



**HAL**  
open science

# A volume integral method for solving scattering problems from locally perturbed infinite periodic layers

Housseem Haddar, Thi Phong Nguyen

► **To cite this version:**

Housseem Haddar, Thi Phong Nguyen. A volume integral method for solving scattering problems from locally perturbed infinite periodic layers. *Applicable Analysis*, 2016, pp.29. 10.1080/00036811.2016.1221942 . hal-01374892

**HAL Id: hal-01374892**

**<https://inria.hal.science/hal-01374892v1>**

Submitted on 2 Oct 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# A Volume integral method for solving scattering problems from locally perturbed infinite periodic layers

Housseem Haddar\*      Thi-Phong Nguyen\*

October 2, 2016

## Abstract

We investigate the scattering problem for the case of locally perturbed periodic layers in  $\mathbb{R}^d$ ,  $d = 2, 3$ . Using the Floquet-Bloch transform in the periodicity direction we reformulate this scattering problem as an equivalent system of coupled volume integral equations. We then apply a spectral method to discretize the obtained system after periodization in the direction orthogonal to the periodicity directions of the medium. The convergence of this method is established and validating numerical results are provided.

## 1 Introduction

We are concerned in this work by the design of a numerical method to compute the solution of the scattering problem from locally perturbed infinite periodic layers. This problem is encountered for instance in applications related to photonics or non destructive testing of gratings. Our work can be seen as an extension of the numerical method in [16, 20] where the solution for periodic media is computed based on quasi-periodicity of the incident wave. The principle of our method is similar to the method employed in [4, 5] and relies on the use of the Floquet-Bloch transform in the periodicity directions to transform the problem into coupled quasi-periodic problems. Our methodologies then differs in the way we formulate the problem and discretize it with respect to the space variable. While in [5] a finite element method is employed using a Dirichlet to Neumann map to bound the computation domain as introduced in [7, 11], we shall rely in the present work on a volume integral formulation of the problem. Discretizing the volume integral equation efficiently through the use of spectral method and FFT technique to evaluate the matrix-vector product has been introduced in [20] and studied in a number of papers [9, 10, 15, 20]. We here mainly rely on the numerical algorithm implemented in [16] for quasi-periodic problems but we shall consider TE modes. The TM mode can be treated in a similar manner using the adaptations proposed in [16]. As in [4], the first main difficulty in the convergence analysis is to establish convergence of the discretization of the problem with respect to the Floquet-Bloch variable. Using a trapezoidal rule to discretize the Floquet-Bloch transform induces a semi-discrete problem which is equivalent, up to a change in the source term, to the periodic problem with period  $ML$  where  $M$  is the number of discretization points and  $L$  is the periodicity length. The convergence (with respect to the parameter  $M$ ) then relies on the decay of the solution in the periodicity directions. The latter is indeed not guaranteed in general and usually requires some special assumptions on the material properties, such as strict monotonicity in the direction orthogonal to the periodicity directions [2, 3, 8, 14, 17]. We shall restrict ourselves here to a simplified problem where (small) absorption in the media is assumed through the consideration of complex wave numbers with positive imaginary parts. This assumption has also been done in [4]. Indeed this greatly simplifies the analysis since the underlying differential equation operator becomes coercive. This assumption allows us to focus more on the specificity of coupling spectral discretization of the volume integral method with discretization in the Floquet-Bloch variable. Indeed this step will enable us, in a future work, to treat the case without absorption and focus

---

\*INRIA, Ecole Polytechnique (CMAP) and Université Paris Saclay, Route de Saclay, 91128 Palaiseau Cedex, France. (Housseem.haddar@inria.fr, tpnguyen@cmap.polytechnique.fr).

on the technicality related to establishing Rellich type identities that can be exploited in the convergence analysis. Our assumption allows also to evacuate problems that would occur at the Wood anomalies. Let us notice however that the incorporation of a radiation condition for absorption free problems (as in [14]) would not induce major additional difficulties (for the numerical implementation of the method) since this condition is encoded in the volume integral equation of the problem.

The convergence analysis of the spectral approximation relies on establishing that the obtained discrete system coincides with the one for volume integral equation associated with the ML-periodic problem. This property is indeed surprising and one would not expect that it holds in general. In fact it is very specific to the choice of a trapezoidal rule to discretize the Floquet-Bloch transform. It has been also observed for finite element discretization in [4]. The optimality of this choice of discretization is not obvious and may for instance not be convenient in the neighborhood of the Wood anomaly (in the absorption free case).

Our discrete system can then be seen as a special re-arranging of the discretization of the  $ML$ -periodic problem (with a slightly different source term). One of the main advantages of this rearranging is that one obtain a linear growth of the computational cost with respect to the number of discretization points in the Floquet-Bloch variable. This is a gain of a logarithmic factor with respect to the application of the method in [20] to the ML periodic problem. Indeed this would be of interest for large (three dimensional) problems or when the absorption is small (requiring the use of a large discretization points in the Floquet-Bloch variable). We shall discuss this issue in a future work. In the present one we shall content ourselves with preliminary validating results in 2D setting of the problem.

Our paper is organized as follows. We first introduce the problem and some notation and results for the Floquet-Bloch transform. In order to ease the reading of the technical part treating the full discretization of the problem we introduce in Section 3 the method employed in [16] for quasi periodic problem. We complement known results with a uniform convergence result with respect to Fourier-Bloch variable. Section 4 constitutes the core of our work. We discuss in Section 4.1 the semi-discretization in the Floquet-Bloch variable and prove exponential convergence result. Section 4.2 is dedicated to a full discretization of the problem and associated convergence result. The last section is concerned with a rapid description of the algorithm and some numerical validating tests for the two dimensional problem.

## 2 Setting of the problem

### 2.1 Introduction of the problem and notation

We investigate the scattering problem for the case of locally perturbed unbounded periodic layers. The problem is formulated as: Given  $f \in L^2(\mathbb{R}^d)$ ,  $d = 2, 3$ , find solution  $u \in H_{loc}^1(\mathbb{R}^d)$  to the Helmholtz equation:

$$\Delta u + k^2 n u = f \quad \text{in } \mathbb{R}^d \quad (1)$$

where  $n \in L^\infty(\mathbb{R}^d)$  denotes the index of refraction and is such that  $n = n_p$  outside a compact domain  $D$  where  $n_p \in L^\infty(\mathbb{R}^d)$  is a periodic function with respect to the first  $d - 1$  variables of period  $L := (L_1, \dots, L_{d-1}) \in \mathbb{R}^{d-1}$ ,  $L_j > 0$ ,  $j = 1, \dots, d - 1$ .

If the wave number  $k$  is real, then (1) should be supplemented with a radiation condition (see for instance [1]). Although our numerical method can formally be extended to the case of real wave numbers  $k$ , the convergence proof requires some decay properties along the periodicity directions that we are only capable to prove in the case of complex wave numbers. This is why we shall restrict ourselves to the (coercive case) where

$$k^2 = k_0^2 + i\sigma$$

with  $k_0 \geq 0$  and  $\sigma > 0$  and assume that  $\text{Re } n(x) \geq c_0 > 0$  and  $\text{Im } n(x) \geq 0$  in  $\mathbb{R}^d$  for some constant  $c_0$ . In this case one easily check, using the Lax-Milgram theorem that (1) has a unique solution  $u \in H^1(\mathbb{R}^d)$  and that

$$\|u\|_{L^2(\mathbb{R}^d)} \leq \frac{1}{\sigma c_0} \|f\|_{L^2(\mathbb{R}^d)}, \quad \|\nabla u\|_{L^2(\mathbb{R}^d)} \leq \sqrt{\frac{k_0^2 \|n\|_\infty}{\sigma^2 c_0^2} + \frac{1}{\sigma c_0}} \|f\|_{L^2(\mathbb{R}^d)}. \quad (2)$$

We also remark that the solution  $u \in H^2(\mathbb{R}^d)$  (since  $\Delta u \in L^2(\mathbb{R}^d)$ ). For the well posedness of the problem in the case of real wave numbers, known results require some additional monotonicity properties of  $n$  with respect to  $x_d$  (see for instance [17]).

For a point  $x \in \mathbb{R}^d$  we shall use the decomposition  $x = (\bar{x}, x_d)$  with  $x_d \in \mathbb{R}$  and  $\bar{x} = (x_1, \dots, x_{d-1}) \in \mathbb{R}^{d-1}$ . For  $m = (m_1, \dots, m_{d-1}) \in \mathbb{Z}^{d-1}$  we denote by

$$\Omega_m := \left[ \left(m - \frac{1}{2}\right)L, \left(m + \frac{1}{2}\right)L \right] \times \mathbb{R},$$

where we use the notation  $\llbracket a, b \rrbracket := [a_1, b_1] \times \dots \times [a_{d-1}, b_{d-1}]$ . We shall also use the notation

$$\Omega^h := \mathbb{R}^{d-1} \times ]-h, h[ \text{ and } \Omega_m^h := \Omega^h \cap \Omega_m.$$

We shall assume that there exists  $h > 0$  such that

$$\text{supp}(n_p - 1) \subset \Omega^h \text{ and } \text{supp}(f) \subset \Omega^h.$$

Moreover, to further ease the presentation, we suppose that the support of  $n - n_p$  is strictly contained in  $\Omega_0^h$ . Indeed this can always be ensured by increasing the periodicity length. One can also work with a periodicity equal  $\Omega_0^h$  even if the support of  $n - n_p$  is larger  $\Omega_0^h$  but at the cost of additional non essential technicalities (See Remark 2.1 below).

**Remark 2.1.** For a point  $x_0 \in \mathbb{R}^d$ , the total field  $u$  generated by a point source at  $x_0$  is solution to

$$\Delta u + k^2 n u = -\delta_{x_0} \quad \text{in } \mathbb{R}^d. \quad (3)$$

Let  $u^i(\cdot; x_0)$  be the incident wave generated by the point source at  $x_0$ ,

$$u^i(\cdot; x_0) := \begin{cases} \frac{i}{4} H_0^{(1)}(k|\cdot - x_0|) & \text{in } \mathbb{R}^2, \\ \frac{e^{ik|\cdot - x_0|}}{4\pi|\cdot - x_0|} & \text{in } \mathbb{R}^3. \end{cases}$$

The scattering problem for a point source at  $x_0$  can be reformulated as finding  $u^s = u - u^i$  with  $u^s \in H^1(\mathbb{R}^d)$  such that

$$\Delta u^s + k^2 n u^s = -k^2(n-1)u^i \quad \text{in } \mathbb{R}^d. \quad (4)$$

Indeed problem (4) has the same form as (1) with  $f := -k^2(n-1)u^i$ .

## 2.2 Formulation of the problem using the Floquet-Bloch transform

**Definition 2.2.** A regular function  $u$  is called quasi-periodic with parameter  $\xi = (\xi_1, \dots, \xi_{d-1})$  and period  $L = (L_1, \dots, L_{d-1})$ , with respect to the first  $d-1$  variables (briefly denoted as  $\xi$ -quasi-periodic with period  $L$ ) if:

$$u(\bar{x} + (jL), x_d) = e^{i\xi \cdot (jL)} u(\bar{x}, x_d), \quad \forall j \in \mathbb{Z}^{d-1}.$$

We now introduce some notation for functional spaces that will be used in the sequel:

- $C_{\xi, L}^\infty(\mathbb{R}^d) := \{u \in C^\infty(\mathbb{R}^d); u \text{ is } \xi\text{-quasi-periodic with period } L\}$ ;
- $C_\xi^\infty(\Omega_0) := \{u|_{\Omega_0}; u \in C_{\xi, L}^\infty(\mathbb{R}^d)\}$ ;
- $H_\xi^s(\Omega_0)$  is the closure of  $C_\xi^\infty(\Omega_0)$  with respect to the  $H^s(\Omega_0)$  norm;
- $H_{\xi, L}^s(\mathbb{R}^d)$  is the extension of  $H_\xi^s(\Omega_0)$  functions to a  $\xi$ -quasi-periodic function with period  $L$  to all of  $\mathbb{R}^d$ .

For a regular function  $\varphi \in \mathcal{S}(\mathbb{R}^d)$ , the Floquet-Bloch transform of  $\varphi$  in the first  $d-1$  variables with period  $L \in \mathbb{R}^{d-1}$  is defined by:

$$\mathcal{F}\varphi(\bar{x}, x_d; \xi) := \sum_{n \in \mathbb{Z}^{d-1}} \varphi(\bar{x} + mL, x_d) e^{-i(mL) \cdot \xi}, \quad \forall \xi \in \left[-\frac{\pi}{L}, \frac{\pi}{L}\right], x \in \mathbb{R}^d. \quad (5)$$

$\mathcal{F}\varphi(\cdot; \xi)$  defined in (5) is  $\xi$ -quasi-periodic with period  $L$ . The inverse Floquet-Bloch transform  $\mathcal{F}^{-1}$  is given by

$$\varphi(\bar{x} + mL, x_d) = \frac{\llbracket L \rrbracket}{(2\pi)^{d-1}} \int_{\left[-\frac{\pi}{L}, \frac{\pi}{L}\right]} \mathcal{F}\varphi(\bar{x}, x_d; \xi) e^{i(mL) \cdot \xi} d\xi, \quad m \in \mathbb{Z}^{d-1}, x \in \mathbb{R}^d, \quad (6)$$

where  $\llbracket L \rrbracket := L_1 \cdots L_{d-1}$ . We shall also use the following notations for vector in  $\mathbb{Z}^{d-1}$ :

$$|L| := |L_1| + \cdots + |L_{d-1}|, \quad \|L\| := \sqrt{L_1^2 + \cdots + L_{d-1}^2}.$$

Operators on vectors such as multiplication or division should be understood as component-wise operations. Since  $\mathcal{S}(\mathbb{R}^d)$  is dense in  $L^2(\mathbb{R}^d)$  and

$$\|\mathcal{F}\varphi\|_{L^2(\Omega_0 \times \left[-\frac{\pi}{L}, \frac{\pi}{L}\right])} = \|\varphi\|_{L^2(\mathbb{R}^d)}$$

for  $\varphi \in \mathcal{S}(\mathbb{R}^d)$ , then the operator  $\mathcal{F}$  extends to an isometry between  $L^2(\mathbb{R}^d)$  and  $L^2(\Omega_0 \times \left[-\frac{\pi}{L}, \frac{\pi}{L}\right])$ . Furthermore, for all integers  $s > 0$ ,  $\mathcal{F}(H^s(\mathbb{R}^d)) = L^2\left(\left[-\frac{\pi}{L}, \frac{\pi}{L}\right], H_\xi^s(\Omega_0)\right)$  and for all  $u \in H^s(\mathbb{R}^d)$ :

$$\mathcal{F}(\partial_x^j u) = \partial_x^j \mathcal{F}(u), \quad \forall 0 \leq j \leq s. \quad (7)$$

This result can be found in [4]. We also refer to [13] for an extensive study of the Floquet-Bloch transform.

We now propose to use the Floquet-Bloch transform in the first  $d-1$  directions to study the spectral numerical approximation of problem (1). We first remark that since  $n_p$  is  $L$  periodic, then  $\mathcal{F}(n_p u) = n_p \mathcal{F}(u)$ . Moreover, if a function  $\varphi$  has a support in  $\Omega_0$ , then  $\mathcal{F}\varphi(\bar{x}, x_d; \xi) = \varphi(\bar{x}, x_d)$  for  $(\bar{x}, x_d) \in \Omega_0$ . Consequently, applying the Floquet-Bloch transform to equation (1) we get

$$\Delta(\mathcal{F}u)(\bar{x}, x_d; \xi) + k^2 n_p(\mathcal{F}u)(\bar{x}, x_d; \xi) + k^2(n - n_p)u(\bar{x}, x_d) = (\mathcal{F}f)(\bar{x}, x_d; \xi); \quad (\bar{x}, x_d) \in \Omega_0. \quad (8)$$

Let us set  $f_\xi := (\mathcal{F}f)(\cdot; \xi)$  and

$$\tilde{u}(\bar{x}, x_d; \xi) := \mathcal{F}u(\bar{x}, x_d; \xi) \quad (x, x_d) \in \Omega_0.$$

Then  $\tilde{u} \in L^2\left(\left[-\frac{\pi}{L}, \frac{\pi}{L}\right], H_\xi^2(\Omega_0)\right)$  and

$$\begin{cases} \Delta \tilde{u}(\cdot; \xi) + k^2 n_p \tilde{u}(\cdot; \xi) + k^2(n - n_p) \mathcal{M}(\tilde{u}) = f_\xi & \text{in } \Omega_0, \xi \in \left[-\frac{\pi}{L}, \frac{\pi}{L}\right] \\ \mathcal{M}(\tilde{u}) := \frac{\llbracket L \rrbracket}{(2\pi)^{d-1}} \int_{\left[-\frac{\pi}{L}, \frac{\pi}{L}\right]} \tilde{u}(\cdot; \xi) d\xi, \end{cases} \quad (9)$$

where the second equation comes from the inverse Floquet-Bloch transform expression

$$u(\bar{x} + mL, x_d) = \frac{\llbracket L \rrbracket}{(2\pi)^{d-1}} \int_{\left[-\frac{\pi}{L}, \frac{\pi}{L}\right]} \tilde{u}(\bar{x}, x_d; \xi) e^{i(mL) \cdot \xi} d\xi, \quad m \in \mathbb{Z}^{d-1}, (\bar{x}, x_d) \in \Omega_0. \quad (10)$$

One can also check that the converse is true, i.e. if  $(\mathcal{F}f) \in L^2(\Omega_0 \times \left[-\frac{\pi}{L}, \frac{\pi}{L}\right])$ , the  $u$  defined by (10) is in  $H^2(\mathbb{R}^d)$  and is solution to (1).

**Remark 2.3.** *If the support of  $n - n_p$  is contained in  $\Omega_{P^-, P^+}^h := \cup_{m \in [P^-, P^+]} \Omega_m^h$  for  $P_j^-, P_j^+ \geq 0$ ,  $j = 1, \dots, d$  then the third term of the first equation of system (9) has to be modified and the system becomes*

$$\begin{cases} \Delta \tilde{u}(\cdot; \xi) + k^2 n_p \tilde{u}(\cdot; \xi) + k^2 \sum_{m=-P^-}^{P^+} (n - n_p)(\cdot + mL) \mathcal{M}_m(\tilde{u}) e^{-i(mL) \cdot \xi} = f_\xi & \text{in } \Omega_0, \xi \in \left[-\frac{\pi}{L}, \frac{\pi}{L}\right] \\ \mathcal{M}_m(\tilde{u}) := \frac{\llbracket L \rrbracket}{(2\pi)^{d-1}} \int_{\left[-\frac{\pi}{L}, \frac{\pi}{L}\right]} \tilde{u}(\cdot; \xi) e^{i(mL) \cdot \xi} d\xi. \end{cases} \quad (11)$$

*The analysis below can be extended to treat this case without any major additional difficulty.*

### 3 Volume integral formulation of the $\xi$ -quasi-periodic problem

Before dealing with problem (9) where the equations for different values of  $\xi$  are coupled through  $\mathcal{M}(\tilde{u})$ , we shall first recall the principles of the spectral volume integral method for fixed  $\xi$  as in [16] for periodic media. This may be helpful in making the technical details of the general case treated in next section more digest. We shall also supplement the convergence results with a result on uniform convergence properties with respect to  $\xi$ . Although the latter is not necessary to treat problem (9), we thought that it has its own interest and would complement the picture for the use of this type of methods in periodic media. Some of the technical results presented will also be useful for Section 4.

#### 3.1 Setting of the volume integral equation

We first consider problem (9) for the case of periodic media, i.e.  $n = n_p$ . Let us set  $u_\xi := \mathcal{F}u(\cdot; \xi)$  and set  $f_\xi := \mathcal{F}f(\cdot; \xi)$ . Then we are led to consider the problem,  $u_\xi \in H_{\xi, L}^2(\mathbb{R}^d)$  verifying

$$\Delta u_\xi + k^2 n_p u_\xi = f_\xi \quad \text{in } \mathbb{R}^d \quad (12)$$

where  $f_\xi \in L_{\xi, L}^2(\mathbb{R}^d)$ . Indeed considering problem (12) in  $\mathbb{R}^d$  or in  $\Omega_0$  is equivalent. In the following we shall not distinguish  $u_\xi$  from its restriction to  $\Omega_0$ . We recall that

$$\|u_\xi\|_{H_\xi^1(\Omega_0)} \leq C \|f_\xi\|_{L^2(\Omega_0^h)}, \quad \text{where } C := \max \left\{ \frac{1}{\sigma c_0}, \sqrt{\frac{k_0^2 \|n\|_\infty}{\sigma^2 c_0^2} + \frac{1}{\sigma c_0}} \right\}. \quad (13)$$

Let  $G_\xi(\cdot) \in L_{\xi, L}^2(\mathbb{R}^d)$  be the  $\xi$ -quasi-periodic Green function of the Helmholtz equation, i.e. satisfying,

$$\Delta G_\xi + k^2 G_\xi(\cdot) = -\delta_0 \quad \text{in } \Omega_0. \quad (14)$$

Then  $G_\xi$  can be expressed as (see [12]),

$$G_\xi(x) = \frac{i}{2\llbracket L \rrbracket} \sum_{j \in \mathbb{Z}^{d-1}} \frac{1}{\beta_\xi(j)} \exp(i\alpha_\xi(j) \cdot \bar{x} + i\beta_\xi(j)|x_d|), \quad (\bar{x}, x_d) \in \mathbb{R}^d, \quad (15)$$

where

$$\alpha_\xi(j) := \xi + \frac{2\pi}{L} j = \left( \xi_1 + \frac{2\pi}{L_1} j_1, \dots, \xi_{d-1} + \frac{2\pi}{L_{d-1}} j_{d-1} \right), \quad \beta_\xi(j)^2 := k^2 - \|\alpha_\xi(j)\|^2, \quad \text{Im } \beta_\xi(j) \geq 0.$$

(Remark that  $\beta_\xi(j) \neq 0$  for all  $j$  and  $\xi$  since  $\sigma > 0$ .) We now consider the volume potential  $V_\xi$  on  $L^2(\Omega_0^h)$  with kernel  $G_\xi$  defined by

$$V_\xi g(x) := \int_{\Omega_0^h} G_\xi(x-y) g(y) dy, \quad x \in \mathbb{R}^d. \quad (16)$$

**Lemma 3.1.** (See for instance [6, 16]) *The volume potential  $V_\xi$  is a linear bounded operator from  $L^2(\Omega_0^h)$  into  $H_{\xi, L}^2(\mathbb{R}^d)$ . Moreover, for all  $g \in L^2(\Omega_0^h)$ , the potential  $w := V_\xi g \in H_{\xi, L}^2(\mathbb{R}^d)$  and is the unique solution to  $\Delta w + k^2 w = -g$  in  $\mathbb{R}^d$ .*

Thus, if we set  $w := V_k(k^2(n_p - 1)u_\xi - f_\xi)$ , then from Lemma 3.1,  $w \in H_{\xi, L}^2(\mathbb{R}^d)$  and satisfies

$$\Delta w + k^2 w = -k^2(n_p - 1)u_\xi + f_\xi \quad \text{in } \mathbb{R}^d.$$

Therefore  $w = u_\xi$  in  $\mathbb{R}^d$  and  $u_\xi \in L^2(\Omega_0^h)$  and satisfies

$$u_\xi = k^2 V_\xi((n_p - 1)u_\xi) - V_\xi f_\xi \quad \text{in } L^2(\Omega_0^h). \quad (17)$$

Conversely, if (17) is verified, then obviously  $u_\xi$  can be obtained in all  $\mathbb{R}^d$  using the expression of  $w$ . The numerical scheme we shall consider is based on the spectral discretization of (17). In order to be efficient, by the use of FFT in evaluating the matrix-vector product, we need to periodize the equation in the direction  $x_d$ . This is the step we shall discuss now.

### 3.2 Periodization of the integral equation

Let  $R \in \mathbb{R}$  such that  $R > 2h$ , we define  $G_\xi^R$  as

$$G_\xi^R(x) := G_\xi(x), \quad \forall x = (\bar{x}, x_d) \in \mathbb{R}^{d-1} \times ]-R, R[, \quad (18)$$

and extend  $G_\xi^R$  to all  $\mathbb{R}^d$  as a  $2R$ -periodic function with respect to  $x_d$ . We then define the periodized volume potential  $V_\xi^R : L^2(\Omega_0^h) \rightarrow L^2(\Omega_0^R)$  using the same expression as (16) where the kernel  $G_\xi$  is replaced by  $G_\xi^R$  and consider the periodized volume integral equation,  $u_\xi^R \in L^2(\Omega_0^R)$ ,

$$u_\xi^R = k^2 V_\xi^R((n_p - 1)u_\xi^R) - V_\xi^R f_\xi \quad \text{in } L^2(\Omega_0^R). \quad (19)$$

We then have the following result.

**Lemma 3.2.** 1. For all  $g \in L^2(\Omega_0^h)$ ,  $V_\xi^R g = V_\xi g$  in  $\mathbb{R}^d$ .

2. For for all  $f_\xi \in L^2(\Omega_0^h)$ , equation (19) has a unique solution  $u_\xi^R \in L^2(\Omega_0^R)$  and we have

$$u_\xi^R = u_\xi \text{ in } \Omega_0^h$$

where  $u_\xi \in L^2(\Omega_0^h)$  is the unique solution of (17).

*Proof.* The equality  $V_\xi^R g = V_\xi g$  in  $\mathbb{R}^d$  holds for all  $g \in L^2(\Omega_0^h)$  since  $G_\xi^R$  and  $G_\xi$  coincide in  $\Omega^R$ ,  $|x_d - y_d| < R$  for all  $(x, y) \in \Omega_0^h \times \Omega_0^h$  and  $R > 2h$ .

Let  $f_\xi \in L^2(\Omega_0^h)$ . Since the support of  $n_p - 1$  is included in  $\Omega^h$ , we deduce from the first point that a solution  $u_\xi^R \in L^2(\Omega_0^R)$  of (19) is such that  $u_\xi^R|_{\Omega_0^h} \in L^2(\Omega_0^h)$  and verifies (17). Since  $u_\xi^R|_{\Omega_0^h} = 0$  and  $f_\xi = 0$  imply  $u_\xi^R = 0$  in  $\Omega_0^R$  we easily conclude from the well-posedness of (17) that (19) has at most one solution. In addition, a solution of (19) can be constructed as

$$u_\xi^R = u_\xi \text{ in } \Omega_0^h \text{ and } u_\xi^R = k^2 V_\xi^R((n_p - 1)u_\xi) - V_\xi^R f_\xi \text{ in } \Omega_0^R \setminus \Omega_0^h$$

where  $u_\xi \in L^2(\Omega_0^h)$  is the solution of (17). □

The spectral discretization method is based on the Fourier basis defined as

$$\varphi_\xi^j(x) := \frac{1}{\sqrt{2\lfloor L \rfloor R}} \exp\left(i\left(\xi + \frac{2\pi}{L}\bar{j}\right) \cdot \bar{x} + i\frac{\pi}{R}j_d x_d\right), \quad j = (j_1, \dots, j_d) = (\bar{j}, j_d) \in \mathbb{Z}^d. \quad (20)$$

For all  $u \in L^2(\Omega_0^R)$ ,  $u = \sum_{j \in \mathbb{Z}^d} \hat{u}(j; \xi) \varphi_\xi^j(x)$  where  $\hat{u}(j; \xi)$  to denotes the  $j$ -th coefficient of  $u$  with respect to this Fourier basis, which is defined as:

$$\hat{u}(j; \xi) := \int_{\Omega_0^R} u(x) \bar{\varphi}_\xi^j(x) dx. \quad (21)$$

If  $u \in H_\xi^s(\Omega_0^R)$ ,  $0 \leq s < +\infty$  then an equivalent norm in  $H_\xi^s(\Omega_0^R)$  is given by

$$\|u\|_{H_\xi^s(\Omega_0^R)}^2 := \sum_{j \in \mathbb{Z}^d} \left(1 + \left\| \frac{\bar{j}}{L} \right\|^2 + \left\| \frac{j_d}{R} \right\|^2\right)^s |\hat{u}(j; \xi)|^2. \quad (22)$$

Remark that since  $\|u\|_{L^2(\Omega_0^R)} = \|e^{i\xi x} u\|_{L^2(\Omega_0^R)}$ , using Plancherel theorem we conclude the norm in  $L^2(\Omega_0^R)$  defined by (22) (for  $s = 0$ ) is independent from  $\xi$ , i.e.,

$$\|u\|_{L^2(\Omega_0^R)}^2 = \sum_{j \in \mathbb{Z}^d} |\hat{u}(j; 0)|^2 = \sum_{j \in \mathbb{Z}^d} |\hat{u}(j; \xi)|^2, \quad \forall \xi. \quad (23)$$

We can easily establish the link between the Fourier coefficients with respect to  $\xi$ -quasi-periodic Fourier basis and periodic Fourier basis ( $\xi \in \llbracket -\frac{\pi}{L}, \frac{\pi}{L} \rrbracket$ ,  $\xi^\ell \neq 0, \forall \ell = 1, \dots, d$ ) as

$$\widehat{u}(t; 0) = 2R \sum_{j \in \mathbb{Z}^d} \widehat{u}(\bar{j}, t_d; \xi) \prod_{\ell=1}^{d-1} \frac{2 \cos(\pi(j_\ell - t_\ell)) \sin(\xi^\ell \frac{L_\ell}{2})}{\xi^\ell + \frac{2\pi}{L_\ell}(j_\ell - t_\ell)} \quad (24)$$

where here  $\xi^\ell$  denotes the component  $\ell$  of  $\xi$ .

Using (15), one can easily compute the Fourier coefficients of  $G_\xi^R$  and get

$$\widehat{G_\xi^R}(j; \xi) = \gamma \frac{e^{i\beta_\xi(\bar{j})R} \cos(\pi j_d) - 1}{\beta_\xi(\bar{j})^2 - (\frac{\pi}{R}j_d)^2} = \gamma \frac{e^{i\beta_\xi(\bar{j})R} \cos(\pi j_d) - 1}{k^2 - \|\alpha_\xi(\bar{j})\|^2 - (\frac{\pi}{R}j_d)^2} \quad (25)$$

with  $\gamma := \frac{1}{\sqrt{2\llbracket L \rrbracket}R}$ . Since  $k^2 = k_0^2 + i\sigma$  with positive  $\sigma$  and  $\text{Im} \beta_\xi(\bar{j}) \geq 0$ , we have the estimate, for sufficiently small  $\varepsilon$  such that  $\sigma - \varepsilon(k_0^2 + 2|\xi|^2) > 0$ ,

$$|\widehat{G_\xi^R}(j; \xi)| \leq \gamma \frac{2}{\sigma + \varepsilon(-k_0^2 + \|\alpha_\xi(\bar{j})\|^2 + (\frac{\pi}{R}j_d)^2)}.$$

Choosing  $\varepsilon$  such that  $\sigma - \varepsilon(k_0^2 + 2\|\pi/L\|^2) > 0$ , we infer the existence of a constant  $C$  independent from  $\xi$ ,  $L$ ,  $R$  and  $j$  such that

$$|\widehat{G_\xi^R}(j; \xi)| \leq C \frac{\gamma}{\left(1 + \left\|\frac{\bar{j}}{L}\right\|^2 + \left\|\frac{j_d}{R}\right\|^2\right)}. \quad (26)$$

Since  $V_\xi^R$  is a convolution operator

$$\widehat{V_\xi^R g}(j; \xi) = \frac{1}{\gamma} \widehat{G_\xi^R}(j; \xi) \widehat{g}(j; \xi).$$

We then immediately get from (26) the following result.

**Lemma 3.3.** *Let  $s \in \mathbb{N}$ . The operator  $V_\xi^R : H_\xi^s(\Omega_0^R) \rightarrow H_\xi^{2+s}(\Omega_0^R)$  is continuous. Moreover, there exists a constant  $C$  independent of  $\xi$  such that*

$$\|V_\xi^R g\|_{H_\xi^{2+s}(\Omega_0^R)} \leq C \|g\|_{H_\xi^s(\Omega_0^R)} \quad \forall g \in H_\xi^s(\Omega_0^R).$$

We then obtain the following uniform regularity result.

**Proposition 3.4.** *Let  $f_\xi \in L^2(\Omega_0^h)$  and  $u_\xi^R \in L^2(\Omega_0^R)$  be the solution of (19). Then there exists a positive constant  $C$  independent from  $\xi$  and  $f_\xi$  such that*

$$\|u_\xi^R\|_{H^2(\Omega_0^R)} \leq C \|f_\xi\|_{L^2(\Omega_0^h)}. \quad (27)$$

*Proof.* Thanks to Lemma 3.2,  $u_\xi^R = u_\xi$  in  $\Omega_0^h$  where  $u_\xi$  denotes the solution of (17). Then, using (13),

$$\|u_\xi^R\|_{L^2(\Omega_0^h)} \leq C_1 \|f_\xi\|_{L^2(\Omega_0^h)}$$

for some constant  $C_1$  independent from  $\xi$ . Using equation (19), Lemma (3.3) and the fact that  $\text{supp}(1-n_p) \subset \Omega_0^h$ , we get the existence of a constant  $C_2$  independent from  $\xi$  such that

$$\|u_\xi^R\|_{H^2(\Omega_0^h)} \leq C_2 (\|f_\xi\|_{L^2(\Omega_0^h)} + \|u_\xi^R\|_{L^2(\Omega_0^h)}).$$

The result of the proposition then immediately follows from the two inequalities.  $\square$



Let us introduce for later use  $T_\xi^R : L^2(\Omega_0^R) \rightarrow L^2(\Omega_0^R)$  defined by

$$T_\xi^R g := k^2 V_\xi^R ((n_p - 1)g) \quad (28)$$

and  $A_\xi^R : L^2(\Omega_0^R) \rightarrow L^2(\Omega_0^R)$  defined by

$$A_\xi^R g := g - T_\xi^R g. \quad (29)$$

Then, as a straightforward consequences of Lemma 3.3, we have:

**Lemma 3.5.** *The operator  $T_\xi^R : L^2(\Omega_0^R) \rightarrow L^2(\Omega_0^R)$  is compact and there exists a positive constant  $C$  independent of  $\xi$  such that:*

$$\|T_\xi^R g\|_{H_\xi^2(\Omega_0^R)} \leq C \|g\|_{L^2(\Omega_0^R)}, \quad \forall g \in L^2(\Omega_0^R). \quad (30)$$

An immediate corollary of this Lemma and Lemma 3.2 is the following.

**Corollary 3.6.** *The operator  $A_\xi^R : L^2(\Omega_0^R) \rightarrow L^2(\Omega_0^R)$  is bounded and injective. Moreover, there exists a positive constant  $C$  independent of  $\xi$  such that*

$$\|A_\xi^R g\|_{L^2(\Omega_0^R)} \leq C \|g\|_{L^2(\Omega_0^R)}.$$

### 3.3 Spectral approximation of problem (19)

Let  $N = (N_1, \dots, N_d) \in \mathbb{N}^d$  and  $\mathbb{Z}_N^d := \{j = (j_1, \dots, j_d) \in \mathbb{Z}^d \mid -\frac{N_\ell}{2} + 1 \leq j_\ell \leq \frac{N_\ell}{2}, \ell = 1, \dots, d\}$ , we define the discrete space  $\mathcal{T}_\xi^N := \text{span}\{\varphi_\xi^j; j \in \mathbb{Z}_N^d\}$  and the orthogonal projection

$$P_\xi^N : L^2(\Omega_0^R) \longrightarrow \mathcal{T}_\xi^N, \quad u \longmapsto P_\xi^N(u) := \sum_{j \in \mathbb{Z}_N^d} \widehat{u}(j; \xi) \varphi_\xi^j. \quad (31)$$

Let  $u_\xi^R$  be the solution of (19). We now consider  $u_\xi^{R,N} \in \mathcal{T}_\xi^N$  an approximation of  $u_\xi^R$ , which is solution to following variational equation

$$\left\langle u_\xi^{R,N} - k^2 V_\xi^R \left( (n_p - 1) u_\xi^{R,N} \right), v_\xi^N \right\rangle_{L^2(\Omega_0^R)} = \left\langle -V_\xi^R f_\xi, v_\xi^N \right\rangle_{L^2(\Omega_0^R)}, \quad \forall v_\xi^N \in \mathcal{T}_\xi^N. \quad (32)$$

The next theorem shows that the convergence rate of this method using standard convergence result of Galerkin discretization, is independent from  $\xi$ .

**Theorem 3.7.** *There exists  $N_0 \in \mathbb{N}^d$  independent from  $\xi$  such that problem (32) admits a unique solution for all  $N > N_0$  and  $f_\xi \in L^2(\Omega_0^R)$ . Moreover, there is a positive constant  $C > 0$  independent of  $\xi$  and  $N$  such that,*

$$\|u_\xi^R - u_\xi^{R,N}\|_{L^2(\Omega_0^R)} \leq C \inf_{v_\xi^N \in \mathcal{T}_\xi^N} \|u_\xi^R - v_\xi^N\|_{L^2(\Omega_0^R)}. \quad (33)$$

*Proof.* We concentrate on the second claim since the first one follows from the similar arguments. The solution  $u_\xi^R$  of (19) satisfies

$$\left\langle A_\xi^R u_\xi^R, v_\xi \right\rangle_{L^2(\Omega_0^R)} = \left\langle -V_\xi^R f_\xi, v_\xi \right\rangle_{L^2(\Omega_0^R)}, \quad \forall v_\xi \in L^2(\Omega_0^R). \quad (34)$$

We then obtain

$$\left\langle A_\xi^R (u_\xi^R - u_\xi^{R,N}), v_\xi^N \right\rangle_{L^2(\Omega_0^R)} = 0, \quad \forall v_\xi^N \in \mathcal{T}_\xi^N. \quad (35)$$

Since  $u_\xi^{R,N} - v_\xi^N \in \mathcal{T}_\xi^N$ , we get

$$\begin{aligned} \left\langle A_\xi^R (u_\xi^R - u_\xi^{R,N}), u_\xi^R - u_\xi^{R,N} \right\rangle_{L^2(\Omega_0^R)} &= \left\langle A_\xi^R (u_\xi^R - u_\xi^{R,N}), u_\xi^R - v_\xi^N \right\rangle_{L^2(\Omega_0^R)} \\ &\leq C \|u_\xi^R - u_\xi^{R,N}\|_{L^2(\Omega_0^R)} \|u_\xi^R - v_\xi^N\|_{L^2(\Omega_0^R)}. \end{aligned} \quad (36)$$

We shall prove for all  $\varepsilon > 0$ , there exists  $N_0 > 0$  independent from  $\xi$  such that  $\forall N > N_0, \forall \xi \in \llbracket -\frac{\pi}{L}, \frac{\pi}{L} \rrbracket$ :

$$\left\langle T_\xi^R(u_\xi^R - u_\xi^{R,N}), u_\xi^R - u_\xi^{R,N} \right\rangle_{L^2(\Omega_0^R)} \leq \varepsilon \|u_\xi^R - u_\xi^{R,N}\|_{L^2(\Omega_0^R)}^2. \quad (37)$$

Let us define

$$w_\xi^N := \frac{u_\xi^R - u_\xi^{R,N}}{\|u_\xi^R - u_\xi^{R,N}\|_{L^2(\Omega_0^R)}} \quad (38)$$

and suppose that (37) is not true. Then there exists a sequence  $N(n) \rightarrow \infty$  such that

$$\left\langle T_{\xi_n}^R w_{\xi_n}^{N(n)}, w_{\xi_n}^{N(n)} \right\rangle_{L^2(\Omega_0^R)} \geq \varepsilon. \quad (39)$$

for some  $\xi_n \in \llbracket -\frac{\pi}{L}, \frac{\pi}{L} \rrbracket$ . To shorten the notation, we set  $w_n := w_{\xi_n}^{N(n)}$ . Since  $(\xi_n)_{n \in \mathbb{N}} \subset \llbracket -\frac{\pi}{L}, \frac{\pi}{L} \rrbracket$  then there is a sub-sequence, also denoted  $(\xi_n)_{n \in \mathbb{N}}$ , that converges to  $\xi \in \llbracket -\frac{\pi}{L}, \frac{\pi}{L} \rrbracket$ . Furthermore,  $(w_n)$  is bounded in  $L^2(\Omega_0^R)$  then there exist a sub-sequence, also denoted  $(w_n)$  that converges weakly to  $w$  in  $L^2(\Omega_0^R)$ . Let  $v \in L^2(\Omega_0^R)$  and set  $v_n := P_{\xi_n}^{N(n)} v \in \mathcal{T}_{\xi_n}^{N(n)}$ . Then

$$\begin{aligned} \langle A_\xi^R w_n, v \rangle_{L^2(\Omega_0^R)} &= \langle (A_\xi^R - A_{\xi_n}^R) w_n, v \rangle_{L^2(\Omega_0^R)} + \langle A_{\xi_n}^R w_n, v_n \rangle_{L^2(\Omega_0^R)} + \langle A_{\xi_n}^R w_n, v - v_n \rangle_{L^2(\Omega_0^R)} \\ &= \langle (A_\xi^R - A_{\xi_n}^R) w_n, v \rangle_{L^2(\Omega_0^R)} + \langle A_{\xi_n}^R w_n, v - v_n \rangle_{L^2(\Omega_0^R)} \end{aligned}$$

where we used (35) for the last equality. Observe that, using Corollary 3.6,

$$\left| \langle A_{\xi_n}^R w_n, v - v_n \rangle_{L^2(\Omega_0^R)} \right| \leq \|A_{\xi_n}^R\|_{L^2(\Omega_0^R) \rightarrow L^2(\Omega_0^R)} \|v - v_n\|_{L^2(\Omega_0^R)} \xrightarrow{n \rightarrow \infty} 0 \quad (40)$$

and from Lemma 3.10 we have,

$$\begin{aligned} \left| \langle (A_\xi^R - A_{\xi_n}^R) w_n, v \rangle_{L^2(\Omega_0^R)} \right| &= \left| \langle (T_{\xi_n}^R - T_\xi^R) w_n, v \rangle_{L^2(\Omega_0^R)} \right| \\ &\leq \|T_{\xi_n}^R - T_\xi^R\|_{L^2(\Omega_0^R) \rightarrow L^2(\Omega_0^R)} \|v\|_{L^2(\Omega_0^R)} \xrightarrow{n \rightarrow \infty} 0. \end{aligned} \quad (41)$$

Therefore, taking the limit as  $n \rightarrow +\infty$  we get

$$\langle A_\xi^R w, v \rangle_{L^2(\Omega_0^R)} = 0, \quad \forall v \in L^2(\Omega_0^R). \quad (42)$$

Since  $A_\xi^R$  is injective (see Corollary 3.6), we deduce that  $w = 0$ . We now observe that

$$\langle T_{\xi_n}^R w_n, w_n \rangle_{L^2(\Omega_0^R)} = \langle (T_{\xi_n}^R - T_\xi^R) w_n, w_n \rangle_{L^2(\Omega_0^R)} + \langle T_\xi^R w_n, w_n \rangle_{L^2(\Omega_0^R)}$$

where

$$\left| \langle (T_{\xi_n}^R - T_\xi^R) w_n, w_n \rangle_{L^2(\Omega_0^R)} \right| \leq \|T_{\xi_n}^R - T_\xi^R\|_{L^2(\Omega_0^R) \rightarrow L^2(\Omega_0^R)} \xrightarrow{n \rightarrow \infty} 0,$$

and

$$\langle T_\xi^R w_n, w_n \rangle_{L^2(\Omega_0^R)} \xrightarrow{n \rightarrow \infty} 0,$$

since  $T_\xi^R$  is compact. We then conclude that

$$\langle T_{\xi_n}^R w_n, w_n \rangle_{L^2(\Omega_0^R)} \xrightarrow{n \rightarrow \infty} \langle T_\xi^R w, w \rangle_{L^2(\Omega_0^R)} = 0, \quad (43)$$

which contradicts (39) and proves (37). Now choose  $\varepsilon = \frac{1}{2}$  and recall that  $A_\xi^R = I - T_\xi^R$ . We obtain from (36)

$$\frac{1}{2} \|u_\xi^R - u_\xi^{R,N}\|_{L^2(\Omega_0^R)}^2 \leq \left\langle A_\xi^R (u_\xi^R - u_\xi^{R,N}), u_\xi^R - u_\xi^{R,N} \right\rangle \leq C \|u_\xi^R - u_\xi^{R,N}\| \|u_\xi^R - u_\xi^{R,N}\| \quad (44)$$

which implies

$$\|u_\xi^R - u_\xi^{R,N}\|_{L^2(\Omega_0^R)} \leq C \inf_{v_\xi^N \in \mathcal{T}_\xi^N} \|u_\xi^R - v_\xi^N\|_{L^2(\Omega_0^R)}. \quad (45)$$

□

It is possible to characterize the convergence rate due to the regularity result of Proposition 3.4 and the following classical interpolation result (See for instance [18]). We here make precise the dependence on  $L$  and  $R$  for later use in the analysis of the discretization of the full problem.

**Lemma 3.8.** *If  $u \in H_\xi^\mu(\Omega_0^R)$ , ( $\mu \in \mathbb{R}$ ) then*

$$\|u - P_\xi^N u\|_{H_\xi^\lambda(\Omega_0^R)} \leq \left( \sum_{\ell=1}^{d-1} \frac{1}{\left(1 + \frac{N_\ell^2}{L^2}\right)^{\mu-\lambda}} + \frac{1}{\left(1 + \frac{N_d^2}{R^2}\right)^{\mu-\lambda}} \right)^{1/2} \|u\|_{H_\xi^\mu(\Omega_0^R)}, \quad \text{for all } \lambda \leq \mu. \quad (46)$$

*Proof.* We give the proof for the reader's convenience. For  $j \in \mathbb{Z}^d$  let us denote by  $I(j) \in \mathbb{Z}^d$

$$I(j)_\ell = \frac{j_\ell}{L}; \quad \ell = 1, \dots, d-1 \quad \text{and} \quad I(j)_d = \frac{j_d}{R}.$$

$$\begin{aligned} \|u - P_\xi^N u\|_{H_\xi^\lambda}^2 &= \sum_{j \in \mathbb{Z}^d \setminus \mathbb{Z}_N^d} (1 + \|I(j)\|^2)^\lambda |\widehat{u}(j; \xi)|^2 = \sum_{\ell=1}^d \sum_{j \in \mathbb{Z}^d, |j_\ell| > N_\ell} (1 + \|I(j)\|^2)^\lambda |\widehat{u}(j; \xi)|^2 \\ &\leq \sum_{\ell=1}^d \frac{1}{(1 + I(N)_\ell^2)^{\mu-\lambda}} \sum_{j \in \mathbb{Z}^d, |j_\ell| > N_\ell} (1 + \|I(j)\|^2)^\mu |\widehat{u}(j; \xi)|^2 \\ &\leq \sum_{\ell=1}^d \frac{1}{(1 + I(N)_\ell^2)^{\mu-\lambda}} \|u\|_{H_\xi^\mu}^2, \end{aligned}$$

which proves the Lemma.  $\square$

Combining Proposition 3.4, Lemma 3.8 and Theorem 3.7 we obtain the following theorem.

**Theorem 3.9.** *There exists  $N_0 \in \mathbb{N}^d$  and a positive constant  $C > 0$  that are independent of  $\xi$  such that,*

$$\|u_\xi^R - u_{\xi'}^{R,N}\|_{L^2(\Omega_0^R)} \leq C \left( \sum_{\ell=1}^d \frac{1}{(1 + N_\ell^2)^2} \right)^{1/2} \|f_\xi\|_{L^2(\Omega_0^h)}. \quad (47)$$

for all  $N \geq N_0$ .

To prove Theorem 3.7 we used some auxiliary results given in Lemma 3.10 and Lemma 3.11 below.

**Lemma 3.10.** *The operator  $T_\xi^R$  in Lemma 3.5 is uniform Lipschitz continuous with respect to  $\xi$ , i.e.*

$$\|T_\xi^R - T_{\xi'}^R\|_{L^2(\Omega_0^R) \rightarrow L^2(\Omega_0^R)} \leq C \|\xi - \xi'\| \quad (48)$$

for some constant  $C > 0$  independent of  $\xi, \xi' \in \llbracket -\frac{\pi}{L}, \frac{\pi}{L} \rrbracket$ .

*Proof.* To prove this lemma we use the norm in  $L^2(\Omega_0^R)$  defined in (23). We first estimate the Fourier coefficients of  $G_\xi^R - G_{\xi'}^R$  with respect to the periodic Fourier basis. For  $j = (\bar{j}, j_d)$ , from (24) and (25) we have

$$\widehat{G_\xi^R}(t; 0) = 2R\gamma \sum_{j \in \mathbb{Z}^d} \frac{e^{i\beta_\xi(\bar{j})R} \cos(\pi j_d) - 1}{k^2 - \|\alpha_\xi(\bar{j})\|^2 - (\frac{\pi}{R}j_d)^2} \prod_{\ell=1}^{d-1} \frac{2 \cos(\pi(j_\ell - t_\ell)) \sin(\xi_\ell \frac{L_\ell}{2})}{\xi_\ell + \frac{2\pi}{L_\ell}(j_\ell - t_\ell)}. \quad (49)$$

(In this proof,  $\xi_\ell$  is used to indicate the  $\ell$ -component of  $\xi$ ,  $\ell = 1, \dots, d$ ). Then,

$$\begin{aligned} \left| \widehat{G_\xi^R}(t; 0) - \widehat{G_{\xi'}^R}(t; 0) \right| &= 2R\gamma \left| \sum_{j \in \mathbb{Z}^d} \frac{e^{i\beta_\xi(\bar{j})R} \cos(\pi j_d) - 1}{k^2 - \|\alpha_\xi(\bar{j})\|^2 - (\frac{\pi}{R}j_d)^2} \prod_{\ell=1}^{d-1} \frac{2 \cos(\pi(j_\ell - t_\ell)) \sin(\xi_\ell \frac{L_\ell}{2})}{\xi_\ell + \frac{2\pi}{L_\ell}(j_\ell - t_\ell)} \right. \\ &\quad \left. - \sum_{j \in \mathbb{Z}^d} \frac{e^{i\beta_{\xi'}(\bar{j})R} \cos(\pi j_d) - 1}{k^2 - \|\alpha_{\xi'}(\bar{j})\|^2 - (\frac{\pi}{R}j_d)^2} \prod_{\ell=1}^{d-1} \frac{2 \cos(\pi(j_\ell - t_\ell)) \sin(\xi'_\ell \frac{L_\ell}{2})}{\xi'_\ell + \frac{2\pi}{L_\ell}(j_\ell - t_\ell)} \right| \end{aligned} \quad (50)$$

To simplify notation we set

$$A_j(\xi) := \frac{e^{i\beta_\xi(\bar{j})R} \cos(\pi j d) - 1}{k^2 - \|\alpha_\xi(\bar{j})\|^2 - (\frac{\pi}{R} j d)^2} \quad \text{and} \quad B_j(\xi) := \prod_{\ell=1}^{d-1} \frac{2 \cos(\pi(j_\ell - t_\ell)) \sin(\xi_\ell \frac{L_\ell}{2})}{\xi_\ell + \frac{2\pi}{L_\ell}(j_\ell - t_\ell)}.$$

So we can simply write (50) as

$$\left| \widehat{G}_\xi^R(t; 0) - \widehat{G}_{\xi'}^R(t; 0) \right| = 2R\gamma \left| \sum_{j \in \mathbb{Z}^d} A_j(\xi) B_j(\xi) - A_j(\xi') B_j(\xi') \right|.$$

We now can prove that there exists positive constants  $C_1, C_2, C_3, C_4$  independent of  $\xi, \xi'$  such that

$$\begin{aligned} |A_j(\xi) - A_j(\xi')| &\leq C_1 \frac{\|\xi - \xi'\|}{1 + \|j\|^2} \quad \text{and} \quad |A_j(\xi)| \leq \frac{C_3}{1 + \|j\|^2} \\ |B_j(\xi) - B_j(\xi')| &\leq C_2 \frac{\|\xi - \xi'\|}{\prod(-\frac{1}{2} + |j_\ell - t_\ell|)}, \quad \text{and} \quad |B_j(\xi)| \leq \frac{C_4}{\prod(-\frac{1}{2} + |j_\ell - t_\ell|)} \end{aligned}$$

So we finally have

$$\begin{aligned} |A_j(\xi) B_j(\xi) - A_j(\xi') B_j(\xi')| &= |(A_j(\xi) - A_j(\xi')) B_j(\xi) + A_j(\xi') (B_j(\xi) - B_j(\xi'))| \\ &\leq |A_j(\xi) - A_j(\xi')| |B_j(\xi)| + |A_j(\xi')| |B_j(\xi) - B_j(\xi')| \\ &\leq \frac{C \|\xi - \xi'\|}{(1 + \|j\|^2) \prod_{\ell=1}^{d-1} (-\frac{1}{2} + |j_\ell - t_\ell|)} \end{aligned}$$

Since

$$\sum_{j \in \mathbb{Z}_N^d} \frac{1}{(1 + \|j\|^2) \prod_{\ell=1}^{d-1} (-\frac{1}{2} + |j_\ell - t_\ell|)} \leq C_0 \quad (51)$$

So,

$$2R\gamma \left| \sum_{j \in \mathbb{Z}_N^d} A_j(\xi) B_j(\xi) - A_j(\xi') B_j(\xi') \right| \leq 2R\gamma \sum_{j \in \mathbb{Z}^d} |A_j(\xi) B_j(\xi) - A_j(\xi') B_j(\xi')| \leq 2R\gamma C C_0 \|\xi - \xi'\|$$

We then obtain

$$|\widehat{G}_\xi^R(t; 0) - \widehat{G}_{\xi'}^R(t; 0)| \leq C \|\xi - \xi'\| \quad (52)$$

for some constant  $C$  independent from  $\xi$  and  $\xi'$ . Since  $V_\xi^R - V_{\xi'}^R$  is a convolution operator

$$(V_\xi^R - V_{\xi'}^R)g(t; 0) = \frac{1}{\gamma} (\widehat{G}_\xi^R(t; 0) - \widehat{G}_{\xi'}^R(t; 0)) \widehat{g}(t; 0).$$

We then have

$$\|(T_\xi^R - T_{\xi'}^R)g\|_{L^2(\Omega_0^R)} = \|(V_\xi^R - V_{\xi'}^R)(n-1)g\|_{L^2(\Omega_0^R)} \leq C \|\xi - \xi'\| \|n\|_{L^\infty} \|g\|_{L^2(\Omega_0^R)}, \quad (53)$$

which proves the lemma.  $\square$

**Lemma 3.11.** *Let  $\xi \in [-\frac{\pi}{L}, \frac{\pi}{L}]$  and a sequence  $(\xi_N)_N$  such that  $\xi_N \rightarrow \xi$  as  $N \rightarrow +\infty$  then*

$$\|v - P_{\xi_N}^N v\|_{L^2(\Omega_0^R)} \xrightarrow{n \rightarrow \infty} 0, \quad \forall v \in L^2(\Omega_0^R). \quad (54)$$

*Proof.* We first see that

$$\|v - P_{\xi_N}^N v\|_{L^2(\Omega_0^R)} \leq \|v - P_{\xi}^N v\|_{L^2(\Omega_0^R)} + \|P_{\xi}^N v - P_{\xi_N}^N v\|_{L^2(\Omega_0^R)}. \quad (55)$$

Since  $\|v - P_{\xi}^N v\|_{L^2(\Omega_0^R)} \rightarrow 0$  as  $N \rightarrow \infty$ , it remains to prove that

$$\|P_{\xi}^N v - P_{\xi_N}^N v\|_{L^2(\Omega_0^R)} \rightarrow 0 \quad \text{as } N \rightarrow \infty.$$

From (31) we have that

$$\begin{aligned} \|P_{\xi}^N v - P_{\xi_N}^N v\|_{L^2(\Omega_0^R)} &= \left\| \sum_{m \in \mathbb{Z}_N} \widehat{v}(m; \xi) \varphi_{\xi}^m - \widehat{v}(m; \xi_N) \varphi_{\xi_N}^m \right\|_{L^2(\Omega_0^R)} \\ &\leq \left\| \sum_{m \in \mathbb{Z}_N} (\widehat{v}(m; \xi) - \widehat{v}(m; \xi_N)) \varphi_{\xi}^m \right\|_{L^2(\Omega_0^R)} + \left\| \sum_{m \in \mathbb{Z}_N} \widehat{v}(m; \xi_N) (\varphi_{\xi}^m - \varphi_{\xi_N}^m) \right\|_{L^2(\Omega_0^R)}. \end{aligned}$$

We observe that

$$\begin{aligned} \left\| \sum_{m \in \mathbb{Z}_N} (\widehat{v}(m; \xi) - \widehat{v}(m; \xi_N)) \varphi_{\xi}^m \right\|_{L^2(\Omega_0^R)}^2 &= \sum_{m \in \mathbb{Z}_N} |\widehat{v}(m; \xi) - \widehat{v}(m; \xi_N)|^2 = \sum_{m \in \mathbb{Z}_N} \left| \int_{\Omega_0^R} (1 - e^{i\bar{x} \cdot (\xi - \xi_N)}) v(x) \varphi_{\xi}^m dx \right|^2 \\ &\leq \left\| (1 - e^{i\bar{x} \cdot (\xi - \xi_N)}) v \right\|_{L^2(\Omega_0^R)}^2 \leq \left\| 1 - e^{i\bar{x} \cdot (\xi - \xi_N)} \right\|_{L^\infty(\Omega_0^R)}^2 \|v\|_{L^2(\Omega_0^R)}^2 \end{aligned}$$

and

$$\begin{aligned} \left\| \sum_{m \in \mathbb{Z}_N} \widehat{v}(m; \xi_N) (\varphi_{\xi}^m - \varphi_{\xi_N}^m) \right\|_{L^2(\Omega_0^R)}^2 &= \left\| (e^{i\bar{x} \cdot (\xi - \xi_N)} - 1) \sum_{m \in \mathbb{Z}_N} \widehat{v}(m; \xi_N) \varphi_{\xi_N}^m \right\|_{L^2(\Omega_0^R)}^2 \\ &\leq \left\| e^{i\bar{x} \cdot (\xi - \xi_N)} - 1 \right\|_{L^\infty(\Omega_0^R)}^2 \|P_{\xi_N}^N v\|_{L^2(\Omega_0^R)}^2 \leq \left\| 1 - e^{i\bar{x} \cdot (\xi - \xi_N)} \right\|_{L^\infty(\Omega_0^R)}^2 \|v\|_{L^2(\Omega_0^R)}^2. \end{aligned}$$

We also see that

$$\left\| 1 - e^{i\bar{x} \cdot (\xi - \xi_N)} \right\|_{L^\infty(\Omega_0^R)}^2 = \left\| 2 \sin \frac{\bar{x} \cdot (\xi_N - \xi)}{2} \right\|_{L^\infty(\Omega_0^R)}^2 \xrightarrow{N \rightarrow \infty} 0,$$

which proves the lemma.  $\square$

## 4 Discretization of the locally perturbed periodic problem and convergence analysis

We now address the discretization of the original problem (9). Let us set again  $f_{\xi} = \mathcal{F}f(\cdot, \xi)$ . The idea is to perform first a discretization with respect to the Floquet-Bloch variable, then perform the discretization in space as done in the previous section. We recall that  $\tilde{u} \in (L^2(\llbracket -\frac{\pi}{L}, \frac{\pi}{L} \rrbracket], H_{\xi}^2(\Omega_0))$  and

$$\begin{cases} \Delta \tilde{u}(\cdot; \xi) + k^2 n_p \tilde{u}(\cdot; \xi) + k^2 (n - n_p) \mathcal{M}(\tilde{u}) = f_{\xi} & \text{in } \Omega_0, \quad \xi \in \llbracket -\frac{\pi}{L}, \frac{\pi}{L} \rrbracket \\ \mathcal{M}(\tilde{u}) := \frac{\llbracket L \rrbracket}{(2\pi)^{d-1}} \int_{\llbracket -\frac{\pi}{L}, \frac{\pi}{L} \rrbracket} \tilde{u}(\cdot; \xi) d\xi & \text{in } \Omega_0, \end{cases} \quad (56)$$

and that  $u$  can be reconstructed from  $\tilde{u}$  using (10).

### 4.1 Discretization and convergence in the Floquet-Bloch variable

We discretize the integral  $\mathcal{M}(\tilde{u})$  using the trapezoidal rule. Consider a uniform partition of  $\llbracket -\frac{\pi}{L}, \frac{\pi}{L} \rrbracket$  into  $\prod_{\ell=1}^{d-1} M_{\ell}$  sub-domains of size  $\Delta \xi_1 \times \dots \times \Delta \xi_{d-1}$  with  $\Delta \xi_{\ell} := 2\pi / (M_{\ell} L_{\ell})$ . Set  $M := (M_1, \dots, M_{d-1}) \in \mathbb{Z}_+^{d-1}$  and define  $\mathbb{Z}_M^{d-1} := \{j = (j_1, \dots, j_{d-1}), [-\frac{M_{\ell}}{2}] + 1 \leq j_{\ell} \leq [\frac{M_{\ell}}{2}], \ell = 1, \dots, d-1\}$ . Here we use the notation

$[\cdot]$  to denote the floor function and  $[a] = ([a_1], \dots, [a_d])$  if  $a = (a_1, \dots, a_d) \in \mathbb{R}^d$ . The discretization points are

$$\xi_j = (j_1 \Delta \xi_1, \dots, j_{d-1} \Delta \xi_{d-1}); \quad j = (j_1, \dots, j_{d-1}) \in \mathbb{Z}_M^{d-1}.$$

Denote by  $\tilde{u}_M(\cdot; \xi_j)$  an approximation of  $\tilde{u}(\cdot, \xi_j)$ . Then, using the fact that  $\tilde{u}(\cdot, -\pi/L) = \tilde{u}(\cdot, \pi/L)$ , we set as discretized equations associated with (56) the following system:  $\tilde{u}_M(\cdot; \xi_j) \in H_{\xi_j}^2(\Omega_0)$

$$\begin{cases} \Delta \tilde{u}_M(\cdot; \xi_j) + k^2 n_p \tilde{u}_M(\cdot; \xi_j) + k^2(n - n_p)u_M = f_{\xi_j} & \text{in } \Omega_0, \quad j \in \mathbb{Z}_M^{d-1} \\ u_M := \frac{1}{\llbracket M \rrbracket} \sum_{j \in \mathbb{Z}_M^{d-1}} \tilde{u}_M(\cdot; \xi_j), & \text{in } \Omega_0. \end{cases} \quad (57)$$

Since  $\text{supp}((n - n_p)u_M) \subset \Omega_0^h$ , one deduces (using similar arguments as in establishing (17)) that system (57) is equivalent to

$$\begin{cases} \tilde{u}_M(\cdot; \xi_j) - k^2 V_{\xi_j}((n_p - 1)\tilde{u}_M(\cdot; \xi_j)) - k^2 V_{\xi_j}((n - n_p)u_M) = -V_{\xi_j}(f_{\xi_j}) & \text{in } L^2(\Omega_0^h), \\ u_M := \frac{1}{\llbracket M \rrbracket} \sum_{j \in \mathbb{Z}_M^{d-1}} \tilde{u}_M(\cdot; \xi_j) & \text{in } \Omega_0. \end{cases} \quad (58)$$

We will prove later that the function  $u_M$  constitutes in fact an approximation of  $u$  in  $\Omega_0^h$ . We recall that  $\tilde{u}_M(\cdot; \xi_j)$  can be extended to  $\Omega_0$  using the first equation in (58) and then extended to all  $\mathbb{R}^d$  using the quasi-periodicity property as

$$\tilde{u}_M((\bar{x} + mL, x_d); \xi_j) := e^{i(mL) \cdot \xi_j} \tilde{u}_M((\bar{x}, x_d); \xi_j), \quad x = (\bar{x}, x_d) \in \Omega_0, \quad m \in \mathbb{Z}^{d-1}.$$

Then we extend  $u_M$  to  $\mathbb{R}^d$  using the same formula as in the second equation of (58), namely

$$u_M := \frac{1}{\llbracket M \rrbracket} \sum_{j \in \mathbb{Z}_M^{d-1}} \tilde{u}_M(\cdot; \xi_j) \quad \text{in } \mathbb{R}^d.$$

We also define  $f_M$  as

$$f_M := \frac{1}{\llbracket M \rrbracket} \sum_{j \in \mathbb{Z}_M^{d-1}} f_{\xi_j} \quad \text{in } \mathbb{R}^d.$$

In particular,

$$u_M(\bar{x} + mL, x_d) = \frac{1}{\llbracket M \rrbracket} \sum_{j \in \mathbb{Z}_M^{d-1}} e^{i(mL) \cdot \xi_j} \tilde{u}_M((\bar{x}, x_d); \xi_j), \quad x = (\bar{x}, x_d) \in \Omega_0, \quad m \in \mathbb{Z}^{d-1}. \quad (59)$$

The following notation will be used

$$\Omega_M^h := \cup_{m \in \mathbb{Z}_M^{d-1}} \Omega_m^h,$$

$$\Omega_M := \cup_{m \in \mathbb{Z}_M^{d-1}} \Omega_m = \llbracket (-\frac{M}{2} + \frac{1}{2})L, (\frac{M}{2} + \frac{1}{2})L \rrbracket \times \mathbb{R},$$

$$\Gamma_{\ell, M}^- := \{x \in \Omega_M, x_\ell = (-\frac{M_\ell}{2} + \frac{1}{2})L_\ell\}, \quad \Gamma_{\ell, M}^+ := \{x \in \Omega_M, x_\ell = (\frac{M_\ell}{2} + \frac{1}{2})L_\ell\}, \quad \forall \ell = 1, \dots, d-1$$

and the term  $ML$ -periodic will refer to periodic functions with respect to the first  $d-1$  variables with period  $ML$ . To distinguish the notation for the space of  $ML$ -periodic we shall set, for  $m \in N$ ,

$$H_{\#, M}^m(\mathbb{R}^d) := H_{0, ML}^m(\mathbb{R}^d) \quad \text{and} \quad H_{\#, M}^m(\Omega_M) := \{u|_{\Omega_M}; u \in H_{\#, M}^m(\mathbb{R}^d)\}.$$

We shall make extensive use of the following identity (that somehow occurs naturally from the conjugation of the  $\xi$ -quasi-periodicity with the use of the trapezoidal rule). For  $\ell, \ell' \in \mathbb{Z}_M^{d-1}$

$$\sum_{m \in \mathbb{Z}_M^{d-1}} e^{-i(mL) \cdot (\xi_\ell - \xi_{\ell'})} = \sum_{m \in \mathbb{Z}_M^{d-1}} e^{-i(m \frac{2\pi}{M}) \cdot (\ell - \ell')} = \begin{cases} 0 & \text{if } \ell \neq \ell' \\ \llbracket M \rrbracket & \text{if } \ell = \ell' \end{cases}. \quad (60)$$

**Lemma 4.1.** *Let us define  $G_{\#,M} := \frac{1}{\llbracket M \rrbracket} \sum_{j \in \mathbb{Z}_M^{d-1}} G_{\xi_j}$ . Then  $G_{\#,M} \in L^2_{\#,M}(\mathbb{R}^d)$  and satisfies:*

$$\Delta G_{\#,M} + k^2 G_{\#,M} = -\delta_{x_0} \quad \text{in } \Omega_M. \quad (61)$$

*Proof.* This lemma can be proved using the Fourier representation of the  $\xi$ -quasi periodic Green function. Recall that the  $ML$ -periodic Green's function, denoted by  $G_{ML}$ , has the series representation (apply (15) with  $\xi = 0$  and  $L = ML$ )

$$G_{ML}(x) = \frac{i}{2\llbracket ML \rrbracket} \sum_{j \in \mathbb{Z}^{d-1}} \frac{1}{\beta_{\#}(j)} \exp(i\alpha_{\#}(j) \cdot \bar{x} + i\beta_{\#}(j)|x_d|), \quad x = (\bar{x}, x_d) \in \mathbb{R}^d \quad (62)$$

where  $\alpha_{\#}(j) := \frac{2\pi}{ML}j$ ,  $\beta_{\#}(j) := \sqrt{k^2 - |\alpha_{\#}(j)|^2}$ . Also recall that the  $\xi$ -quasi periodic Green's function  $G_{\xi}$  has the series representation (15). Then, forming the expression of  $G_{\#,M}$  we deduce

$$\begin{aligned} G_{\#,M} &= \frac{1}{\llbracket M \rrbracket} \frac{i}{2\llbracket L \rrbracket} \sum_{\ell \in \mathbb{Z}_M^{d-1}} \sum_{j \in \mathbb{Z}^{d-1}} \frac{1}{\sqrt{k^2 - (\xi_{\ell} + \frac{2\pi}{L}j)^2}} \exp(i(\xi_{\ell} + \frac{2\pi}{L}j) \cdot \bar{x} + i\sqrt{k^2 - (\xi_{\ell} + \frac{2\pi}{L}j)^2}|x_d|) \\ &= \frac{i}{2\llbracket ML \rrbracket} \sum_{\ell \in \mathbb{Z}_M^{d-1}} \sum_{j \in \mathbb{Z}^{d-1}} \frac{1}{\sqrt{k^2 - (\frac{2\pi}{L}J(\ell, j))^2}} \exp(i\frac{2\pi}{L}J(\ell, j) \cdot \bar{x} + i\sqrt{k^2 - (\frac{2\pi}{L}J(\ell, j))^2}|x_d|), \end{aligned}$$

where  $J(\ell, j) := \ell + Mj$ . We see that,  $\text{span}\{J(\ell, j), \ell \in \mathbb{Z}_M^{d-1}, j \in \mathbb{Z}^{d-1}\} = \mathbb{Z}^{d-1}$  and  $J(\ell_1, j_1) = J(\ell_2, j_2)$  if and only if  $j_1 = j_2$  and  $\ell_1 = \ell_2$ . Therefore, reordering the summation, we get

$$G_{\#,M} = \frac{i}{2\llbracket ML \rrbracket} \sum_{J \in \mathbb{Z}^{d-1}} \frac{1}{\beta_{\#}(J)} \exp(i\alpha_{\#}(J) \cdot \bar{x} + i\beta_{\#}(J)|x_d|) \quad (63)$$

which clearly proves that  $G_{\#,M} = G_{ML}$ . □

We now introduce the potential  $V_{\#,M}$  with kernel  $G_{\#,M}$ :

$$V_{\#,M}f(x) := \int_{\Omega_M^h} G_{\#,M}(x-y)f(y) \, dy. \quad (64)$$

We have the following properties for the operator  $V_{\#,M}$ .

**Theorem 4.2.** *1. The volume potential  $V_{\#,M}$  is a linear bounded operator from  $L^2(\Omega_M)$  into  $H^2_{\#,M}(\mathbb{R}^2)$  and for all  $g \in L^2(\Omega_M)$ ,  $u := V_{\#,M}g \in H^2_{\#}(\Omega_M)$  is the unique variational solution to:  $\Delta u + k^2u = -g$  in  $\Omega_M$ .*

*2. Let  $\tilde{u}_M(\cdot; \xi_j)$  be the solution to system (58). Then the function  $u_M$  defined in (59) belongs to  $H^1_{\#}(\Omega_M)$  and verifies the following volume integral equation*

$$u_M - k^2V_{\#,M}((n-1)u_M) = -V_{\#,M}(f_M) \quad \text{in } L^2(\Omega_M). \quad (65)$$

*Proof.* Proving the first claim can be done similarly to Lemma (3.1). We now prove the second claim. We first introduce the operator  $\tilde{V}_{\xi} : L^2(\Omega_M^h) \rightarrow L^2(\Omega_M^h)$ :

$$\tilde{V}_{\xi}g(x) := \int_{\Omega_M^h} G_{\xi}(x-y)g(y) \, dy. \quad (66)$$

Then,  $V_{\#,M} = \frac{1}{\llbracket M \rrbracket} \sum_{j \in \mathbb{Z}_M^{d-1}} \tilde{V}_{\xi_j}$  and if  $g \in L^2(\mathbb{R}^2)$  with compact support in  $\Omega_0^h$ ,

$$\tilde{V}_{\xi_{\ell}}g(x) = \int_{\Omega_M^h} G_{\xi_{\ell}}(x-y)g(y) \, dy = \int_{\Omega_0^h} G_{\xi_{\ell}}(x-y)g(y) \, dy = V_{\xi_{\ell}}g(x), \quad \forall x \in \mathbb{R}^2. \quad (67)$$

We now observe that for  $\tilde{v}(\cdot; \xi_{\ell'})$  in  $L^2(\Omega_M)$  a  $\xi_{\ell'}$ -quasi-periodic function with period  $L$ , we have

$$\tilde{V}_{\xi_{\ell}}(\tilde{v}(\cdot; \xi_{\ell'})) = \begin{cases} \llbracket M \rrbracket V_{\xi_{\ell}}(\tilde{v}(\cdot; \xi_{\ell})) & \text{if } \ell' = \ell, \\ 0 & \text{if } \ell' \neq \ell. \end{cases} \quad (68)$$

To prove this claim we write,

$$\begin{aligned} \tilde{V}_{\xi_{\ell}}(\tilde{v}(\cdot; \xi_{\ell'}))(x) &= \int_{\Omega_M^h} G_{\xi_{\ell}}(x-y) \tilde{v}(y; \xi_{\ell'}) \, dy = \sum_{m \in \mathbb{Z}_M^{d-1}} \int_{\Omega_M^h} G_{\xi_{\ell}}(x-y) \tilde{v}(y; \xi_{\ell'}) \, dy \\ &= \sum_{m \in \mathbb{Z}_M^{d-1}} \int_{\Omega_0^h} G_{\xi_{\ell}}(x-y) \tilde{v}(y; \xi_{\ell'}) e^{-i(mL) \cdot (\xi_{\ell} - \xi_{\ell'})} \, dy \\ &= \int_{\Omega_0^h} G_{\xi_{\ell}}(x-y) \tilde{v}(y; \xi_{\ell'}) \, dy \left( \sum_{m \in \mathbb{Z}_M^{d-1}} e^{-i(mL) \cdot (\xi_{\ell} - \xi_{\ell'})} \right). \end{aligned}$$

Using property (60) we then obtain (68). Since  $n - n_p$  has compact support in  $\Omega_0^h$ , then from (67) we have:

$$V_{\xi_{\ell}}((n - n_p)u_M) = \tilde{V}_{\xi_{\ell}}((n - n_p)u_M). \quad (69)$$

Since  $(n_p - 1)\tilde{u}_M(\cdot; \xi_{\ell})$  is  $\xi_{\ell}$ -quasi periodic with period  $L$ , then from (68) we have:

$$V_{\xi_{\ell}}((n_p - 1)\tilde{u}_M(\cdot; \xi_{\ell})) = \frac{1}{\llbracket M \rrbracket} \tilde{V}_{\xi_{\ell}}((n_p - 1)\tilde{u}_M(\cdot; \xi_{\ell})) \quad \text{and} \quad \tilde{V}_{\xi_{\ell}}((n_p - 1)\tilde{u}_M(\cdot; \xi_{\ell'})) = 0, \quad \ell \neq \ell'.$$

Therefore,

$$V_{\xi_{\ell}}((n_p - 1)\tilde{u}_M(\cdot; \xi_{\ell})) = \tilde{V}_{\xi_{\ell}}((n_p - 1)u_M). \quad (70)$$

Similarly we also get

$$V_{\xi_{\ell}} f_{\xi_{\ell}} = \tilde{V}_{\xi_{\ell}} f_M.$$

Therefore

$$\tilde{u}_M(\cdot; \xi_{\ell}) - k^2 \tilde{V}_{\xi_{\ell}}((n_p - 1)u_M) - k^2 \tilde{V}_{\xi_{\ell}}((n - n_p)u_M) = -\tilde{V}_{\xi_{\ell}} f_M \quad \text{in } \Omega_M,$$

in other words,

$$\tilde{u}_M(\cdot; \xi_{\ell}) - k^2 \tilde{V}_{\xi_{\ell}}((n - 1)u_M) = -\tilde{V}_{\xi_{\ell}} f_M, \quad \text{in } \Omega_M. \quad (71)$$

Taking the sum with respect to  $\ell \in \mathbb{Z}_M^{d-1}$  we finally obtain

$$u_M - k^2 V_{\#,M}((n - 1)u_M) = -V_{\#,M} f_M, \quad (72)$$

which proves the theorem.  $\square$

The following Corollary can be immediately deduced from Theorem 4.2, Lemma 4.3 and estimate (2).

**Corollary 4.3.** *The function  $u_M \in H_{\#}^2(\Omega_M)$  and is solution to*

$$\Delta u_M + k^2 n u_M = f_M \quad \text{in } \Omega_M. \quad (73)$$

Furthermore, there exists a positive constant  $C$  independent from  $M$  such that

$$\|u_M\|_{H_{\#}^2(\Omega_M)} \leq C \|f_M\|_{L^2(\Omega_M^h)}. \quad (74)$$

Due to the non vanishing imaginary part of  $k^2$ , we now can formulate the following exponential convergence result with respect to  $M$  that is a consequence of the exponential decay of the solution  $u$ . A proof of this theorem can also be found in [4]. For the reader's convenience, we here give a sketch of the proof.



**Theorem 4.4.** *Let  $f \in L^2(\mathbb{R}^d)$  such that  $e^{\tau^* \|x\|} f \in L^2(\mathbb{R}^d)$  for some  $\tau^* > 0$ . Let  $u \in H^1(\mathbb{R}^d)$  be the solution of (1) and  $u_M \in H_{\#}^1(\Omega_M)$  be the solution of (73). Then there exists  $\tau_0 > 0$  and a constant  $C$  independent of  $M$  and  $f$  and  $\tau$  such that*

$$\|u - u_M\|_{H^1(\Omega_M)} \leq C \left( e^{-\tau \min(ML)/2} \|e^{\tau \|x\|} f\|_{L^2(\mathbb{R}^d)} + \|f - f_M\|_{L^2(\mathbb{R}^d)} \right) \quad \forall 0 \leq \tau \leq \tau_0. \quad (75)$$

*Proof.* The first observation is that, if we set  $v = e^{\tau \|x\|} u$ , then  $v$  satisfies

$$\Delta v - 2\tau \frac{x}{\|x\|} \cdot \nabla v + (k^2 n + \tau^2) v = e^{\tau \|x\|} f \text{ in } \mathbb{R}^d.$$

It is easy to see (similarly to (2)) that

$$\|v\|_{L^2(\mathbb{R}^d)} \leq \frac{1}{\sigma c_0} (\|e^{\tau \|x\|} f\|_{L^2(\mathbb{R}^d)} + 2\tau \|\nabla v\|_{L^2(\mathbb{R}^d)})$$

and

$$\|\nabla v\|_{L^2(\mathbb{R}^d)}^2 \leq (\|e^{\tau \|x\|} f\|_{L^2(\mathbb{R}^d)} \|v\|_{L^2(\mathbb{R}^d)} + 2\tau \|\nabla v\|_{L^2(\mathbb{R}^d)}) \|v\|_{L^2(\mathbb{R}^d)} + |k^2 n + \tau^2| \|v\|_{L^2(\mathbb{R}^d)}^2.$$

Choosing  $\tau_0 > 0$  such that  $1 - \frac{(2\tau_0)^2}{(\sigma c_0)^2} (\sigma c_0 + |k^2 n + \tau_0^2|^2 + 1) > 1/2$  and  $\tau_0 < \tau^*$  we get the existence of a constant  $C(\tau_0)$  independent from  $\tau$  such that

$$\|\Delta v\|_{L^2(\mathbb{R}^d)}^2 + \|\nabla v\|_{H^1(\mathbb{R}^d)}^2 \leq C \|e^{\tau \|x\|} f\|_{L^2(\mathbb{R}^d)}^2 \quad \forall 0 \leq \tau \leq \tau_0. \quad (76)$$

Let us denote by  $\phi_j^\pm := u|_{\Gamma_{j,M}^\pm}$  and  $\psi_j := \frac{\partial}{\partial x_j} u|_{\Gamma_{j,M}^\pm}$ . Then, from trace theorems and since  $\Gamma_{j,M}^\pm$  can be identified as a part of  $\mathbb{R}^{d-1}$ , there exists a constant  $C$  independent from  $j$  such that

$$e^{\tau(M_j \pm 1)L_j} (\|\phi_j^\pm\|_{H^{1/2}(\Gamma_{j,M}^\pm)}^2 + \|\psi_j^\pm\|_{H^{-1/2}(\Gamma_{j,M}^\pm)}^2) \leq C (\|\Delta v\|_{L^2(\mathbb{R}^d)}^2 + \|\nabla v\|_{H^1(\mathbb{R}^d)}^2). \quad (77)$$

Since  $w := u - u_M$  solves

$$\begin{cases} \Delta w + k^2 n w = f - f_M & \text{in } \Omega_M, \\ w|_{\Gamma_{j,M}^+} - w|_{\Gamma_{j,M}^-} = \phi_j^+ - \phi_j^-, \\ \frac{\partial}{\partial x_1} w|_{\Gamma_{j,M}^+} - \frac{\partial}{\partial x_1} w|_{\Gamma_{j,M}^-} = \psi_j^+ - \psi_j^-, \end{cases}$$

then, due to  $\sigma > 0$ , there exists constant  $C' > 0$  independent from  $M$  such that

$$\|w\|_{H^1(\Omega_M)}^2 \leq C' \left( \sum_j (\|\phi_j^\pm\|_{H^{1/2}(\Gamma_{j,M}^\pm)}^2 + \|\psi_j^\pm\|_{H^{-1/2}(\Gamma_{j,M}^\pm)}^2) + \|f - f_M\|_{L^2(\Omega_M)}^2 \right). \quad (78)$$

The result directly follows from combining (76), (77) and (78).  $\square$

## 4.2 Discretization in the spatial variable

The discretization procedure in the spatial variable is the same as the one discussed in Section 3.

### 4.2.1 Periodization of the coupled integral equations

To apply our method we first periodize the volume integral equations of system (58) in  $x_d$ -direction and get the following system.

$$\begin{cases} \tilde{u}_M^R(\cdot; \xi_j) - k^2 V_{\xi_j}^R ((n_p - 1) \tilde{u}_M^R(\cdot; \xi_j)) - k^2 V_{\xi_j}^R ((n - n_p) u_M^R) = -V_{\xi_j}^R f_{\xi_j} & \text{in } L^2(\Omega_0^R), \\ u_M^R = \frac{1}{\llbracket M \rrbracket} \sum_{j \in \mathbb{Z}_M^{d-1}} \tilde{u}_M^R(\cdot; \xi_j). \end{cases} \quad (79)$$

From Lemma 3.2, the solution to equation (58) and the solution to equation (79) coincides in  $\Omega_0^h$ . Therefore, the extension by quasi-periodicity of the solution to equation (58) coincide with the extension by quasi-periodicity of the solution to equation (79) in  $\Omega_M^h$ . Let

$$G_{\#,M}^R := \frac{1}{\llbracket M \rrbracket} \sum_{j \in \mathbb{Z}_M^{d-1}} G_{\xi_j}^R$$

be the periodization of  $G_{\#,M}$  in  $x_d$  direction. We introduce the volume potential  $V_{\#,M}^R$  with kernel  $G_{\#,M}^R$  as:

$$V_{\#,M}^R f(x) := \int_{\Omega_M^h} G_{\#,M}^R(x-y) f(y) dy. \quad (80)$$

Then, following the same arguments as in the proof of the second item in Theorem 4.2, we obtain that  $u_M^R \in L^2(\Omega_M^R)$  verifies:

$$u_M^R = k^2 V_{\#,M}^R((n-1)u_M^R) - V_{\#,M}^R f_M \quad \text{in } L^2(\Omega_M^R). \quad (81)$$

The link between the solution of the periodized equation (81) and the solution of (65) can be established in a similar way as for  $\xi$ -quasi-periodic problems (apply the procedure with  $\xi = 0$  and a period =  $ML$ ). We here regroup these results:

**Proposition 4.5.** • Equation (65) has a unique solution  $u_M \in L^2(\Omega_M^h)$  for all  $f_M \in L^2(\Omega_M^h)$  if and only if equation (81) has a unique solution  $u_M^R \in L^2(\Omega_M^R)$  for all  $f_M \in L^2(\Omega_M^h)$ .

- Let  $f_M \in L^2(\Omega_M^h)$ ,  $u_M^R \in L^2(\Omega_M^R)$  be the solution to (81) and  $u_M \in L^2(\Omega_M^h)$  be the solution to (65). Then

$$u_M^R = u_M \quad \text{in } \Omega_M^h.$$

Moreover  $u_M^R \in H_{\#}^2(\Omega_M^R)$  and there exists a positive constant  $C$  independent of  $M$  and  $f_M$  such that:

$$\|u_M^R\|_{H_{\#}^2(\Omega_M^R)} \leq C \|f_M\|_{L^2(\Omega_M^h)}. \quad (82)$$

The definition of  $H_{\#}^s(\Omega_M^R)$  is the same as  $H_{\xi}^s(\Omega_0^R)$  with  $\xi = 0$  and the period  $L$  replaced with  $ML$ . The uniform bound with respect to  $M$  is a consequence of the following lemma that can be proved in the same manner as in Lemma 3.3 using (26).

**Lemma 4.6.** Let  $s \in \mathbb{N}$ . The operator  $V_{\#,M}^R : H_{\#}^s(\Omega_M^R) \rightarrow H_{\#}^{s+2}(\Omega_M^R)$  is continuous. Moreover, there exists a constant  $C$  independent of  $M$  such that

$$\|V_{\#,M}^R g\|_{H_{\#}^{2+s}(\Omega_M^R)} \leq C \|g\|_{H_{\#}^s(\Omega_M^R)}, \quad \forall g \in H_{\#}^s(\Omega_M^R). \quad (83)$$

Let us introduce for later use in the convergence analysis  $T_{\#,M}^R : L^2(\Omega_M^R) \rightarrow L^2(\Omega_M^R)$  defined by

$$T_{\#,M}^R g := k^2 V_{\#,M}^R((n_p - 1)g) \quad (84)$$

and  $A_{\#,M}^R : L^2(\Omega_M^R) \rightarrow L^2(\Omega_M^R)$  defined by

$$A_{\#,M}^R g := g - T_{\#,M}^R g. \quad (85)$$

Then, as straightforward consequences of Lemma 4.6 and the first item in Proposition 4.5 we get the following.

**Lemma 4.7.** The operator  $T_{\#,M}^R : L^2(\Omega_M^R) \rightarrow L^2(\Omega_M^R)$  is compact. The operator  $A_{\#,M}^R : L^2(\Omega_M^R) \rightarrow L^2(\Omega_M^R)$  is bounded and injective. Moreover, there exists a positive constant  $C$  independent of  $M$  such that

$$\|A_{\#,M}^R g\|_{L^2(\Omega_M^R)} \leq C \|g\|_{L^2(\Omega_M^R)}.$$

### 4.2.2 Spectral approximation

Let  $\tilde{u}_M^R(\cdot; \xi_j) \in L^2(\Omega_0^R)$  be the solution of (79). We consider  $\tilde{u}_M^{R,N}(\cdot; \xi_j) \in \mathcal{T}_{\xi_j}^N$  an approximation of  $\tilde{u}_M^R(\cdot; \xi_j)$ , which is solution to following variational equations

$$\begin{aligned} & \left\langle \tilde{u}_M^{R,N}(\cdot; \xi_j), v_{\xi_j}^N \right\rangle_{L^2(\Omega_0^R)} - k^2 \left\langle V_{\xi_j}^R((n_p - 1)\tilde{u}_M^{R,N}(\cdot; \xi_j)), v_{\xi_j}^N \right\rangle_{L^2(\Omega_0^R)} \\ & - k^2 \left\langle V_{\xi_j}^R((n - n_p)(\Delta \xi \sum_{\ell} u_M^{R,N}(\cdot; \xi_{\ell}))), v_{\xi_j}^N \right\rangle_{L^2(\Omega_0^R)} = - \left\langle V_{\xi_j}^R(f_{\xi_j}), v_{\xi_j}^N \right\rangle_{L^2(\Omega_0^R)}, \quad \forall v_{\xi_j}^N \in \mathcal{T}_{\xi_j}^N(\Omega_0^R). \end{aligned} \quad (86)$$

$u_M^{R,N}$  is defined as

$$u_M^{R,N} := \frac{1}{\llbracket M \rrbracket} \sum_{j \in \mathbb{Z}_M^{d-1}} \tilde{u}_M^{R,N}(\cdot; \xi_j). \quad (87)$$

We now define  $\mathcal{T}_{\#, \xi_j}^N$  as the space of functions that are extensions by  $\xi_j$ -quasi-periodicity of functions in  $\mathcal{T}_{\xi_j}^N$  to the domain  $\Omega_M^R$ . We then set

$$\mathcal{T}_{\#, M}^N := \bigoplus_{j \in \mathbb{Z}_M^{d-1}} \mathcal{T}_{\#, \xi_j}^N.$$

We extend  $\tilde{u}_M^{R,N}(\cdot; \xi_j)$  by  $\xi_j$ -quasi-periodicity to  $\Omega_M^R$ , which then can be considered as element of  $\mathcal{T}_{\#, \xi_j}^N$ . The function  $u_M^{R,N}$  is then extended to  $\Omega_M^R$  using (87). We now prove that  $u_M^{R,N}$  is also an approximation of the solution  $u_M^R$  of equation (81) using the Fourier basis on  $L_{\#}^2(\Omega_M^R)$ .

We first observe that

$$\mathcal{T}_{\#, M}^N = \text{span}\{\varphi_{\#, M}^J, J \in \mathbb{Z}_{M, N}^2\} \quad (88)$$

where for  $J = (\bar{J}, J_d) \in \mathbb{Z}^{d-1} \times \mathbb{Z}$

$$\varphi_{\#, M}^J(\bar{x}, x_d) := \frac{1}{\sqrt{2R\llbracket ML \rrbracket}} \exp\left(i \left( \frac{2\pi}{ML} \bar{J} \right) \cdot \bar{x} + i \frac{\pi}{R} J_d x_d\right)$$

and where  $\mathbb{Z}_{M, N}^d \subset \mathbb{Z}^d$  is defined for  $N = (\bar{N}, N_d) \in \mathbb{Z}^{d-1} \times \mathbb{Z}$  and  $M \in \mathbb{Z}^{d-1}$  by

$$\mathbb{Z}_{M, N}^2 := \{J = (\bar{J}, J_d) \in \mathbb{Z}^{d-1} \times \mathbb{Z}; [-\frac{M}{2}] - \frac{MN}{2} + M + 1 \leq \bar{J} \leq [\frac{M}{2}] + \frac{MN}{2} \text{ and } -N_d/2 + 1 \leq J_d \leq N_d/2\}.$$

This a consequence of the fact that for any  $J = (\bar{J}, J_d) \in \mathbb{Z}^{d-1} \times \mathbb{Z}$  we can uniquely associate a couple  $(j, \ell) \in \mathbb{Z}_N^d \times \mathbb{Z}_N^{d-1}$  such that

$$\bar{J} = \bar{j} + \ell M \text{ and } j_d = J_d \quad (89)$$

and that if (89) holds then

$$\varphi_{\#, M}^J = \frac{1}{\sqrt{\llbracket M \rrbracket}} \varphi_{\xi_{\ell}}^j \text{ in } \Omega_M^R. \quad (90)$$

We now can prove the following.

**Theorem 4.8.** *The function  $u_M^{R,N}$  defined in (87) belongs to  $\mathcal{T}_{\#, M}^N$  and verifies*

$$\left\langle u_M^{R,N}, v_M^N \right\rangle_{L^2(\Omega_M^R)} - k^2 \left\langle V_{\#, M}^R((n-1)u_M^{R,N}), v_M^N \right\rangle_{L^2(\Omega_M^R)} = - \left\langle V_{\#, M}^R f_M, v_M^N \right\rangle_{L^2(\Omega_M^R)}, \quad \forall v_M^N \in \mathcal{T}_{\#, M}^N. \quad (91)$$

*Proof.* Similarly to the proof of Theorem 4.2 let us introduce the periodized potential  $\tilde{V}_{\xi_j}^R$  as

$$\tilde{V}_{\xi_j}^R g(x) := \int_{\Omega_M^R} G_{\xi_j}^R(x-y)g(y) \, dy. \quad (92)$$

Then, by definition,  $V_{\#,M}^R = \frac{1}{\llbracket M \rrbracket} \sum_{j \in \mathbb{Z}_M^{d-1}} \tilde{V}_{\xi_j}^R$ . Since  $(n - n_p)u_M^{R,N}$  has support in  $\Omega_0^R$ , then

$$V_{\xi_j}^R \left( (n - n_p)u_M^{R,N} \right) = \tilde{V}_{\xi_j}^R \left( (n - n_p)u_M^{R,N} \right). \quad (93)$$

As in the proof of Theorem 4.2, the main ingredient here is that for  $\tilde{v}(\cdot; \xi_{\ell'})$  in  $L^2(\Omega_M^R)$  a  $\xi_{\ell'}$ -quasi-periodic function with period  $L$ , we have

$$\tilde{V}_{\xi_{\ell'}}^R(\tilde{v}(\cdot; \xi_{\ell'})) = \begin{cases} \llbracket M \rrbracket V_{\xi_{\ell'}}^R(\tilde{v}(\cdot; \xi_{\ell})) & \text{if } \ell' = \ell, \\ 0 & \text{if } \ell' \neq \ell. \end{cases}$$

The proof of this property is the same as the proof of (68) and relies on identity (60). We then get

$$\tilde{V}_{\xi_j}^R \left( (n_p - 1)\tilde{u}_M^{R,N}(\cdot; \xi_j) \right) = \tilde{V}_{\xi_j}^R \left( (n_p - 1)u_M^{R,N} \right) \quad \text{and} \quad \tilde{V}_{\xi_j}^R(f_{\xi_j}) = \tilde{V}_{\xi_j}^R(f_M). \quad (94)$$

Combined with equation (86), we obtain that  $\tilde{u}_M^{R,N}(\cdot; \xi_j)$  verifies

$$\left\langle \tilde{u}_M^{R,N}(\cdot; \xi_j), v_{\xi_j}^N \right\rangle_{L^2(\Omega_0^R)} - k^2 \left\langle \tilde{V}_{\xi_j}^R \left( (n - 1)u_M^{R,N} \right), v_{\xi_j}^N \right\rangle_{L^2(\Omega_0^R)} = - \left\langle \tilde{V}_{\xi_j}^R f_M, v_{\xi_j}^N \right\rangle_{L^2(\Omega_0)}, \quad \forall v_{\xi_j}^N \in \mathcal{T}_{\xi_j}^N. \quad (95)$$

Thanks to the the quasi-periodicity property we deduce that

$$\begin{aligned} \left\langle \tilde{u}_M^{R,N}(\cdot; \xi_j), v_{\xi_j}^N \right\rangle_{L^2(\Omega_m^R)} - k^2 \left\langle \tilde{V}_{\xi_j}^R \left( (n - 1)u_M^{R,N} \right), v_{\xi_j}^N \right\rangle_{L^2(\Omega_m^R)} \\ = - \left\langle \tilde{V}_{\xi_j}^R f_M, v_{\xi_j, M}^N \right\rangle_{L^2(\Omega_m)}, \quad \forall v_{\xi_j, M}^N \in \mathcal{T}_{\xi_j, M}^N(\Omega_m), \quad \forall m \in \mathbb{Z}_M^{d-1}. \end{aligned} \quad (96)$$

Taking the sum with respect to  $m$  we then obtain

$$\left\langle \tilde{u}_{\#,M}^{R,N}(\cdot; \xi_j), v_{\xi_j}^N \right\rangle_{L^2(\Omega_M^R)} - k^2 \left\langle \tilde{V}_{\xi_j}^R \left( (n - 1)u_M^{R,N} \right), v_{\xi_j}^N \right\rangle_{L^2(\Omega_M^R)} = - \left\langle \tilde{V}_{\xi_j}^R f_M, v_{\xi_j}^N \right\rangle_{L^2(\Omega_M)}, \quad \forall v_{\xi_j}^N \in \mathcal{T}_{\#, \xi_j}^N. \quad (97)$$

Using identity (60) we also get that if  $\tilde{v}(\cdot; \xi_j)$  in  $L^2(\Omega_M^R)$  is a  $\xi_j$ -quasi-periodic function with period  $L$  then

$$\left\langle \tilde{v}(\cdot; \xi_j), v_{\xi_{\ell}}^N \right\rangle_{L^2(\Omega_M^R)} = 0 \quad \text{if } \ell \neq j.$$

Therefore we can replace  $v_{\xi_j}^N$  in (97) with any test function  $v_M^N \in \mathcal{T}_{\#, M}^N$ . That is,

$$\begin{aligned} \left\langle \tilde{u}_M^{R,N}(\cdot; \xi_j), v_M^N \right\rangle_{L^2(\Omega_M^R)} - k^2 \left\langle \tilde{V}_{\xi_j}^R \left( (n - 1)u_M^{R,N} \right), v_M^N \right\rangle_{L^2(\Omega_M^R)} \\ = - \left\langle \tilde{V}_{\xi_j}^R f_M, v_M^N \right\rangle_{L^2(\Omega_M)}, \quad \forall v_M^N \in \mathcal{T}_{\#, M}^N. \end{aligned} \quad (98)$$

Now, taking the average with respect to  $j \in \mathbb{Z}_M^{d-1}$  we obtain (91).  $\square$

### 4.2.3 Convergence analysis for spatial discretization

The convergence analysis is based on the interpretation of our discrete system (86) as the one obtained with the discretization of the periodic problem with period  $= ML$  (Theorem 4.8). Indeed the following convergence results for  $u_M^R$  implies uniform convergence results for each  $\xi_{\ell}$ -quasi-periodic component. The latter can be seen as a consequence of the results of Section 3.

**Theorem 4.9.** *Let  $u_M^R \in H^1(\Omega_M)$  and  $u_M^{R,N} \in \mathcal{T}_{\#, M}^N$  be solutions to (81) and (87) respectively. Then for all positive  $M \in \mathbb{N}^{d-1}$ , there exists  $N_0(M) \in \mathbb{N}^d$  such that for all  $N \in \mathbb{N}^d$ ,  $N > N_0$ ,  $u_M^{R,N}$  is uniquely defined and there exists a positive constant  $C$  independent from  $M$  and  $N$  such that*

$$\|u_M^R - u_M^{R,N}\|_{L^2(\Omega_M^R)} \leq C \inf_{v_M^N \in \mathcal{T}_M^N} \|u_M^R - v_M^N\|_{L^2(\Omega_M^R)}. \quad (99)$$

*Proof.* The proof of this Theorem can be seen as a consequence of Theorem 4.2.7 in [19]. The solution  $u_M^R$  of (81) satisfies

$$\langle A_{\#,M}^R u_M^R, v_M \rangle_{L^2(\Omega_M^R)} = \langle -V_{\#,M}^R f_M, v_M \rangle_{L^2(\Omega_M^R)}, \quad \forall v_M \in L^2(\Omega_M^R). \quad (100)$$

We then obtain from (91),

$$\langle A_{\#,M}^R (u_M^R - u_M^{R,N}), v_M^N \rangle_{L^2(\Omega_M^R)} = 0, \quad \forall v_M^N \in \mathcal{T}_{\#,M}^N. \quad (101)$$

Since  $u_M^{R,N} - v_M^N \in \mathcal{T}_{\#,M}^N$ , then  $\langle A_{\#,M}^R (u_M^R - u_M^{R,N}), u_M^{R,N} - v_M^N \rangle_{L^2(\Omega_M^R)} = 0$ .

$$\begin{aligned} \langle A_{\#,M}^R (u_M^R - u_M^{R,N}), u_M^R - u_M^{R,N} \rangle_{L^2(\Omega_M^R)} &= \langle A_{\#,M}^R (u_M^R - u_M^{R,N}), u_M^R - v_M^N \rangle_{L^2(\Omega_M^R)} \\ &\leq C \|u_M^R - u_M^{R,N}\|_{L^2(\Omega_M^R)} \|u_M^R - v_M^N\|_{L^2(\Omega_M^R)}. \end{aligned} \quad (102)$$

For a fixed  $M$ , let us define

$$w_M^{R,N} := \frac{u_M^R - u_M^{R,N}}{\|u_M^R - u_M^{R,N}\|_{L^2(\Omega_M^R)}}, \quad (103)$$

For a given  $\varepsilon > 0$ , there exists  $N_0(M) \in \mathbb{N}^d$  such that for all  $N > N_0$

$$|\langle T_{\#,M}^R w_M^{R,N}, w_M^{R,N} \rangle| < \varepsilon. \quad (104)$$

This claim is proved by contradiction. Suppose that there exists a sequence  $w_M^{R,N'}$  such that  $|\langle T_{\#,M}^R w_M^{R,N'}, w_M^{R,N'} \rangle| \geq \varepsilon$ . We see that  $(w_M^{R,N'})$  is bounded in  $L^2(\Omega_M^R)$  and therefore there exists a subsequence, which we denoted by  $(w_M^{R,N'})$ , that converges weakly to  $w \in L^2(\Omega_M^R)$ . Observe that (by use of (91))

$$\langle A_{\#,M}^R w_M^{R,N'}, v \rangle_{L^2(\Omega_M^R)} = \langle A_{\#,M}^R w_M^{R,N'}, v - P_{\#}^{N'} v \rangle_{L^2(\Omega_M^R)}$$

where  $P_{\#}^{N'}$  denotes the projection operator from  $L^2(\Omega_M^R)$  onto  $\mathcal{T}_{\#,M}^{N'}$ . Taking the limit as  $N' \rightarrow \infty$  implies that  $w = 0$  by the injectivity of  $A_{\#,M}^R$ . The compactness of  $T_{\#,M}^R$  then implies  $\langle T_{\#,M}^R w_M^{R,N'}, w_M^{R,N'} \rangle \rightarrow \langle T_{\#,M}^R w, w \rangle = 0$ . This is a contradiction with  $\varepsilon > 0$ . Now, we choose  $\varepsilon = \frac{1}{2}$  in (104) and use that  $A_{\#,M}^R = I - T_{\#,M}^R$  to deduce from (102) that:

$$\frac{1}{2} \|u_M^R - u_M^{R,N}\|_{L^2(\Omega_M^R)}^2 \leq C \|u_M^R - u_M^{R,N}\|_{L^2(\Omega_M^R)} \|u_M^R - v_M^N\|_{L^2(\Omega_M^R)}.$$

Thus,

$$\|u_M^R - u_M^{R,N}\|_{L^2(\Omega_M^R)} \leq C \inf_{v_M^N \in \mathcal{T}^N} \|u_M^R - v_M^N\|_{L^2(\Omega_M^R)}. \quad (105)$$

□

It is now possible to immediately deduce the convergence rate thanks to the regularity result of Proposition 3.4 and using the interpolation result of Lemma 3.8 with  $\xi = 0$ ,  $L$  replaced with  $ML$  and  $N_\ell$  replaced with  $M_\ell N_\ell$  ( $\ell = 1, d-1$ ).

**Theorem 4.10.** *Let  $u_M^R \in H^1(\Omega_M)$  and  $u_M^{R,N} \in \mathcal{T}_{\#,M}^N$  be verifying (81) and (87) respectively. Then for all positive  $M \in \mathbb{N}^{d-1}$ , there exists  $N_0(M) \in \mathbb{N}^d$  and a positive constant  $C$  independent from  $M$  such that for all  $N \in \mathbb{N}^d$ ,  $N > N_0$ ,*

$$\|u_M^R - u_M^{R,N}\|_{L^2(\Omega_M^R)} \leq C \left( \sum_{\ell=1}^d \frac{1}{(1 + N_\ell^2)^2} \right)^{1/2} \|f_M\|_{L^2(\Omega_M^R)}. \quad (106)$$

#### 4.2.4 A convergence result for a combined discretization

We are now in position to combine the previous convergence results and deduce a convergence result for the discretization of our original problem. Let  $f \in L^2(\mathbb{R}^d)$  with support in  $\Omega^h$  and let  $u \in H^1(\mathbb{R}^d)$  be the solution to (1) and  $u_M^{R,N} \in \mathcal{T}_{\#,M}^N$  be verifying (87). We shall assume that there exists  $\tau_* > 0$  such that

$$e^{\tau_* \|x\|} f \in L^2(\mathbb{R}^d). \quad (107)$$

To obtain the following convergence result, we first write the decomposition

$$u - u_M^{R,N} = (u - u_M) + (u_M - u_M^{R,N}).$$

We apply Theorem 4.4 to treat the term  $u - u_M$  and for the second term we first use Proposition 4.5 to see that

$$u_M - u_M^{R,N} = u_M^R - u_M^{R,N} \text{ in } \Omega_M^h,$$

then apply Theorem 4.10.

**Theorem 4.11.** *For all positive  $M \in \mathbb{N}^{d-1}$ , there exists  $N_0(M) \in \mathbb{N}^d$  and two positive constants  $C$  and  $\tau_0$  independent from  $M$  such that for all  $N \in \mathbb{N}^d$ ,  $N > N_0$ ,*

$$\|u - u_M^{R,N}\|_{L^2(\Omega_M^h)} \leq C \left( e^{-\tau_0 \min(ML)/2} \|e^{\tau_0 \|x\|} f\|_{L^2(\mathbb{R}^d)} + \|f - f_M\|_{L^2(\Omega_M)} + \left( \sum_{\ell=1}^d \frac{1}{(1 + N_\ell^2)^2} \right)^{1/2} \|f\|_{L^2(\mathbb{R}^d)} \right).$$

Indeed in this theorem we used that  $\|f_M\|_{L^2(\Omega_M)} \leq \|f\|_{L^2(\mathbb{R}^d)}$ . We also notice that  $\|f - f_M\|_{L^2(\Omega_M)}$  converges exponentially fast to 0 as  $M \rightarrow \infty$  thanks to assumption (107). In the cases where  $f$  has compact support, assumption (107) is automatically verified and  $f_M = f$  in  $\Omega_M$  as soon as the support of  $f$  is contained in  $\Omega_M$ . Let us also note that the rate of convergence with respect to  $N$  can be higher if additional smoothness on the coefficients and the source term is imposed.

**Theorem 4.12.** *Assume in addition that  $n \in W^{m,\infty}(\mathbb{R}^d)$  and  $f \in H_{loc}^m(\mathbb{R}^d)$  for some integer  $m > 0$ . Then, for all positive  $M \in \mathbb{N}^{d-1}$ , there exists  $N_0(M) \in \mathbb{N}^d$  and two positive constants  $C$  and  $\tau_0$  independent from  $M$  such that for all  $N \in \mathbb{N}^d$ ,  $N > N_0$ ,*

$$\|u - u_M^{R,N}\|_{L^2(\Omega_M^h)} \leq C \left( e^{-\tau_0 \min(ML)/2} \|e^{\tau_0 \|x\|} f\|_{L^2(\mathbb{R}^d)} + \|f - f_M\|_{L^2(\Omega_M)} + \left( \sum_{\ell=1}^d \frac{1}{(1 + N_\ell^2)^{m+2}} \right)^{1/2} \|f_M\|_{H_{\#}^m(\Omega_M)} \right).$$

*Proof.* The proof of this theorem follows the same lines as the proof of theorem 4.11 and the observation and that under the regularity assumption of the theorem, there exists a constant independent from  $M$  such that

$$\|u_M^R\|_{H^{2+m}(\Omega_M^R)} \leq C \|f_M\|_{H_{\#}^m(\Omega_M)}.$$

The latter is a direct consequence of (81) and Lemma 4.6.  $\square$

## 5 Numerical Algorithm and Experiments

We first notice that for any function in  $L^2(\Omega_0^R)$ ,

$$P_{\xi_j}^N V_{\xi_j}^R v = V_{\xi_j}^R P_{\xi_j}^N v.$$

Therefore, the discrete system (86) can be written as

$$\begin{cases} \tilde{u}_M^{R,N}(\cdot; \xi_j) - k^2 V_{\xi_j}^R (P_{\xi_j}^N ((n_p - 1) \tilde{u}_M^{R,N}(\cdot; \xi_j))) - k^2 V_{\xi_j}^R (P_{\xi_j}^N (n - n_p) u_M^{R,N}) = -V_{\xi_j}^R (P_{\xi_j}^N f) \\ u_M^{R,N} = \frac{1}{M} \sum_{\ell \in \mathbb{Z}_M^d} \tilde{u}_M^{R,N}(\cdot; \xi_\ell). \end{cases} \quad (108)$$

The Fourier coefficients of  $V_{\xi_j}^R P_{\xi_j}^N v$  can be expressed as

$$V_{\xi_j}^R P_{\xi_j}^N v(\ell; \xi_j) = \frac{1}{\gamma} \widehat{G_{\xi_j}}(\ell; \xi_j) \widehat{v}(\ell; \xi_j). \quad (109)$$

The coefficients  $\widehat{G_{\xi_j}}(\ell; \xi_j)$  have explicit expressions given in (25). Writing (108) in the Fourier domain then leads to simple algebraic equations for the  $\xi_j$ -Fourier coefficients (the term  $\xi_j$ -Fourier coefficients refers to the values  $\widehat{v}(\ell; \xi_j)$  of a function  $v$ ) of  $\tilde{u}_M^{R,N}(\cdot; \xi_j)$ , which constitute the unknowns of our problem. In these equations one needs in particular to evaluate the  $\xi_j$ -Fourier coefficients of

$$(n_p - 1) \tilde{u}_M^{R,N}(\cdot; \xi_j) \text{ and } (n - n_p) u_M^{R,N}.$$

This step is numerically done using (inverse) FFT to evaluate first the nodal points of  $\tilde{u}_M^{R,N}(\cdot; \xi_j)$  (from the  $\xi_j$ -Fourier coefficients) then FFT to compute the  $\xi_j$ -Fourier coefficients of  $(n_p - 1) \tilde{u}_M^{R,N}(\cdot; \xi_j)$  and  $(n - n_p) u_M^{R,N}$ .

As in [16], we use the Matlab version of GMRES algorithm to invert the linear system in the Fourier domain. In the numerical experiments below, we used a tolerance  $=10^{-9}$ . We also note that we did not use any preconditioning of our system since our objective here is to rather give some preliminary validating tests of the method in a two dimensional setting of the problem.

### Example 1: Case of $n_p = 1$

We here consider the scattering problem described in Remark 2.1 for a locally perturbed homogeneous domain, where the local perturbation is a disk of center  $0 = (0, 0)$  and radius  $r < \min\{\frac{L}{2}, x_d^0\}$ , which has refractive index  $n > 1$ . Thanks to separation of variables we can calculate an analytic expression for the solution as

$$\begin{cases} u = \sum_m \beta_m J_m(k\sqrt{n}|x|) e^{im\theta} & |x| < r, \\ u^s = \sum_m \alpha_m H_m^{(1)}(k|x|) e^{im\theta} & |x| > r, \end{cases} \quad (110)$$

where,

$$\begin{cases} \alpha_m = \frac{i}{4} H_m^{(1)}(k|x^0|) e^{-in\theta^0} \frac{\sqrt{n} J_m'(k\sqrt{n}|x|) J_m(k|x|) - J_m(k\sqrt{n}|x|) J_m'(k|x|)}{-\sqrt{n} J_m'(k\sqrt{n}|x|) H_m^{(1)}(k|x|) + J_m(k\sqrt{n}|x|) H_m^{(1)'}(k|x|)}, \\ \beta_m = \frac{i}{4} H_m^{(1)}(k|x^0|) e^{-in\theta^0} \frac{\sqrt{n} H_m^{(1)'}(k|x|) J_m(k|x|) - H_m^{(1)}(k|x|) J_m'(k|x|)}{H_m^{(1)'}(k|x|) J_m(k\sqrt{n}|x|) - \sqrt{n} J_m'(k\sqrt{n}|x|) H_m^{(1)}(k|x|)}. \end{cases} \quad (111)$$

Here  $\theta^0$  denotes the angle between  $x$  and  $x^0$  and  $J_m$  and  $H_m^{(1)}$  respectively denote the Bessel function and the Hankel function of the first kind of order  $m$ . Figure 1 reports the results obtained using the numerical approximation (right) and the one obtained using the analytic expression (left). These results correspond with a refractive index  $n = 3$ , a wave number  $k = \pi + 0.1i$ , a period length  $L = 2\pi$ ,  $x^0 = (0, 2.5\lambda)$  where  $\lambda := \frac{2\pi}{|k|}$  and  $r = 0.5\lambda$ . For our numerical solution, we used  $h = 1.5\lambda$  and  $R = \pi\lambda$ . We observe that we obtain qualitatively the same solution.

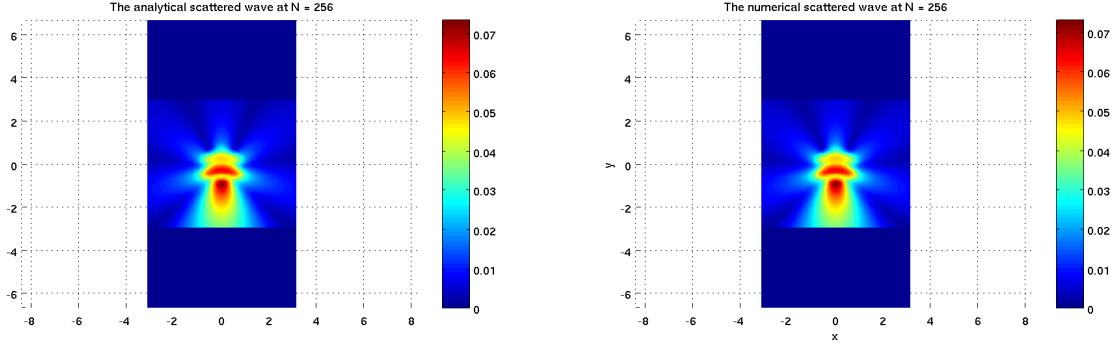


Figure 1: The analytic solution in  $\Omega_0^h$  (left) and the numerical solution in  $\Omega_0^h$  (right) for the first example (here  $h = 3$ ).

Figure 2-right reports the relative error for fixed  $N = 256$  and varying  $M$ . We observe a fast convergence rate with a rapid saturation of the error, meaning that the approximation error with respect to  $M$  becomes smaller than the one with respect to  $N$ .

Figure 2-left indicates the behavior of the  $L^2(\Omega_0^h)$  error with respect to  $N$  (we choose  $N_1 = N_2 = N$ ) for fixed  $M = 9$ . We observe that the slope of the log-log curve is  $-1/2$  less than the predicted one (that should be  $-2$ ). We think that this is due to the fact that we do not evaluate exactly the  $\xi_j$ -Fourier coefficients of  $(n - n_p)u_M^{R,N}$  or those of  $(n - 1)f$ . In our algorithm, we first project the function on the nodal points then evaluate numerically the FFT of the obtained discrete vector. Indeed in our convergence theorem we did not take into account this type of error.

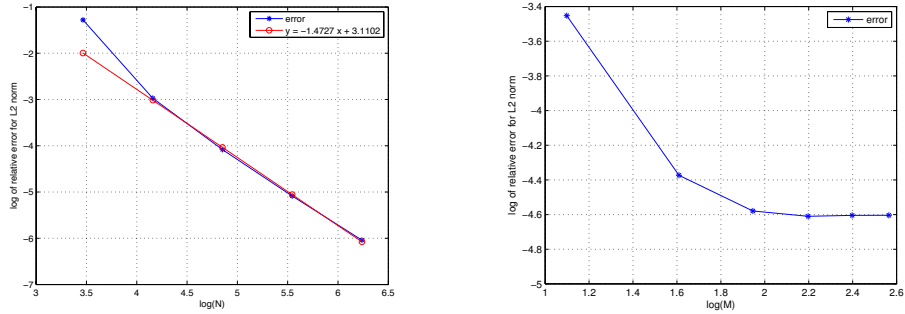


Figure 2: Left: The  $L^2$  relative error in  $\Omega_0^h$  with respect to  $N$  when  $M = 9$ . Right: The  $L^2$  relative error in  $\Omega_0^h$  with respect to  $M$  when  $N = 256$ .

To illustrate the effect of the numerical approximation of the Fourier transform of discontinuous functions on the convergence rate, we compare in Figure 3 the error in evaluating  $V_\xi^R f$  with different methods in computing the Fourier coefficients of  $f$ . Consider the piecewise constant  $f \in L^2(\Omega_0^R)$  where  $f = 2$  in  $\Omega_0^h$  and  $f = 0$  otherwise. We first compute the convolution operator  $F_N := P_\xi^N V_\xi^R f$  using the exact Fourier coefficients of  $f$ . We second compute  $\tilde{F}_N \approx P_\xi^N V_\xi^R f$  by evaluating the Fourier coefficients of  $f$  using FFT. The Fourier coefficients of  $V_\xi^R$  are exactly computed in both cases. In Figure 3 we report the  $L^2(\Omega_0^R)$  norm of  $E_N = F_N - F_{N_{\max}}$  and  $\tilde{E}_N = \tilde{F}_N - \tilde{F}_{N_{\max}}$  for  $N = 2^n, n = 5, \dots, 9$  and with  $N_{\max} = 2^{11}$ . While the rate of convergence is equal to the one predicted by the theory for  $E_N$  (Figure 3-left) we observe a non monotone rate of convergence for  $\tilde{E}_N$  (Figure 3-right).



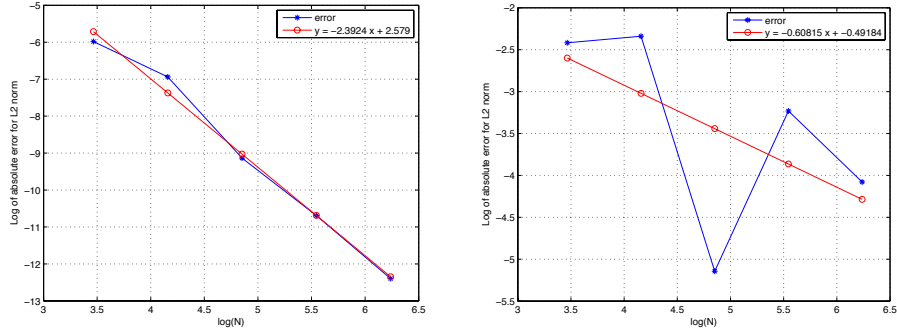


Figure 3: Left: The rate of convergence of  $E_N$  with respect to  $N$  when evaluating the convolution operator using exact Fourier coefficients. The rate of convergence of  $\tilde{E}_N$  with respect to  $N$  when evaluating the convolution operator using FFT.

Let us notice that the observation made for Figure 3 is less visible for regular functions  $f$  as attested by the following experiment.

In the experiment illustrated by Figure 4 we again consider the scattering problem with  $n_p = 1$  but here the local perturbation (the support of  $n - 1$ ) is a square  $\omega$  centered at the origin with an edge length  $r = 0.6$ . We give in Figure 4 the convergence rate with respect to  $N$  (for  $M = 9$ ) for three different choices of the refractive index  $n$  inside the domain  $\omega$ :  $n = n_1 := 3$ ,  $n = n_2 := 1 + \frac{1}{2} \cos(\frac{\pi x}{r}) \cos(\frac{\pi y}{r})$  and  $n = n_3 := 1 + 2 \cos^2(\frac{\pi x}{r}) \cos^2(\frac{\pi y}{r})$ . The error is computed as the  $L^2(\Omega_0^h)$  norm of  $u_M^{R,N} - u_M^{R,N_{\max}}$  where  $N_{\max} = 2^{10}$  and  $N = 2^n$ ,  $n = 5, \dots, 9$ . As we can observe, the correct rate of convergence is only observed for the third choice of  $n$ . For the second choice, the rate is closer to the correct one (that should be 3) for the first values of  $N$ .

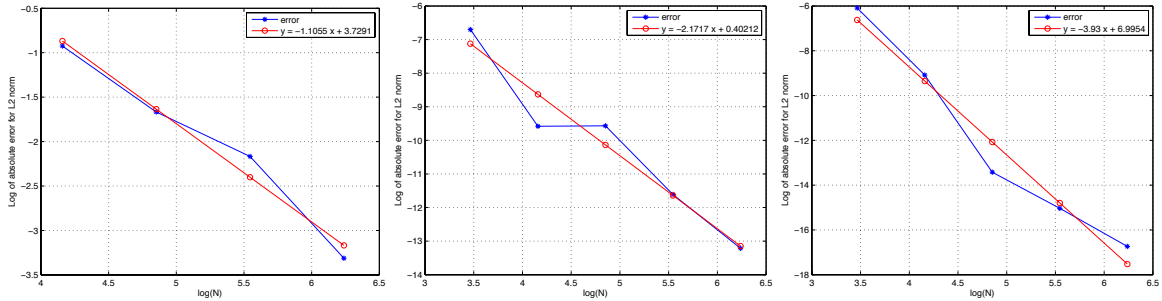


Figure 4:  $L^2(\Omega_0^h)$  norm of  $u_M^{R,N} - u_M^{R,N_{\max}}$  where  $N_{\max} = 2^{10}$  and  $N = 2^n$ ,  $n = 5, \dots, 9$  in the case of  $n = n_1$  (left),  $n = n_2$  (middle) and  $n = n_3$  (right).

### Example 2: The case of local perturbation

We end our numerical examples with the case of locally perturbed periodic media where  $n_p \neq n$ . The experiment corresponds with a scattering from a point source as in the first example. The parameters are the same except for the definition of  $n_p$  and  $n$ . The choice of the refractive index  $n_p$  is such that  $n_p = 3$  in  $B_0(r)$ ,  $n_p = 1$  in  $\Omega_0^h \setminus B_0(r)$  and  $r = 0.5 * \lambda$ . The local perturbation is such that  $n = 1$  in  $\Omega_0^h$ . We compare in Figure 5 the solution obtained by our numerical algorithm and the numerical solution computed on  $\Omega_M^R$  using the finite element code FreeFem++ with a periodicity conditions imposed of the boundary of  $\Omega_M^R$  for  $M = 9$ . We observe a slight difference between the two solutions that indeed comes from difference between the two numerical schemes but also from the difference between  $f_M$  and  $f$  in  $\Omega_M$  where  $f = (n - 1)u_i$ .

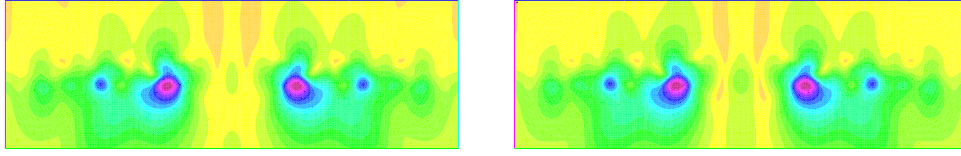


Figure 5: Left: Numerical solution obtained by finite element discretization of the periodic problem using Freefem++. Right: Numerical solution obtained by our algorithm.

We illustrate in Figure 6 the convergence rate with respect to  $N$  by computing the  $L^2(\Omega_0^h)$  norm of  $u_M^{R,N} - u_M^{R,N_{\max}}$  for  $M = 9$  with  $N_{\max} = 2^{11}$  and  $N = 2^n, n = 3, \dots, 7$ . We observe a rate of convergence compatible with the one observed for the first example.

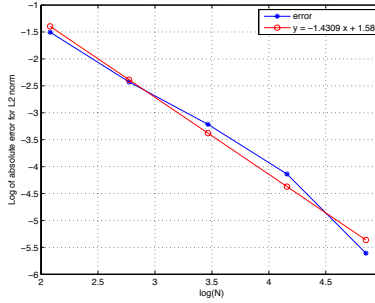


Figure 6:  $L^2(\Omega_0^h)$  norm of  $u_M^{R,N} - u_M^{R,N_{\max}}$  where  $N_{\max} = 2^{11}$  and  $N = 2^n, n = 3, \dots, 7$  in the case of locally perturbed periodic media and piecewise constant refractive index.

## References

- [1] T. ARENS AND T. HOHAGE, *On radiation conditions for rough surface scattering problems*, IMA J. Appl. Math., 70 (2005), pp. 839–847.
- [2] A.-S. BONNET-BENDHIA AND F. STARLING, *Guided waves by electromagnetic gratings and nonuniqueness examples for the diffraction problem*, Math. Methods Appl. Sci., 17 (1994), pp. 305–338.
- [3] S. N. CHANDLER-WILDE, P. MONK, AND M. THOMAS, *The mathematics of scattering by unbounded, rough, inhomogeneous layers*, J. Comput. Appl. Math., 204 (2007), pp. 549–559.
- [4] J. COATLÉVEN, *Analyse mathématique et numérique de quelques problèmes d’ondes en milieu périodique*, PhD thesis, 10 2011.
- [5] ———, *Helmholtz equation in periodic media with a line defect*, Journal of Computational Physics, 231 (2012).
- [6] D. COLTON AND R. KRESS, *Inverse acoustic and electromagnetic scattering theory*, vol. 93 of Applied Mathematical Sciences, Springer, New York, third ed., 2013.
- [7] S. FLISS, *Analyse mathématique et numérique de problèmes de propagation des ondes dans des milieux périodiques infinis localement perturbés*, PhD thesis, Ecole Polytechnique X, 2009.
- [8] H. HADDAR AND A. LECHLEITER, *Electromagnetic wave scattering from rough penetrable layers*, SIAM J. Math. Anal., 43 (2011), pp. 2418–2443.
- [9] T. HOHAGE, *On the numerical solution of a three-dimensional inverse medium scattering problem*, Inverse Problems, 17 (2001), pp. 1743–1763.

- [10] ———, *Fast numerical solution of the electromagnetic medium scattering problem and applications to the inverse problem*, J. Comp. Phys., 214 (2006), pp. 224–238.
- [11] P. JOLY, J.-R. LI, AND S. FLISS, *Exact boundary conditions for periodic waveguides containing a local perturbation*, Commun. Comput. Phys., 1 (2006), pp. 945–973.
- [12] A. KIRSCH, *Diffraction by periodic structures*, in Proc. Lapland Conf. on Inverse Problems, L. Pävarinta and E. Somersalo, eds., Springer, 1993, pp. 87–102.
- [13] P. KUCHMENT, *Floquet theory for partial differential equations*, vol. 60 of Operator Theory: Advances and Applications, Birkhäuser Verlag, Basel, 1993.
- [14] A. LECHLEITER AND D.-L. NGUYEN, *On uniqueness in electromagnetic scattering from biperiodic structures*, ESAIM Math. Model. Numer. Anal., 47 (2013), pp. 1167–1184.
- [15] A. LECHLEITER AND D.-L. NGUYEN, *Volume integral equations for scattering from anisotropic diffraction gratings*, Math. Methods Appl. Sci, 36 (2013), pp. 262–274.
- [16] A. LECHLEITER AND D.-L. NGUYEN, *A trigonometric Galerkin method for volume integral equations arising in TM grating scattering*, Adv. Comput. Math., 40 (2014), pp. 1–25.
- [17] A. LECHLEITER AND S. RITTERBUSCH, *A variational method for wave scattering from penetrable rough layers*, IMA J. Appl. Math., 75 (2010), pp. 366–391.
- [18] J. SARANEN AND G. VAINIKKO, *Periodic integral and pseudodifferential equations with numerical approximation*, Springer, 2002.
- [19] S. SAUTER AND C. SCHWAB, *Boundary Element Methods*, Springer, 1. ed., 2007.
- [20] G. VAINIKKO, *Fast solvers of the Lippmann-Schwinger equation*, in Direct and Inverse Problems of Mathematical Physics, D. Newark, ed., Int. Soc. Anal. Appl. Comput. 5, Dordrecht, 2000, Kluwer, p. 423.