measures on \( X \) such that \( \mu^1, \mu^2 \) have a Radon-Nikodym Derivative with respect to \( \nu \). Then the Hellinger distance between \( \mu^1 \) and \( \mu^2 \) is defined as

\[
d_{\text{Hell}}(\mu^1, \mu^2) = \left( \int_X \left( \frac{1}{2} \left( \frac{d\mu^1}{d\nu} \right)^\frac{1}{2} - \left( \frac{d\mu^2}{d\nu} \right)^\frac{1}{2} \right)^2 d\nu \right)^\frac{1}{2}.
\]

(24)

The Hellinger distance tells us how well two measures agree. Using Gaussian noise we obtain exponential terms in the Radon-Nikodym derivatives. In order to estimate those, we require an simple lemma, which immediately follows from the mean value theorem.

**Lemma 1** Let \( a, b, c \geq 0 \). Then the following bound holds:

\[
|\exp (-ab) - \exp (-ac)| \leq a|b - c|.
\]

### 3.1 The Prior

Recall that the Bayesian approach needs prior knowledge in form of a probability distribution which fits to problem [1]. As we want to recover the number of sound sources, their amplitudes, and their positions in \( \mathbb{D} \) we require a prior distribution which is suits these requirements. Let \( (\ell^1_k, \mathcal{F}) \) be the measurable set where \( \mathcal{F} \) is the Borel \( \sigma \)-algebra associated with the open sets generated by the norm \( \| \cdot \|_\ell^1 \) on the open set \( \ell^1_k \). For \( n \in \mathbb{N}_0 \) we define the sets

\[
\ell^1_{n,k} := \{ (\alpha_k, x_k)_{k=1}^\infty \in \ell^1_k | \alpha_k = 0, x_k = 0, \forall k > n \} \subseteq \ell^1.
\]

Using this notation we are able to construct a specific probability measure on \( (\ell^1_k, \mathcal{F}) \).
Theorem 7 Let $q$ be a probability mass function on $(\mathbb{N}_0, 2^{\mathbb{N}_0})$ and for every $n \in \mathbb{N}_0$ let $\mu_n^0$ be a probability measure on $(\ell^1_{\kappa}, \mathcal{F})$ such that $\mu_n^0(\ell^1_{n,k}) = 1$. Then $\mu^0$ defined by

$$\mu^0(F) = \sum_{n \in \mathbb{N}_0} q(n)\mu_n^0(F), \quad \text{for all } F \in \mathcal{F},$$

is a well-defined probability measure on $(\ell^1_{\kappa}, \mathcal{F})$. This probability measure satisfies

$$\mu^0(\{u \in \ell^1_{\kappa} \mid u \in \ell^1_{n,k} \text{ for some } n\}) = 1. \quad (25)$$

This implies that samples from $\mu^0$ are $\mu^0$-almost surely in $\ell^1_{n,k}$ for some $n$.

Proof The proof is straight forward.

The motivation to choose such a measure follows from property $(25)$. This ensures that the forcing term for $[\ ]$ consists $\mu^0$-almost surely of a finite number of Diracs, that is

$$\tau(u) = \tau((\alpha_{\ell}, x_{\ell})_{\ell=1}^{\infty}) = \sum_{\ell=1}^{n} \alpha_{\ell}\delta_{x_{\ell}},$$

because $u \in \ell^1_{n,k}$ for some $n \in \mathbb{N}_0$. A practical approach to construct such a measure is now to express $u$ in terms of random variables $n, \alpha, x$ such that the Diracs on the right hand side of the Helmholtz equation are given by

$$\tau(u) = \sum_{\ell=1}^{n} \alpha_{\ell}\delta_{x_{\ell}}.$$ 

The function $q$ is the probability mass function of the random variable $n$, and given a particular $n$ we are able to define the probability measure $\mu_n^0$ for the intensities $(\alpha_1^n, ..., \alpha_n^n) \in \mathbb{C}^n$ and positions $(x_1^n, ..., x_n^n) \in D^n_k$. We now state a simple to prove proposition from which we can deduce whether $u \sim \mu^0$ has moments.

Proposition 8 Assume that $\mu^0$ is constructed as in Theorem 7 and that for any $n \in \mathbb{N}$ the measure $\mu_n^0$ have $p$-th moment. Let $id$ denote the identity on $\ell^1_{\kappa}$. Then there holds

$$\mathbb{E}_{\mu^0}[\|id\|_{\Sigma}] = \sum_{n \in \mathbb{N}_0} q(n)\mathbb{E}_{\mu_n^0}[\|id\|_{\ell^1_{\kappa}}],$$

and if the last series is finite, then $\mu^0$ has $p$-th moment.

3.2 Properties of the observation operator and potential

Now we investigate in the observation operator $G$. It is important to see that the Helmholtz equation only influences the posterior through the observation operator $G$. Therefore, we derive suitable properties of $G$ from the underlying PDE model.

Proposition 9 The observation operator $G$ is $\mu^0$-measurable and satisfies

$$\|G(u)\|_{\Sigma} \leq C \left(\|u\|_\Theta + \|g\|_{H^\frac{1}{2}(\Gamma_N)} \right).$$

If the prior $\mu^0$ has $p$-th moment then $G \in L^p(\ell^1_{\kappa}; \mu^0)$.

Proof Because $\mu^0$ is a measure on $(\ell^1_{\kappa}, \mathcal{F})$ and $\mathcal{F}$ is the Borel $\mathcal{F}$-Algebra we can apply Corollary 3 for every point $z \in M_k$ to show that $G : \ell^1_{\kappa} \to \mathbb{C}^m$ is continuous and in particular $\mu^0$-measurable. Again Corollary 3 applied to every observation point $z \in M_k$ yields

$$\|G(u)\|_{\Sigma} \leq C \left(\|u\|_\Theta + \|g\|_{H^\frac{1}{2}(\Gamma_N)} \right),$$

which shows the bound and $G \in L^p(\ell^1_{\kappa}; \mu^0)$ if $\mu^0$ has $p$-th moment.

Definition 12 Define the potential $\Psi$ as follows:

$$\Psi : \ell^1_k \times \mathbb{C}^m \to \mathbb{R}, \quad \Psi(u, y) = \frac{1}{2} \|y - G(u)\|_{\Sigma}^2. \quad (26)$$

We assume that $\| \cdot \|_{\Sigma}$ is derived from an inner product $(\cdot, \cdot)_\Sigma$ on $\mathbb{C}^m$.

Because the norm $\| \cdot \|_{\Sigma}$ is defined on a finite dimensional space, where all norms are equivalent, we exclusively use the $\| \cdot \|_{\Sigma}$-norm w.l.o.g.

Given the regularity of the observation operator $G$ we are able to deduce several properties of the potential $\Psi$.

Lemma 2 Assume that $\mu^0$ has second moment. Then the potential $\Psi$ and the prior measure $\mu^0$ satisfy:

(i) $\Psi \geq 0$.

(ii) There exists a $\mu^0$-measurable set $X \subseteq \ell^1_k$ and constants $K, C > 0$ such that $\mu^0(X) > 0$ and

$$\Psi(u, y) \leq K + C\|y\|_{\Sigma}^2, \quad \forall u \in X, \quad \forall y \in \mathbb{C}^m.$$

(iii) For every $y \in \mathbb{C}^m$ the map $\Psi(\cdot, y) : \ell^1_k \to \mathbb{R}$ is $\mu^0$-measurable.

(iv) If $\mu^0$ has $p$-th moment for $p \geq 2$ then for every $y \in \mathbb{C}^m$ we have that $\Psi(\cdot, y) \in L^p(\ell^1_k; \mu^0)$.

(v) For every $r > 0$ there exists a $C > 0$ such that

$$|\Psi(u, y_1) - \Psi(u, y_2)| \leq C (\|u\|_\Theta + 1) \|y_1 - y_2\|_{\Sigma}$$

for all $u \in \ell^1_k$ and for all $y_1, y_2 \in B_r(0)$. 

Proof Property (i) follows directly from the definition of the potential $\Psi$. To prove property (ii) apply Markov’s inequality to the nonnegative random variable $\|G\|_{\Sigma}^2 \in L^1(\ell_k^1; \mu^0)$ (see Proposition 9) to get
\[
\mu^0\left(\left\{ u \in \ell_k^1 : \|G(u)\|_{\Sigma}^2 \geq 2\mathbb{E}_{\mu^0}[\|G\|_{\Sigma}^2]\right\}\right) \leq \frac{\mathbb{E}_{\mu^0}[\|G\|_{\Sigma}^2]}{2\mathbb{E}_{\mu^0}[\|G\|_{\Sigma}^2]} = \frac{1}{2}.
\]
Thus we are able to define $X$ as the complementary set of events
\[
X := \left\{ u \in \ell_k^1 : \|G(u)\|_{\Sigma}^2 < 2\mathbb{E}_{\mu^0}[\|G\|_{\Sigma}^2]\right\}
\]
and conclude $\mu^0(X) \geq \frac{1}{2}$. Using the triangle and Young’s inequality we obtain for $u \in X, y \in \mathbb{C}^m$
\[
|\Psi(y, u) - \Psi(y, u_2)| = \frac{1}{2} \|y_1 - G(u)\|_{\Sigma}^2 - \|y_2 - G(u)\|_{\Sigma}^2
\]
\[
= \frac{1}{2}((y_1 - G(u), y_1 - G(u))_\Sigma
- (y_1 - G(u), y_2 - G(u))_\Sigma
+ (y_1 - G(u), y_2 - G(u))_\Sigma
- (y_2 - G(u), y_2 - G(u))_\Sigma
\]
\[
\leq \frac{1}{2} \|y_1 - y_2\|_\Sigma \|y_1 + y_2 - 2G(u)\|_\Sigma,
\]
which proves the claim using Proposition 9 to estimate $\|G(u)\|_\Sigma$.

3.3 Posterior

In the following, we define the posterior by means of a Radon-Nikodym derivative:

**Definition 13** Let the observation $y \in \mathbb{C}^m$ be given. The posterior density with respect to the prior measure is defined by
\[
\frac{d\mu^y}{d\mu^0} = \frac{1}{A(y)} \exp(-\Psi(u, y)), \tag{27}
\]
where $A(y)$ is a normalization constant given by
\[
A(y) = \int_{\ell_k^2} \exp(-\Psi(u, y))d\mu^0(u).
\]
Next we prove that the posterior measure $\mu^y$ is well-defined.

**Theorem 8** Let $y \in \mathbb{C}^m$. Assume that $\mu^0$ has second moment. Then the posterior defined by (27) is a well-defined probability measure on $\ell_k^1$. Moreover there exist constants $C_1, C_2 > 0$ independent of $y$ such that the normalization constant satisfies
\[
C_1 \exp(-C_2\|y\|_{\Sigma}^2) \leq A(y) \leq 1.
\]

**Proof** We follow [Stuart 2010, Theorem 4.1]. First notice that the potential $\Psi(\cdot, y)$ is $\mu^0$-measurable. To prove the lower bound of $A(y)$ we apply Lemma 2 (ii) to get an appropriate set $X \subseteq \ell_k^1$ with $\mu^0(X) > 0$ such that
\[
\int_{\ell_k^2} \exp(-\Psi(u, y))d\mu^0(u) \geq \int_X \exp(-\Psi(u, y))d\mu^0(u)
\]
\[
\geq \int_X \exp(-K - C\|y\|_{\Sigma}^2)d\mu^0(u)
\]
\[
= \mu^0(X)\exp(-K)\exp(-C\|y\|_{\Sigma}^2)
\]
\[
\geq C_1 \exp(-C_2\|y\|_{\Sigma}^2).
\]
This implies $A(y) > 0$. For the upper bound we use Lemma 2 (i) and that $\mu^0$ is a probability measure to immediately get
\[
A(y) = \int_{\ell_k^2} \exp(-\Psi(u, y))d\mu^0(u) \leq \int_{\ell_k^2} 1 d\mu^0(u) = 1.
\]
This proves well-definedness of the posterior.

In the next theorem we show that the posterior measure is stable with respect to small variations in $y$ in the Hellinger distance $d_{\text{Hell}}$.

**Theorem 9** Assume that $\mu^0$ has second moment. Then for all $r > 0$ there exists $C > 0$ such that
\[
d_{\text{Hell}}(\mu^{y_1}, \mu^{y_2}) \leq C\|y_1 - y_2\|_\Sigma, \quad \forall y_1, y_2 \in B_r(0).
\]

**Proof** First let us define the function
\[
f(u, y) := \exp\left(-\frac{1}{2}\Psi(u, y)\right).
\]
This function satisfies
\[
\|f(\cdot, y_1) - f(\cdot, y_2)\|_{L^2(\ell_k^1; \mu^0)}^2
= \int_{\ell_k^2} \left(\exp\left(-\frac{1}{2}\Psi(u, y_1)\right) - \exp\left(-\frac{1}{2}\Psi(u, y_2)\right)\right)^2 d\mu^0(u)
\]
\[
\leq C \int_{\ell_k^2} (\Psi(u, y_1) - \Psi(u, y_2))^2 d\mu^0(u)
\]
\[
\leq C \int_{\ell_k^2} (\|u\|_\Sigma + 1)^2 \|y_1 - y_2\|_{\Sigma}^2 d\mu^0(u)
\]
\[
\leq C\|y_1 - y_2\|_{\Sigma}^2.
\]
The function \( f : \mathbb{R}^d \to \mathbb{R} \) is locally Lipschitz for \( x > 0 \) and applying Theorem 11 so that

\[
A(y) \geq C_1 \exp(-C_2 \|y\|_2) \geq C > 0, \quad \forall y \in B_r(0),
\]

we obtain the estimate

\[
|A(y_1)^{-\frac{1}{2}} - A(y_2)^{-\frac{1}{2}}| \leq C|A(y_1) - A(y_2)| \leq C\|y_1 - y_2\|_2.
\]

With this result in mind take a look at two times the square of the Hellinger distance:

\[
2d_{\text{Hell}}(\mu^{y_1}, \mu^{y_2})^2 = \int_{\ell_2^d} \left( A(y_1)^{-\frac{1}{2}} \exp \left( -\frac{1}{2} \Psi(u, y_1) \right) - A(y_2)^{-\frac{1}{2}} \exp \left( -\frac{1}{2} \Psi(u, y_2) \right) \right)^2 d\mu^0(u)
\]

\[
\leq \int_{\ell_2^d} \left( A(y_1)^{-\frac{1}{2}} \|f(\cdot, y_1) - f(\cdot, y_2)\|_{L^2(\ell_2^d; \mu^0)} + \|f(\cdot, y_2)\|_{L^2(\ell_2^d; \mu^0)} |A(y_1)^{-\frac{1}{2}} - A(y_2)^{-\frac{1}{2}}| \right)^2 d\mu^0(u)
\]

\[
\leq C\|y_1 - y_2\|_2^2.
\]

Here, the last inequality follows from the previous estimates together with \( f \leq 1 \) and \( A(y) \geq C \).

Theorem 11 has shown that the posterior is in fact well-defined. Often it is desirable to show existence of moments of it.

**Theorem 10** Let \( y \in \mathbb{C}^m \). Assume that \( \mu^0 \) has \( p \)-th moment with \( p \geq 2 \) and let \( (X, \| \cdot \|_X) \) be a Banach space. Furthermore, let \( f : \ell_2^d \to X \) be a \( \mu^0 \)-measurable function such that for \( \mu^0 \)-almost every \( u \in \ell_2^d \) there holds

\[
\|f(u)\|_X \leq C\|u\|_{\ell_2^d}^p.
\]

Then \( f \in L^1(\ell_2^d; \mu^0) \). Moreover, \( \mu^p \) has \( p \)-th moment.

**Proof** There holds

\[
\int_{\ell_2^d} \|f(u)\|_X d\mu^0(u) = A(y)^{-1} \int_{\ell_2^d} \|f(u)\|_X \exp(-\Psi(u, y)) d\mu^0(u)
\]

\[
\leq CA(y)^{-1} \int_{\ell_2^d} \|u\|_p^p \exp(-\Psi(u, y)) d\mu^0(u)
\]

\[
\leq C \int_{\ell_2^d} \|u\|_p^p d\mu^0(u),
\]

where we have used that \( \Psi \geq 0 \). This shows \( f \in L^1(\ell_2^d; \mu^0) \). Observe that the posterior \( \mu^p \) has \( p \)-th moment since we can choose \( f = \text{id} \), where \( \text{id} \) is the identity on \( \ell_2^d \).

**Theorem 11** Assume that \( \mu^0 \) has second moment and let \( (X, \| \cdot \|_X) \) be a Banach space. Let \( f : \ell_2^d \to X \) satisfy \( f \in L^2(\ell_2^d; \mu^0) \) and \( f \in L^2(\ell_2^d; \mu^y) \). Let \( r > 0 \) and \( y_1, y_2 \in B_r(0) \subset \mathbb{C}^m \). Then there is a \( C > 0 \) such that

\[
\|\mathbb{E}_{\mu^y}[f] - \mathbb{E}_{\mu^{y_2}}[f]\|_X \leq C\|y_1 - y_2\|_2.
\]

In particular, computing moments is stable under small perturbations in the measurements \( y \).

**Proof** ([Stuart] 2010) Lemma 6.37 delivers

\[
\|\mathbb{E}_{\mu^y}[f] - \mathbb{E}_{\mu^{y_2}}[f]\|_X \leq 2 (\|\mathbb{E}_{\mu^1}[f]\|_X^2 + \|\mathbb{E}_{\mu^2}[f]\|_X^2)^{\frac{1}{2}} d_{\text{Hell}}(\mu^1, \mu^2).
\]

Applying Theorem 11 yields the desired result.

## 4 Sampling

In this section, we derive a method to sample from the posterior. First, we introduce a discretized observation operator to obtain a computable discretized posterior measure and show several properties of it. We proceed to apply a Sequential Monte Carlo method to generate samples from the posterior. Finally, we derive a method to decrease the complexity of evaluating the observation operator multiple times.

### 4.1 Discrete Approximation of the Posterior Measure

In practice we are not able to compute the solution of the Helmholtz equation exactly. As a consequence we have to replace the observation operator \( G \) by an approximation \( G_h \), where \( h \) is a discretization parameter. In our setting \( h \) is the mesh diameter. The discrete observation operator is then given by

\[
G_h(u) := (y_{u,h}(z_j))_{j=1}^m.
\]
The next lemma states several properties of the discrete observation operator and its relation to $G$.

**Lemma 3** There exists a $h_0 > 0$ such that for every $h \in (0, h_0]$ and every $u \in \ell_k^1$, the discrete observation operator satisfies

$$
\|G_h(u)\|_\Sigma \leq C \left(\|u\|_{\ell_k} + \|g\|_{H^{1/2}(\Gamma_N)}\right),
$$

(28)

$$
\|G(u) - G_h(u)\|_\Sigma \leq C \ln h |h|^2 \left(\|u\|_{\ell_k} + \|g\|_{H^{1/2}(\Gamma_N)}\right).
$$

(29)

with constants $C$ independent of $u$ and $h$. Furthermore, $G_h$ is $\mu^0$-measurable and if $\mu^0$ has $p$-th moment then $G_h \in L^p(\ell_k^1; \mu^0)$.

**Proof** From Theorem 6 we have

$$|y_u(u) - y_u(h)(\cdot)| \leq C \ln h |h|^2 \left(\|u\|_{\ell_k} + \|g\|_{H^{1/2}(\Gamma_N)}\right)$$

with a constant $C > 0$ independent of $u$ and $h$. This estimate applies to the measurement points $(z_j)_{j=1}^m \in M_k^N$ shows (29). By the triangle inequality we have

$$\|G_h(u)\|_\Sigma \leq \|G_h(u) - G(u)\|_\Sigma + \|G(u)\|_\Sigma,$$

which implies (28) using the a-priori estimate, $h \leq h_0$ and Proposition 9 $\mu^0$ is $\mu^0$-measurable because it is continuous by Corollary 2. In particular,

$$\|G_h(u) - G_h(v)\|_\Sigma \leq C \sum_{j=1}^m |y_{u,h}(z_j) - y_{v,h}(z_j)|$$

$$\leq C_{\kappa} m \left(\|u\|_{\ell_k} + 1 + |\ln h| |h|^2\right) \|u - v\|_{\ell_k}.$$

An application of (28) under the assumption that $\mu^0$ has $p$-th moment shows $G_h \in L^p(\ell_k^1; \mu^0)$.

We further have to work with a discretized potential $\Psi_h$. Let us fix some arbitrary measurement observation $y \in \mathbb{C}^m$ in (1) and define the discrete potential

$$\Psi_h(u, y) := \frac{1}{2} \|y - G_h(u)\|_{\Sigma}^2.$$

The next lemma states some essential properties of $\Psi_h$.

**Lemma 4** There exist $C, h_0 > 0$ such that for every $h \in (0, h_0]$ the discrete potential satisfies for all $u \in \ell_k^1, y \in \mathbb{C}^m$.

$$|\Psi(u, y) - \Psi_h(u, y)|$$

$$\leq C \ln h |h|^2 \left(\|u\|_{\ell_k}^2 + \|y\|_{\Sigma}^2 + \|g\|_{H^{1/2}(\Gamma_N)}^2\right).$$

Furthermore, Lemma 3 is valid for $\Psi_h$ instead of $\Psi$ with constants and sets independent of $h$.

**Proof** We compute

$$|\Psi(u, y) - \Psi_h(u, y)|$$

$$= \frac{1}{2} \|(y - G_h(u), y - G_h(u)) - (y - G(u), y - G(u))\|_{\Sigma}$$

$$+ \|(y - G(u), y - G_h(u)) - (y - G(u), y - G(u))\|_{\Sigma}$$

$$\leq \frac{1}{2} \|G(u) - G_h(u)\|_{\Sigma} \|2y - G(u) - G_h(u)\|_{\Sigma}$$

$$\leq C \|G(u) - G_h(u)\|_\Sigma \|y\|_{\Sigma} + \|G(u)\|_{\Sigma} + \|G_h(u)\|_{\Sigma}.$$

Now the a-priori error estimate in Lemma 3 together with $\|G(u)\| \leq C \|u\|_{\ell_k}$ and Young’s inequality imply

$$\|\Psi(u, y) - \Psi_h(u, y)\|$$

$$\leq C \ln h |h|^2 \left(\|u\|_{\ell_k} + \|g\|_{H^{1/2}(\Gamma_N)}\right) \|y\|_{\Sigma} + \|u\|_{\ell_k}$$

$$\leq C \ln h |h|^2 \left(\|u\|_{\ell_k}^2 + \|y\|_{\Sigma}^2 + \|g\|_{H^{1/2}(\Gamma_N)}^2\right).$$

That Lemma 2 holds with $\Psi$ replaced by $\Psi_h$ with constants and sets independent of $h$ is shown by a straightforward computation following the proof of Lemma 2. We omit most of it for brevity. The only slight difficulty is to define the set $X \subseteq \ell_k^1$ in (ii) independently of $h$. Therefore, let

$$X_h := \{u \in \ell_k^1 : \|G_h(u)\|_{\Sigma}^2 \leq 2E_{\mu^0}[\|G_h\|_{\Sigma}^2]\}$$

and conclude $\mu^0(X_h) \geq \frac{1}{2}$ for any $h \in (0, h_0]$ as shown in the proof of (ii) of Lemma 2. Using Young’s inequality we estimate

$$\sup_{h \in (0, h_0]} 2E_{\mu^0}[\|G_h\|_{\Sigma}^2]$$

$$\leq 4 \sup_{h \in (0, h_0]} E_{\mu^0}[\|G_h - G\|_{\Sigma}^2 + \|G\|_{\Sigma}^2]$$

$$\leq C \sup_{h \in (0, h_0]} E_{\mu^0} \left(\ln h|h|^2 \left(\|u\|_{\ell_k} + \|g\|_{H^{1/2}(\Gamma_N)}\right)^2\right) + C$$

$$\leq C_1,$$

where we used $h \leq h_0$, that $\mu^0$ has second moment and $G \in L^2(\ell_k^1; \mu^0)$. Hence, setting

$$X := \left\{u \in \ell_k^1 : \|G_h(u)\|_{\Sigma}^2 \leq \sup_{h \in (0, h_0]} 2E_{\mu^0}[\|G_h\|_{\Sigma}^2]\right\}$$

implies $X_h \subseteq X$ and therefore $\mu^0(X) \geq \frac{1}{2}$. Now we apply Young’s inequality to obtain

$$\Psi_h(u, y) \leq \|y\|_{\Sigma}^2 + \|G_h(u)\|_{\Sigma}^2$$

$$\leq \|y\|_{\Sigma}^2 + \sup_{h \in (0, h_0]} 2E_{\mu^0}[\|G_h\|_{\Sigma}^2] \leq \|y\|_{\Sigma}^2 + C_1$$

for all $u \in X, y \in \mathbb{C}^m$. This proves the statement.
We apply Lemma 4 and Lemma 3 to show that Lemma 2 holds if we replace $G$ and $\Psi$ by their discrete counterparts $G_h$ and $\Psi_h$. We emphasize that all the estimates in Lemma 2 are valid uniformly in $h$ as long as $h \in (0, h_0]$ for a suitably small $h_0$. Then applying Theorem 5 ensures that the discrete posterior $\mu_h^\Psi$ is well-defined, i.e. $\mu_h^\Psi$ is defined by the Radon-Nikodym derivative with respect to the prior

$$\frac{d\mu_h^\Psi}{d\mu^0} = \frac{1}{A_h(y)} \exp(-\Psi_h(u, y)).$$

(30)

Here, $A_h(y)$ is a normalization constant defined by

$$A_h(y) = \int_{\ell^2_2} \exp(-\Psi_h(u, y)) d\mu^0(u).$$

We summarize this result in the next theorem.

**Theorem 12** Let $y \in \mathbb{C}^m$. Assume that $\mu^0$ has a finite second moment. Then there exists $h_0 > 0$ such that for any $h \in (0, h_0]$ the discrete posterior measure $\mu_h^\Psi$ is well-defined. There exist constants $C_1, C_2 > 0$ independent of such that the normalization constants satisfy

$$C_1 \exp(-C_2 ||y||_2^2) \leq A_h(y) \leq 1, \quad \forall y \in \mathbb{C}^m.$$

In particular, (30) yields the well-definedness of $\mu_h^\Psi$.

**Proof** The proof is analogous to the proof of Theorem 3 using the fact that Lemma 2 holds for the discrete potential $\Psi_h$, with constants and set independent of $h$.

The next theorem states that under certain conditions the rate of convergence in $h$ of the observation operator carries over to the rate of convergence of $\mu_h^\Psi$ to $\mu^\Psi$ in the Hellinger distance.

**Theorem 13** Let $y \in \mathbb{C}^m$. Assume that $\mu^0$ has a finite fourth moment. Then there exists a $h_0 > 0$ such that for any $h \in (0, h_0]$ we have

$$d_{\text{Hell}}(\mu_h^\Psi, \mu^\Psi) \leq C |\ln h|h^2.

C does not depend on $h$, but on $y$.

**Proof** Proving this theorem is similar to proving Theorem 5. We define the functions

$$f(u) = \exp\left(-\frac{1}{2}\Psi(u, y)\right),$$

$$f_h(u) = \exp\left(-\frac{1}{2}\Psi_h(u, y)\right).$$

Using the properties of the exponential function from Lemma 3 together with Lemma 4 yields

$$\|f - f_h\|^2_{L^2(\ell^2_2, \mu^0)}$$

$$= \int_{\ell^2_2} \left(\exp\left(-\frac{1}{2}\Psi(u, y)\right) - \exp\left(-\frac{1}{2}\Psi_h(u, y)\right)\right)^2 d\mu^0(u)$$

$$\leq C \int_{\ell^2_2} (\Psi(u, y) - \Psi_h(u, y))^2 d\mu^0(u)$$

$$\leq C |\ln h|h^2 \int_{\ell^2_2} \left(\|u\|^2_{\ell^2_2} + \|y\|^2_{\ell^2_2} + \|g\|^2_{\ell^2_2} \right) d\mu^0(u)$$

$$\leq C |\ln h|h^2,$$

where we have used that $\mu^0$ has fourth moment in the last inequality. For the normalization constants observe that

$$|A(y) - A_h(y)|$$

$$\leq \int_{\ell^2_2} |\exp(-\Psi(u, y)) - \exp(-\Psi_h(u, y))| d\mu^0(u)$$

$$\leq \int_{\ell^2_2} |\Psi(u, y) - \Psi_h(u, y)| d\mu^0(u)$$

$$\leq C |\ln h|h^2 \int_{\ell^2_2} \left(\|u\|^2_{\ell^2_2} + \|y\|^2_{\ell^2_2} + \|g\|^2_{\ell^2_2} \right) d\mu^0(u)$$

$$\leq C |\ln h|h^2.$$

We are able to bound $A(y)$ and $A_h(y)$ away from zero independently of $h$ using Theorem 8 and Theorem 12 which shows that

$$|A(y) - \frac{1}{2} A_h(y) - \frac{1}{2} A_h(y)| \leq C |A(y) - A_h(y)| \leq C |\ln h|h^2,$$

because $x \mapsto x - \frac{1}{2}$ is locally Lipschitz for $x > 0$. A straight forward computation yields the result:

$$2d_{\text{Hell}}(\mu_h^\Psi, \mu^\Psi)^2$$

$$= \int_{\ell^2_2} \left(\|A(y) - \frac{1}{2} A_h(y) - \frac{1}{2} A_h(y)\|^2_{L^2(\ell^2_2, \mu^0)} + \|A(y) - \frac{1}{2} A_h(y) - \frac{1}{2} A_h(y)\|_{L^2(\ell^2_2, \mu^0)}^2\right)$$

$$\leq \left(|A(y) - \frac{1}{2} A_h(y) - \frac{1}{2} A_h(y)|\right)^2$$

$$\leq C |\ln h|h^2.$$

Here we have used the previous estimates, $f_h \leq 1$ for any $h$ and $A(y) > 0$.

**Theorem 14** Let $y \in \mathbb{C}^m$. Assume that $\mu^0$ has a finite fourth moment. Let $(X, \|\cdot\|_X)$ be a Banach space and
Let \( f : \ell^1_k \to X \) have second moments with respect to both \( \mu^\Psi \) and \( \mu^\Psi_n \). Then there holds
\[
\|E_{\mu^\Psi}[f] - E_{\mu^\Psi_n}[f]\|_X \leq C|\ln h|^2.
\]

**Proof** As in the proof of Theorem 11 this follows from (Stuart, 2010, Lemma 6.37) using Theorem 13.

### 4.2 Sequential Monte Carlo

Throughout Section 4.2 let \( y \in \mathbb{C}^m \) be a given observation. In the following section we will use the Sequential Monte Carlo Method (SMC) from (Dashti and Stuart, 2015) to draw samples from the posterior measure. We also derive an error estimate.

Let \( J \in \mathbb{N} \) and for \( j \in \{0, \ldots, J\} \) define a sequence of measures \( \mu_j << \mu_0 \) by
\[
\frac{d\mu_j}{d\mu_0}(u) := L_j^{-1} \exp\left(-jJ^{-1}\Psi(u, y)\right),
\]
\[
L_j := \int \exp\left(-jJ^{-1}\Psi(u, y)\right) d\mu_0(u).
\]

(31)

Note that \( \mu_0 \) is equal to the prior \( \mu^0 \) and \( \mu_j \) equal to the posterior measure \( \mu^\Psi \). Our goal is to approximate \( \mu_j \) sequentially using information of each \( \mu_j \) to construct the next approximation \( \mu_{j+1} \). One idea behind SMC is the approximation of each measure \( \mu_j \) by a weighted sum of Dirac measures
\[
\mu_j \approx \mu^N_j := \sum_{n=1}^N w_j^{(n)} \delta_{\tilde{u}_j^{(n)}},
\]

(32)

with \( w_j^{(n)} \in \ell^1_k \) and weights \( w_j^{(n)} \geq 0 \) that sum up to 1. We define the operator
\[
S : \mathbb{P}(\ell^1_k, \mathcal{F}) \to \mathbb{P}(\ell^1_k, \mathcal{F})
\]
\[
\nu \mapsto \frac{1}{N} \sum_{n=1}^N \delta_{u_j^{(n)}}, \quad u_j^{(n)} \sim \nu,
\]

Here \( \mathbb{P}(\ell^1_k, \mathcal{F}) \) denotes the space of probability measures on \( (\ell^1_k, \mathcal{F}) \). We also define
\[
L : \mathbb{P}(\ell^1_k, \mathcal{F}) \to \mathbb{P}(\ell^1_k, \mathcal{F})
\]
\[
\nu \mapsto \frac{1}{L_j} \exp(-jJ^{-1}\Psi(u, y)).
\]

We remark that the operator \( L \) satisfies
\[
\mu_{j+1} = L\mu_j, \quad j = 1, \ldots, J.
\]

We further define \( P_j : \ell^1_k \times \mathcal{F} \to [0, 1] \) as \( \mu_j \)-invariant Markov kernel. That means
\[
\mu_j(A) = \int_{\ell^1_k} P_j(u, A) d\mu_j(u) \quad \forall A \in \mathcal{F}.
\]

The idea of the kernel is to redraw samples in each iteration of the algorithm to better approximate \( \mu_j \). (See 4. in Algorithm 1.)

This allows us to define the discrete measures according to (Dashti and Stuart, 2015, Algorithm 5.12.) as follows
\[
\mu_j^N := S^N \mu_0, \quad \mu_{j+1} := L S^NP_j \mu_j^N, \quad j = 0, \ldots, J - 1.
\]

#### Algorithm 1: SMC-Algorithm

1. Let \( \mu_0^N = \mu_0 \) and set \( j = 0 \).
2. Draw \( u_j^{(n)} \sim \mu_j^N \), \( n = 1, \ldots, N \).
3. Set \( w_j^{(n)} = \frac{1}{N}, n = 1, \ldots, N \) and define \( \mu_j^N \) by (32).
4. Resampling: draw \( \tilde{u}_{j+1}^{(n)} \sim P_j(u_j^{(n)}, \cdot) \).
5. Define \( w_{j+1}^{(n)} = \frac{1}{N} \sum_{n=1}^N \tilde{w}_{j+1}^{(n)} \) with
\[
\tilde{w}_{j+1}^{(n)} = \exp(-(j+1)J^{-1}\Psi(U_{j+1}^{(n)}, y))w_j^{(n)}
\]
\[
\mu_{j+1}^N := \sum_{n=1}^N \tilde{w}_{j+1}^{(n)} \delta_{\tilde{u}_{j+1}^{(n)}}.
\]

In Algorithm 1 we see the SMC-Algorithm described in (Dashti and Stuart, 2015, Section 5.3.). Let us generalize (Dashti and Stuart, 2015, Theorem 5.13) in such a way that it applies to our setting.

**Theorem 15** For every measurable and bounded function \( f \) the measure \( \mu_j^N \) satisfies
\[
\mathbb{E}_{\text{SMC}}[\mathbb{E}_{\mu_j^N}[f] - \mathbb{E}_{\mu^\Psi}[f]]^2 \leq \left( \sum_{j=1}^J (2A_j)^{-1} \right)^2 \frac{\|f\|_\infty^2}{N},
\]

(33)

where \( \mathbb{E}_{\text{SMC}} \) is the expectation with respect to the randomness in the SMC algorithm.

**Proof** We first prove a variant of (Dashti and Stuart, 2015, Lemma 5.17) using similar techniques. Define \( g : \exp(-J^{-1}\Psi) \) and let \( f \) with \( \|f\|_\infty \leq 1 \) be given. Then from the proof of (Dashti and Stuart, 2015, Lemma 5.17) and defining \( \eta_j = S^N P_j \mu_j^N \) we conclude
\[
(L\mu_j)[f] - (L\eta_j)[f]
\]
\[
= \frac{1}{\mathbb{E}_{\mu_j}[g]}(\mathbb{E}_{\mu_j}[fg] - \mathbb{E}_{\eta_j}[fg])
\]
\[
+ \frac{\mathbb{E}_{\eta_j}[fg]}{\mathbb{E}_{\mu_j}[g]}(\mathbb{E}_{\eta_j}[g] - \mathbb{E}_{\mu_j}[g]).
\]

A quick calculation shows
\[
\mathbb{E}_{\mu_j}[g] = \mathbb{E}_{\mu^\Psi}[\exp(-jJ^{-1}\Psi)] \geq \mathbb{E}_{\mu^\Psi}[\exp(-\Psi)] = A_j
\]
and from \( \|f\|_{\infty} \leq 1 \) we conclude \( |E_{\eta_j}[fg]/E_{\eta_j}[g]| \leq 1 \). Hence we obtain

\[
|(L_{\mu_j})[f] - (L_{\eta_j})[f]| \\
\leq A_j^{-1} |E_{\mu_j}[fg] - E_{\eta_j}[fg]| + A_j^{-1} |E_{\mu_j}[g] - E_{\eta_j}[g]|. \tag{34}
\]

We define the distance \( d_{op} \) for probability measures \( \nu, \eta \) on \((\ell^1_k, \mathcal{F})\) as follows

\[
d_{op}(\nu, \eta) := \sup_{\|f\|_{\infty} \leq 1} \text{op}_{\text{SMC}}[|E_{\nu}[f] - E_{\eta}[f]|^2]^{1/2}
\]

and from \( \|g\|_{\infty} \leq 1, \|f\|_{\infty} \leq 1 \) together with \( \|g\|_{\infty} \leq 1 \) we conclude

\[
d_{op}(LS^N P_j \mu_j^N, L\mu_j) \leq 2A_j^{-1} d_{op}(S^N P_j \mu_j^N, \mu_j). \tag{35}
\]

Now we show that \cite{Dashti2015} Theorem 5.13 holds with weaker assumptions. First we use \( \|\mu\|_{\infty} \leq 1 \) and the triangle inequality to conclude

\[
d_{op}(\mu_j^N, \mu_{j+1}) = d_{op}(LS^N P_j \mu_j^N, L\mu_j) \\
\leq 2A_j^{-1} d_{op}(S^N P_j \mu_j^N, \mu_j) \\
\leq 2A_j^{-1} d_{op}(S^N P_j \mu_j^N, P_j \mu_j) \\
\leq 2A_j^{-1} d_{op}(S^N P_j \mu_j^N, P_j \mu_j) + d_{op}(P_j \mu_j^N, P_j \mu_j))
\]

A straightforward continuation similar to the proof of \cite{Dashti2015} Theorem 5.13 leads to the statement

\[
d_{op}(\mu_j^N, \mu_j^N) \leq \sum_{j=1}^{N} (2A_j^{-1})^j \frac{1}{\sqrt{N}}. \tag{36}
\]

Here we remark that both \cite{Dashti2015} Lemma 5.15 and \cite{Dashti2015} Lemma 5.16 needed for the proof of \cite{Dashti2015} Theorem 5.13 still hold. From \( \|g\|_{\infty} \leq 1 \) we conclude that \( \|\mu\|_{\infty} \leq 1 \) and the statement for a measurable bounded function \( f \) follows by a scaling argument.

5 Numerical Experiments

The Set Up

In this section, we present our numerical results for our prior model. Let \( D = [0,1]^2 \) be the physical domain. We want to recover two sound sources placed in \( x_{\text{exact},1} = (0.25,0.75) \) and \( x_{\text{exact},2} = (0.75,0.75) \) with amplitudes \( \alpha_{\text{exact},1} = 10 + 10i \) and \( \alpha_{\text{exact},2} = 10 + 10i \). We also choose \( g = 0 \). The three measurement points \( (z_i)_{i=1}^{3} \) are located at the following positions

\[
z_1 = (0.1, 0.5), \ z_2 = (0.5, 0.5), \ z_3 = (0.9, 0.5).
\]

We choose \( \kappa = 0.05 \) and thus as source domain \( D_{\kappa} = [0,1,0.9] \times [0,0.6,0.9] \). Recall that the measurement domain is defined by

\[
M_{\kappa} := \{x \in D \mid \text{dist}(x,D_{\kappa}) > \kappa \ \text{and} \ \text{dist}(x,G) > \kappa\},
\]

so that we indeed have \( z_1, \ldots, z_3 \in M_{\kappa} \).

For the parameters of the Helmholtz equation we consider the fluid density \( \rho = 1 \), frequency \( \Omega = 30 \), sound speed \( c = 5 \), and coefficients \( \alpha(\zeta) = 1, \beta(\zeta) = 1/30 \) for the isolating material on the boundary \( \Gamma_G = \Gamma \).

We define \( u_{\text{exact}} = (\alpha_{\text{exact},1}, x_{\text{exact},1}, \alpha_{\text{exact},2}, x_{\text{exact},2}, 0,0,\ldots) \in \ell^1_k \) and \( y_{\text{exact}} := G_h(u_{\text{exact}}) \). We chose the prior number of sources \( k \) to be Poisson distributed with expectation 2, that is \( k \sim \text{Poi}(2) \). Given the number of sources \( k \), we choose the amplitudes \( \alpha_k^1,\ldots,\alpha_k^k \) to be independently and identically complex normal distributed with mean 10 + 10i and \( \Sigma = \begin{pmatrix} 0 & 0 \\ 0 & 10 \end{pmatrix} \subset C^{2 \times 2} \).

The prior source positions are uniformly distributed \( x_1^k,\ldots,x_k^k \sim \mathcal{U}(D_{\kappa}) \). We in particular assume that the random variables \( \beta_k^1 \) and \( \beta_k^2 \) are independent. For the observational noise we consider \( \eta = (\eta_i)_{i=1}^{\infty} \) with \( \eta_1, \eta_2, \eta_3 \) all uniformly and identically distributed with mean 0 and \( \Sigma = \begin{pmatrix} 0.1 & 0 \\ 0 & 0.1 \end{pmatrix} \subset C^{2 \times 2} \).

The domain \( D \) is uniformly triangulated with triangles of diameter \( h \). Throughout our experiments we choose the Markov kernel

\[
P_j : \ell^1_k \times \mathcal{F} \to [0,1], \ (u,A) \mapsto \begin{cases} 1 & \text{if } u \in A, \\ 0 & \text{if } u \notin A, \end{cases}
\]

which is clearly invariant with respect to \( \mu_j \) (or any other measure).

The SMC Method in Action

We now present results on the source positions of samples from the posterior. We consider the SMC as introduced in Section 4 with \( J = 1 \) and \( N = 44800 \). Since we then only have one tempering step for \( J = 1 \) we drop the \( j \) from the index. We choose a uniform mesh of size \( h = \sqrt{2} \cdot 2^{-3} \) for all experiments in this section.

Algorithm 1 delivers the following data \( k^{(n)} \in \mathbb{N}_0, u^{(n)} = (u_1^{(n)}, u_2^{(n)}, \ldots, u_k^{(n)}, 0,0,\ldots) \in \ell^1_k \) for \( n = 1,\ldots,N \) and weights \( w^{(1)}(1),\ldots,w^{(N)}(N) \). We define the function \( P_{\text{emp}}(x) \) as an approximation to the probability that a source is located at \( x \in D \) in the following sense

\[
P_{\text{emp}}(x) := \sum_{n=1}^{N} u^{(n)} \max_{1 \leq i \leq k^{(n)}} \left\{ K_{\kappa} (x - x_i^{(n)}) \right\}
\]

\[
\approx \mathbb{P}_\mu(u \text{ has a source in } B_\varepsilon(x)).
\]
Here $\varepsilon = 0.04$ and $K_\varepsilon$ is a smooth cutoff function approximating the indicator function $1_{B_\varepsilon(0)}$. In Figure 1 (a) we can see the function $P_{\text{emp}}$ distinctly recovers the true source positions $x_{\text{exact,1}}$ and $x_{\text{exact,2}}$.

We further want to analyze positions of pairs. In particular we would like to find out if it happens that for a sample two positions are close to the same $x_{\text{exact,1}}$ ($l = 1$ or $l = 2$). That would mean we do not identify two distinct sound sources. Given that a sample has two sources, i.e. $k = 2$, and one is in the set $Q$ we look for the probability that the second source is close to some $x \in D$ but not in $Q$. Formally for the index set $I(Q) = \{n \in \{1, ..., N\} | k(n) = 2, x_{1(n)} \in Q \text{ or } x_{2(n)} \in Q\}$ we define

$$P_{\text{emp}}(x|Q, 2) := \frac{1}{\sum_{n \in I(Q)} w_{(n)}^{(n)}} \sum_{n \in I(Q)} w_{(n)}^{(n)} \max_{x_{l(n)}^{(n)}} K_{\varepsilon}(x - x_{l(n)}^{(n)})$$

$$\approx P_{\mu^J}(u \text{ has a source in } B_{\varepsilon}(x) \setminus Q | u \text{ has a source in } Q, k = 2).$$

The results for this function are given in Figure 1 (b) and (c). We conclude that if one of the sources is located near $x_{\text{exact,1}}$ then the other near $x_{\text{exact,2}}$ and vice versa.

We also would like to address the notion of a maximum a posteriori estimator (MAP). We define the empirical MAP-index and the empirical MAP-index given $k$ sources as follows

$$n_{\text{MAP}} := \arg\max_{n \in \{1, ..., N\}} w_{(n)}^{(n)},$$

$$n_{k\text{MAP}} := \arg\max_{n \in \{1, ..., N\}, k(n) = k} w_{(n)}^{(n)}, \quad (37)$$

In our experiments both quantities are well-defined and unique. The empirical MAP-estimator is the $n_{\text{MAP}}$-th sample of the last (and only) iteration $j = J = 1$ of the SMC. Analogously we define the empirical MAP-estimator given $k$ sources. Further results of this experiment are given in Table 1 and visualized in Figure 2. Note that Table 1 contains only results for $k = 2, 3$ number of sound sources. The occurrence probability for $k \neq 2, 3$ sound sources is below 1% and therefore not mentioned.

Convergence in Mean Square Error for functions

Theorem 15 shows linear convergence in the Mean Square Error (MSE) for bounded measurable functions $f$ integrated over the SMC posterior $\mu_{J}^\text{SMC}$. We are interested testing the convergence of different functions $f_1, ..., f_5$ listed in Table 2. These functions allow us to extract certain information from the posterior measure.

We chose $z_{\text{prediction}} := (1/2, 1/4) \in D$ to predict the pressure at a point distinct from the measurement points $z_1, ..., z_3$. For $f_5$ we chose time $t = 1$. We remark that Theorem 15 shows convergence results only for $f_2$. Nevertheless we observe convergence for the other functions in figure 9. For this experiment we fix the mesh size to $h_{\text{ref}} = \sqrt{2} \cdot 2^{-5}$ and only vary the number of samples. We chose $N_{\text{ref}} = 448.000$ and use $\mu_{N_{\text{ref}}}^\text{SMC}$ as reference measure. We used $j = J = 1$ as tempering step. For the expectation integral "$E_{\text{SMC}}$" we used a standard Monte Carlo method with 50 runs.

Convergence in the Hellinger Distance

We also verify Theorem 13 numerically. It states convergence of the discretized to the true posterior with

\[ Q \]
Table 1 $P_{\mu_Y}(A_k)$ is the occurrence probability of $A_k := \{ u \in \ell^2 \mid u \text{ has } k \text{ sources} \}$ with respect to the approximated posterior measure $\mu_Y^\ast$. $E_{\mu_Y^\ast}(x \mid k)$ and $E_{\mu_Y^\ast}(\alpha \mid k)$ are the expected sound source positions and amplitudes for a fixed number of $k$ sources with respect to $\mu_Y^\ast$. $x^{(\text{MAP})}$ and $\alpha^{(\text{MAP})}$ are the positions and amplitudes resulting from the MAP estimator in (37) with respect to $\mu_Y^\ast$. Further results for the experiment described in section 5.

Table 2 List of functions used in the experiments in Figure 3 and Figure 5.

Table 3 Probability to have exactly two sources $\alpha$ given three sources for $\mu_Y$. Further results for the experiment described in section 5.
Convergence in $h$ for functions

We want to verify Theorem 14 which states the convergence of the error

$$e_h(f) := \|E_{\mu^X}[f] - E_{\mu^Y}[f]\|_X \leq C|\ln h|h^2$$

(38)

for functions $f$ with second moments with respect to both $\mu^X$ and $\mu^Y$. We use the functions $f_1, ..., f_5$ defined in Table 2 and a reference measure $\mu_{ref} := \mu_{h_{ref}}$ with $h_{ref} = \sqrt{2} \cdot 2^{-7}$. For all experiments we choose $N_{ref} = 20,000$ samples from the SMC sampler with temperature $J = 1$. For this experiment we do not use the resampling step of the SMC sampler. For the other measures we vary the mesh parameter $h = \sqrt{2} \cdot 2^{-k}$ for $k = 2, ..., 6$ and average the approximation error $e_h$ in (38) over 20 runs. The results are seen in figure 5 and verify the stated convergence for the functions $f_1, ..., f_5$.

Conclusion

In this paper we studied source identification for the Helmholtz equation using the Bayesian approach. We derived suitable properties and regularity for the solution of the Helmholtz equation with a series of Diracs as forcing term. We defined a prior probability distribution on the number of sources, their position and intensity and proved well-posedness of the Bayesian inverse problem. We have shown that the discretized solution operator for the Helmholtz equation leads to a discretized posterior which converges to the true posterior in the Hellinger distance. We used the Sequential Monte Carlo method to sample from the posterior. In the last section, we applied these techniques to verify our results numerically.

Acknowledgements

All authors gratefully acknowledge support from the International Research Training Group IGDK, funded by the German Science Foundation (DFG) and the Austrian Science Fund (FWF).

Dominik Hafemeyer acknowledges support from the graduate program TopMath of the Elite Network of Bavaria and the TopMath Graduate Center of TUM Graduate School at Technische Universität München. He is also a scholarship student of the Studienstiftung des deutschen Volkes.

We thank Elisabeth Ullmann for the support she gave us by proof reading the draft of this paper and her helpful hints to improve the paper. We further thank Jonas Latz for supplying us with useful background on the Sequential Monte Carlo method.
Fig. 4 The continuous curve of the left figure shows the Hellinger distance of the discretized measure $\mu^y_h$ to the reference measure $\mu^y_{\text{ref}}$ averaged over 20 runs. As a reference $O(|\ln h|h^2)$ is drawn dotted. The right image shows the Variance of the Hellinger distance for different $h$ computed from 20 runs.

Fig. 5 The left figure shows the error $e_h(f_i)$ for the functions from Table 2. The reference convergence rate $O(|\ln h|h^2)$ is drawn dotted. The right figure shows the variance of $e_h(f_i)$ for 20 runs over which the error terms on left figure were averaged over.
References

Ambrosio L, Fusco N, Pallara D (2000) Functions of Bounded Variation and Free Discontinuity Problems. Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, New York

Asano F, Asoh H, Nakadai K (2013) Sound source localization using joint bayesian estimation with a hierarchical noise model. IEEE Transactions on Audio, Speech, and Language Processing 21:1953 – 1965

Bermúdez A, Gamallo P, Rodríguez R (2004) Finite element methods in local active control of sound. SIAM J Control Optim 43(2):437–465

Bowman F (1958) Introduction to Bessel functions. Dover Publications Inc., New York

Dashti M, Stuart AM (2015) The Bayesian Approach To Inverse Problems. ArXiv e-prints [1302.6989v4]

Dautray R, Lions JL (2012) Mathematical Analysis and Numerical Methods for Science and Technology: Volume 1 Physical Origins and Classical Methods. Springer Science & Business Media

Denison DGT, Holmes CC, Mallick BK, Smith AFM (2002) Bayesian methods for nonlinear classification and regression. Wiley Series in Probability and Statistics, John Wiley & Sons, Ltd., Chichester

Green PJ (1995) Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. Biometrika 82(4):711–732

Halliwell LJ (2015) Complex random variables. URL https://www.casact.org/pubs/forum/15fforum/Halliwell_Complex.pdf last access 09.05.2018

Košljakov NS, Smirnov MM, Gliner EB (1964) Differential equations of mathematical physics. North-Holland Publ. Co.

Nakamura K, Nakadai K, Asano F, Hasegawa Y, Tsujino H (2009) Intelligent sound source localization for dynamic environments. In: 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems, pp 664–669

Pieper K, Tang BQ, Trautmann P, Walter D (2018) Inverse point source location with the helmholtz equation. Submitted URL https://arxiv.org/abs/1805.03310v1

Richardson S, Green PJ (1997) On Bayesian analysis of mixtures with an unknown number of components. J Roy Statist Soc Ser B 59(4):731–792

Schatz AH (1974) An observation concerning Ritz-Galerkin methods with indefinite bilinear forms. Math Comp 28:959–962

Schatz AH, Wahlbin LB (1977) Interior maximum norm estimates for finite element methods. Math Comp 31(138):414–442