

# ON THE CONSISTENCY OF COMPRESSED MODES FOR VARIATIONAL PROBLEMS ASSOCIATED WITH THE SCHRÖDINGER OPERATOR\*

FARZIN BAREKAT<sup>†</sup>

**Abstract.** This paper provides theoretical consistency results for compressed modes. We prove that as a  $L^1$  regularization term in certain non-convex variational optimization problems vanishes, the solutions of the optimization problem and the corresponding eigenvalues converge to a sequence of Wannier-like functions and the eigenvalues of the Hamiltonian, respectively.

**Key words.** compressed modes,  $L^1$ -regularization, Wannier-like functions

**AMS subject classifications.** 49Rxx, 47A75

**1. Introduction.** In [6, 7] the authors pioneered a new methodology of using sparsity techniques to obtain localized solutions to a class of problems in mathematical physics that can be recast as variational optimization problems. The typical method used for finding orthogonal functions that span an eigenspace of a Hamiltonian and are spatially localized is to choose a particular unitary transformation of the eigenfunctions of the Hamiltonian. In solid state physics and quantum mechanics these functions are known as Wannier functions [10]. There are many approaches to finding the “right” unitary transformation. The most widely used approach for calculating maximally localized Wannier functions (MLWFs) was introduced in [4]. There are two difficulties associated with this approach. First, the eigenfunctions of the Hamiltonian must be calculated. Second, one needs to determine a distance to manually cut off the resulting MLWFs.

In [6], it was demonstrated that introducing  $L^1$  regularization in the variational formulation of the Schrödinger equation of quantum mechanics and solving the new non-convex optimization problem, results in a set of localized functions called compressed modes (CMs). It was shown numerically that CMs have many desirable features, for example, the energy calculated using CMs approximates the ground state energy of the system. Moreover, there is no requirement to cut off the resulting CMs “by hand”. In a more recent development, the ideas of [6] were used in [7] to generate a new set of spatially localized orthonormal functions, called compressed plane waves (CPWs), with multi resolution capabilities adapted for the Laplace operator.

The idea of using the  $L^1$  norm as a constraint or penalty term to achieve sparsity has attracted a lot of attention in a variety of fields including compressed sensing [3, 2], matrix completion [8], phase retrieval [1], etc. Recently, sparsity techniques began being used in physical science (see for example [5]) and partial differential equations (see for example [9]). In all these examples sparsity means that in the representation of a corresponding vector or function in terms of a well-chosen set of modes (ie. a basis or dictionary), most coefficients are zero. However, [6, 7] for the first time advocates the use of  $L^1$  norm regularization to achieve spatial sparsity (i.e. functions that are spatially localized).

---

\*This work was partially supported by DOE grant number DE-SC0010613, and NSERC PGS-D award.

<sup>†</sup>Mathematics Department, University of California at Los Angeles, Los Angeles, CA 90095-1555 USA. (fbarekat@math.ucla.edu).

In this paper we prove consistency results for compressed modes (CMs) introduced in [6]. In particular, we show that as the  $L^1$  regularization term vanishes the approximate energy calculated using CMs converges from above to the actual energy of the system. This is done in section §2. More importantly, in section §3, we show that under some necessary assumptions on the spectrum of the Hamiltonian, as the  $L^1$  regularization term vanishes, CMs converge to a sequence of unitary transformations of the eigenfunctions (i.e. Wannier-like functions) in  $H^1$  norm. Moreover, we verify a conjecture stated in [6].

Throughout this paper for the ease of exposition we assume that the domain  $\Omega$  is a bounded rectangular subset of  $\mathbb{R}^d$  with periodic boundary conditions. Other boundary conditions can be handled in a similar way. Let  $\hat{H} = -\frac{1}{2}\Delta + V(\mathbf{x})$  denote the Hamiltonian of a system with normalized eigenfunctions  $\phi_1, \phi_2, \dots$ , and corresponding eigenvalues  $\lambda_1 \leq \lambda_2 \leq \dots$ . We assume that the potential  $V(\mathbf{x})$  is bounded; that is,

$$\|V(\mathbf{x})\|_\infty := \sup_{\Omega} |V(\mathbf{x})| < \infty.$$

These assumptions, in particular, imply that the spectrum of the Hamiltonian  $\hat{H}$  is discrete and its eigenvalues grow to positive infinity. Observe that  $\phi_1, \dots, \phi_N$  are a solution to the optimization problem:

$$(1.1) \quad \{\phi_1, \dots, \phi_N\} = \operatorname{argmin}_{\tilde{\phi}_1, \dots, \tilde{\phi}_N} \sum_{i=1}^N \langle \tilde{\phi}_i, \hat{H} \tilde{\phi}_i \rangle \quad \text{s.t.} \quad \langle \tilde{\phi}_j, \tilde{\phi}_k \rangle = \delta_{jk},$$

where  $\langle \phi_j, \phi_k \rangle = \int_{\Omega} \phi_j^*(\mathbf{x}) \phi_k(\mathbf{x}) d\mathbf{x}$ .

In [6], the compressed modes  $\{\psi_i\}_{i=1}^N$ , corresponding to the number  $N$ , are defined as the solution to the  $L^1$  regularized optimization problem

$$(1.2) \quad \{\psi_1, \dots, \psi_N\} = \operatorname{argmin}_{\tilde{\psi}_1, \dots, \tilde{\psi}_N} \sum_{i=1}^N \frac{1}{\mu} \|\tilde{\psi}_i\|_1 + \langle \tilde{\psi}_i, \hat{H} \tilde{\psi}_i \rangle \quad \text{s.t.} \quad \langle \tilde{\psi}_j, \tilde{\psi}_k \rangle = \delta_{jk},$$

where  $\|\psi_i\|_1 = \int_{\Omega} |\psi_i| d\mathbf{x}$ . As shown in [6, 7], compressed modes have many desirable features. In particular, consider the  $N \times N$  matrix  $\langle \Psi_N^T, \hat{H} \Psi_N \rangle$  with the  $(j, k)$ -th entry defined by  $\langle \psi_j, \hat{H} \psi_k \rangle$  and let  $(\sigma_1, \dots, \sigma_N)$  denote its eigenvalues in non-decreasing order. In [6, 7], it was conjectured that as  $\mu \rightarrow \infty$ , the  $\sigma_i$ 's converge to the  $\lambda_i$ 's. In theorem 3.5 we verify this conjecture. We also show that under some necessary assumptions on the spectrum of the Hamiltonian, as  $\mu \rightarrow \infty$ , the CMs  $\psi_i$ 's converge to a sequence of unitary transformations of the eigenfunctions  $\phi_i$ 's in the  $H^1$  norm. Observe that the orthonormality constraints in optimization problems (1.1) and (1.2), render them to be non-standard. In particular note that the space of feasible functions in (1.1) and (1.2) is not a convex set and many convex optimization techniques and analysis cannot be applied here.

Indeed, we show that the results hold in a more general setting: Suppose  $J : L^2 \rightarrow \mathbb{R}^+$  is a nonnegative bounded operator on the space of  $L^2(\Omega)$  functions; that is, there exists a constant  $C$  such that

$$0 \leq J(g) \leq C \|g\|_2 \quad \text{for all} \quad g \in L^2.$$

Let  $\{f_i\}_{i=1}^N$  denote a set of solutions, corresponding to the number  $N$ , of the optimization problem

$$(1.3) \quad \{f_1, \dots, f_N\} = \operatorname{argmin}_{\tilde{f}_1, \dots, \tilde{f}_N} \sum_{i=1}^N \frac{1}{\mu} J(\tilde{f}_i) + \langle \tilde{f}_i, \hat{H} \tilde{f}_i \rangle \quad \text{s.t.} \quad \langle \tilde{f}_j, \tilde{f}_k \rangle = \delta_{jk}.$$

Let  $\langle F_N^T, \hat{H} F_N \rangle$  denote the  $N \times N$  matrix whose  $(j, k)$ -th entry is  $\langle f_j, \hat{H} f_k \rangle$  and let  $(\nu_1, \dots, \nu_N)$  denote its eigenvalues in non-decreasing order. Define the energy associated with this solutions by

$$(1.4) \quad E = \sum_{i=1}^N \langle f_i, \hat{H} f_i \rangle = \text{Tr}(\langle F_N^T, \hat{H} F_N \rangle) = \nu_1 + \dots + \nu_N.$$

Recall that the domain  $\Omega$  is bounded; in particular, the operator  $J(\cdot) = \|\cdot\|_1$  satisfies the above conditions. Therefore, the results shown for solutions  $f_1, \dots, f_N$  and the corresponding  $\nu_1, \dots, \nu_N$ , in particular hold for the CMs  $\psi_1, \dots, \psi_N$  and corresponding  $\sigma_1, \dots, \sigma_N$ . Nevertheless, the case  $J(\cdot) = \|\cdot\|_1$  is the most interesting application and the main motivation for considering this problem in the first place.

The remainder of this paper is as follows. In section §2, we show that  $E$  converges from above to the actual energy of the system as  $\mu \rightarrow \infty$ . Section §3 contains the main results of the paper: assuming  $\lambda_{N+1} > \lambda_N$ , as  $\mu \rightarrow \infty$ , the  $f_i$ 's converge to a sequence of unitary transformations of  $\phi_i$ 's, and the eigenvalues  $(\nu_1, \dots, \nu_N)$  converge to  $(\lambda_1, \dots, \lambda_N)$ . In section §4, we make some concluding remarks.

**2. Convergence of Energies.** In this section we show that  $E$  converges to the ground state energy  $E_0 = \sum_{j=1}^N \lambda_j$  as  $\mu \rightarrow \infty$ . Although the result of this section can be readily deduced from theorem 3.5, we have included it here for its independent interest and simplicity of argument.

First observe that

$$(2.1) \quad E_0 = \lambda_1 + \dots + \lambda_N = \sum_{i=1}^N \langle \phi_i, \hat{H} \phi_i \rangle \leq \sum_{i=1}^N \langle f_i, \hat{H} f_i \rangle = \nu_1 + \dots + \nu_N = E,$$

where we used property (1.1) for the inequality and equation (1.4) for the last equality.

Next choose  $\mu$  large enough such that

$$\frac{1}{\mu} \sum_{i=1}^N J(\phi_i) < \epsilon.$$

This is possible due to the boundedness of the operator  $J$  and the assumption that  $\|\phi_i\|_2 = 1$  for  $i = 1, \dots, N$ .

We have

$$(2.2) \quad \begin{aligned} E = \nu_1 + \dots + \nu_N &= \sum_{i=1}^N \langle f_i, \hat{H} f_i \rangle \leq \frac{1}{\mu} \sum_{i=1}^N J(f_i) + \sum_{i=1}^N \langle f_i, \hat{H} f_i \rangle \\ &\leq \frac{1}{\mu} \sum_{i=1}^N J(\phi_i) + \sum_{i=1}^N \langle \phi_i, \hat{H} \phi_i \rangle < \epsilon + \lambda_1 + \dots + \lambda_N = \epsilon + E_0, \end{aligned}$$

where we used (1.4) for the first equality, positivity of the operator  $J$  for the first inequality, and property (1.3) for the second inequality.

From equations (2.1) and (2.2) it follows that

$$E \downarrow E_0 \quad \text{as } \mu \rightarrow \infty.$$

**3. Consistency Results for  $L^1$  Regularization.** This section contains the main results of the paper. In theorem 3.5, we show that as  $\mu \rightarrow \infty$ , the eigenvalues of matrix  $\langle F_N^T, \hat{H} F_N \rangle$  converges to the first  $N$  eigenvalues of the Hamiltonian  $\hat{H}$ . This provides an affirmative answer to the conjecture stated in [6]. Furthermore, in theorem 3.7, we show that as  $\mu \rightarrow \infty$ , the solutions to the regularized optimization problem (1.3) converge to a sequence of unitary transformations of the eigenfunctions in  $H^1$  norm.

First we prove the following lemmas.

LEMMA 3.1. *Suppose that for  $i, j = 1, \dots, N$*

$$(3.1) \quad \sum_{k=1}^{\infty} a_{ik}^* a_{jk} = \delta_{ij}.$$

Then, for any  $k$ ,

$$\sum_{i=1}^N |a_{ik}|^2 \leq 1.$$

**Proof:**<sup>1</sup> It suffices to show the result for  $k = 1$  (by relabeling the indices, the result follows for other  $k$ 's). Let  $A$  be the  $N \times N$  matrix whose  $ij$ -th entry is  $a_{ij}$ ,  $1 \leq i, j \leq N$ . Hypothesis (3.1) implies that  $AA^\dagger \preceq I$  (i.e.  $I - AA^\dagger$  is positive semidefinite), where  $I$  is the  $N \times N$  identity matrix. Since  $AA^\dagger$  and  $A^\dagger A$  have the same spectrum, we conclude that  $A^\dagger A \preceq I$ . In particular, the first diagonal entry of  $I - A^\dagger A$ ,

$$1 - \sum_{i=1}^N |a_{i1}|^2$$

must be nonnegative. The result follows.  $\square$

LEMMA 3.2. *Suppose  $\lambda_N < \lambda_{N+1}$ . For any  $\epsilon > 0$  there exists an  $\epsilon_0$  such that if the set of orthonormal functions  $\{g_i\}_{i=1}^N$  satisfies*

$$\left| \sum_{i=1}^N \langle g_i, \hat{H} g_i \rangle - \sum_{i=1}^N \langle \phi_i, \hat{H} \phi_i \rangle \right| < \epsilon_0,$$

then  $\|g_i - \varphi_i\|_2 < \epsilon$ , for  $1 \leq i \leq N$ , where  $\varphi_1, \dots, \varphi_N$  is some unitary transformation of  $\phi_1, \dots, \phi_N$  depending on the  $g_i$ 's and  $\epsilon$ .

**Proof:** Since  $\{\phi_i\}_{i=1}^\infty$  form a complete set of orthonormal basis vectors in  $L^2$ , for every  $i = 1, \dots, N$ , we have

$$g_i = \sum_{k=1}^{\infty} a_{ik} \phi_k.$$

Moreover, the orthonormality assumptions on the  $g_i$ 's imply that

$$(3.2) \quad \sum_{k=1}^{\infty} a_{ik}^* a_{jk} = \delta_{ij}.$$

---

<sup>1</sup>The proof of lemma 3.1 was suggested by one of the anonymous referees, and is simpler than the proof that appeared in the first draft of this paper.

By spectral decomposition,

$$\hat{H} = \sum_{l=1}^{\infty} \lambda_l |\phi_l\rangle \langle \phi_l|.$$

Thus, for  $i = 1 \dots, N$

$$\langle g_i, \hat{H} g_i \rangle = \langle g_i, \sum_{l=1}^{\infty} \lambda_l |\phi_l\rangle \langle \phi_l| g_i \rangle = \sum_{l=1}^{\infty} \lambda_l \langle g_i, \phi_l \rangle \langle \phi_l, g_i \rangle = \sum_{l=1}^{\infty} \lambda_l |a_{il}|^2.$$

Summing over  $i$  on both sides of the above equality, we have

$$\sum_{i=1}^N \langle g_i, \hat{H} g_i \rangle = \sum_{i=1}^N \sum_{l=1}^{\infty} \lambda_l |a_{il}|^2 = \sum_{l=1}^{\infty} \lambda_l \sum_{i=1}^N |a_{il}|^2 = \sum_{l=1}^{\infty} \lambda_l b_l,$$

where

$$b_l := \sum_{i=1}^N |a_{il}|^2.$$

Observe that the  $b_l$ 's satisfy the following properties:

$$(3.3) \quad 0 \leq b_l \leq 1 \quad \text{for } l = 1, \dots,$$

and

$$(3.4) \quad \sum_{l=1}^{\infty} b_l = N.$$

We can see the first property using lemma 3.1. To see the second property note that

$$\sum_{l=1}^{\infty} b_l = \sum_{l=1}^{\infty} \sum_{i=1}^N |a_{il}|^2 = \sum_{i=1}^N \sum_{l=1}^{\infty} |a_{il}|^2 = \sum_{i=1}^N 1 = N.$$

Now observe that

$$\begin{aligned} & \sum_{i=1}^N \langle g_i, \hat{H} g_i \rangle - \sum_{i=1}^N \langle \phi_i, \hat{H} \phi_i \rangle \\ &= \sum_{l=1}^{\infty} \lambda_l b_l - (\lambda_1 + \dots + \lambda_N) \\ &\geq (b_1 - 1)\lambda_1 + \dots + (b_N - 1)\lambda_N + \left( \sum_{l=N+1}^{\infty} b_l \right) \lambda_{N+1} \\ &= (b_1 - 1)\lambda_1 + \dots + (b_N - 1)\lambda_N + (N - b_1 - \dots - b_N)\lambda_{N+1} \\ (3.5) \quad &= (1 - b_1)(\lambda_{N+1} - \lambda_1) + \dots + (1 - b_N)(\lambda_{N+1} - \lambda_N), \end{aligned}$$

where we used the nondecreasing ordering of the  $\lambda_i$ 's in the third line, and (3.4) in the fourth line. Now from (3.3) and nondecreasing ordering of  $\lambda_i$ 's, each of the terms in the last sum in (3.5) is positive. Hence

$$\left| \sum_{i=1}^N \langle g_i, \hat{H} g_i \rangle - \sum_{i=1}^N \langle \phi_i, \hat{H} \phi_i \rangle \right| \geq (1 - b_1)(\lambda_{N+1} - \lambda_1) + \dots + (1 - b_N)(\lambda_{N+1} - \lambda_N).$$

Now using the assumption that  $\lambda_{N+1}$  is strictly greater than  $\lambda_1, \dots, \lambda_N$ , we may conclude that for every  $\epsilon_1 > 0$ , there exist an  $\epsilon_0$  such that if the LHS of the above inequality is smaller than  $\epsilon_0$ , then

$$1 - b_l < \epsilon_1 \quad \text{for } l = 1, \dots, N.$$

Moreover, using equation (3.4), we conclude that

$$\sum_{l=N+1}^{\infty} b_l < N\epsilon_1 \implies \sum_{l=N+1}^{\infty} \sum_{i=1}^N |a_{il}|^2 = \sum_{i=1}^N \sum_{l=N+1}^{\infty} |a_{il}|^2 < N\epsilon_1.$$

In particular, for  $i = 1, \dots, N$ , we have

$$(3.6) \quad \sum_{l=N+1}^{\infty} |a_{il}|^2 < N\epsilon_1.$$

Next we show that the  $N \times N$  matrix  $\{a_{ik}\}_{i,k=1}^N$  is ‘‘almost’’ unitary in the sense that

$$\begin{aligned} \left| \delta_{ij} - \sum_{k=1}^N a_{ik}^* a_{jk} \right| &= \left| \sum_{k=N+1}^{\infty} a_{ik}^* a_{jk} \right| \leq \left( \sum_{k=N+1}^{\infty} |a_{ik}|^2 \right)^{1/2} \left( \sum_{k=N+1}^{\infty} |a_{jk}|^2 \right)^{1/2} \\ &< (N\epsilon_1)^{1/2} (N\epsilon_1)^{1/2} \\ &= N\epsilon_1, \end{aligned}$$

where we used (3.2) for the first equality, Cauchy-Schwarz for the first inequality, and (3.6) for the second inequality. We can orthonormalize the rows of the matrix  $\{a_{ik}\}_{i,k=1}^N$  using the Gram-Schmidt process, to form a new unitary matrix  $\{a'_{ik}\}_{i,k=1}^N$ . Indeed, because of the above inequality, for any  $\epsilon_2 > 0$ , we can choose  $\epsilon_1$  small enough such that

$$(3.7) \quad \sum_{k=1}^N |a_{ik} - a'_{ik}|^2 < \epsilon_2 \quad \text{for every } i = 1, \dots, N.$$

Now, for  $i = 1, \dots, N$ , set

$$\varphi_i = \sum_{k=1}^N a'_{ik} \phi_k.$$

Observe that

$$\|g_i - \varphi_i\|_2^2 = \sum_{k=1}^N |a_{ik} - a'_{ik}|^2 + \sum_{k=N+1}^{\infty} |a_{ik}|^2 < \epsilon_2 + N\epsilon_1,$$

where we used (3.7) and (3.6) for the inequality.

Hence, for any  $\epsilon > 0$ , we can choose  $\epsilon_1$  and  $\epsilon_2$  small enough such that  $\|g_i - \varphi_i\|_2 < \epsilon$  for  $i = 1, \dots, N$ . The result follows.  $\square$

Now, we prove the main results of this section.

**THEOREM 3.3.** *Assume  $\lambda_{N+1} > \lambda_N$ . For every  $\epsilon > 0$ , there exists a  $\mu_0$  such that for  $\mu > \mu_0$ , the solutions to the regularized optimization problem (1.3) satisfy*

$$\|f_i - \varphi_i\|_2 < \epsilon \quad \text{for} \quad i = 1, \dots, N,$$

where  $\varphi_1, \dots, \varphi_N$  is some unitary transformation of  $\phi_1, \dots, \phi_N$  depending on  $\mu$  and  $\epsilon$ .

**Proof:** For a given  $\epsilon$ , choose  $\epsilon_0$  as indicated by lemma 3.2. Choose  $\mu_0$  large enough such that

$$(3.8) \quad \frac{1}{\mu_0} \sum_{i=1}^N J(\phi_i) < \epsilon_0.$$

Let  $\mu > \mu_0$ . Observe that

$$\sum_{i=1}^N \langle \phi_i, \hat{H}\phi_i \rangle \leq \sum_{i=1}^N \langle f_i, \hat{H}f_i \rangle \leq \frac{1}{\mu} \sum_{i=1}^N J(f_i) + \sum_{i=1}^N \langle f_i, \hat{H}f_i \rangle \leq \frac{1}{\mu} \sum_{i=1}^N J(\phi_i) + \sum_{i=1}^N \langle \phi_i, \hat{H}\phi_i \rangle.$$

where we used property (1.1) for the first inequality, positivity of operator  $J$  for the second inequality, and property (1.3) for the last inequality. Hence, using (3.8), we have

$$\sum_{i=1}^N \langle \phi_i, \hat{H}\phi_i \rangle \leq \sum_{i=1}^N \langle f_i, \hat{H}f_i \rangle \leq \epsilon_0 + \sum_{i=1}^N \langle \phi_i, \hat{H}\phi_i \rangle,$$

which implies that

$$(3.9) \quad \left| \sum_{i=1}^N \langle f_i, \hat{H}f_i \rangle - \sum_{i=1}^N \langle \phi_i, \hat{H}\phi_i \rangle \right| < \epsilon_0.$$

Now applying lemma 3.2, completes the proof.  $\square$

**REMARK 3.4.** *The assumption  $\lambda_{N+1} > \lambda_N$  is essential. Because otherwise, we can have situations in which  $\psi_N$  converges to  $\phi_{N+1}$  in  $L^2$  norm and the result of theorem 3.3 would clearly not hold.*

**THEOREM 3.5.** *The Eigenvalues  $(\nu_1, \dots, \nu_N)$  of the matrix  $\langle F^T, \hat{H}F \rangle$ , converge to  $(\lambda_1, \dots, \lambda_N)$ , the eigenvalues of the Hamiltonian  $\hat{H}$ .*

**Proof:** Using the same notation as before, let  $\varphi_i$ , for  $i = 1, \dots, N$ , be the unitary transformation of the eigenfunctions  $\phi_1, \dots, \phi_N$  as described by theorem 3.3. Let  $\langle \Phi^T, \hat{H}\Phi \rangle$  be the  $N \times N$  matrix whose  $(j, k)$ -th entry is equal to  $\langle \varphi_j, \hat{H}\varphi_k \rangle$ . Note that both matrices  $\langle F^T, \hat{H}F \rangle$  and  $\langle \Phi^T, \hat{H}\Phi \rangle$  are not fixed and depend on the value of regularization parameter  $\mu$ ; however, we suppress the explicit dependence on  $\mu$  to avoid cumbersome notation. We show that the matrix  $\langle F^T, \hat{H}F \rangle$  converges to the matrix  $\langle \Phi^T, \hat{H}\Phi \rangle$  entry-wise. This yields the result because the matrix  $\langle \Phi^T, \hat{H}\Phi \rangle$  has the same eigenvalues  $(\lambda_1, \dots, \lambda_N)$  as the Hamiltonian, since the  $\varphi_i$ 's are unitary transformation of the eigenfunctions  $\phi_i$ 's. To that end, it suffices to show that

$$(3.10) \quad \langle f_i, \hat{H}f_j \rangle \rightarrow \langle \varphi_i, \hat{H}\varphi_j \rangle \quad \text{for} \quad i, j = 1, \dots, N.$$

Suppose

$$\varphi_i = \sum_{k=1}^N a_{ik}' \phi_k \quad \text{for} \quad i = 1, \dots, N,$$

and

$$f_i = \sum_{k=1}^{\infty} a_{ik} \phi_k \quad \text{for} \quad i = 1, \dots, N.$$

From the result of theorem 3.3 we know that

$$(3.11) \quad \|f_i - \varphi_i\|_2^2 = \sum_{k=1}^N |a_{ik} - a'_{ik}|^2 + \sum_{k=N+1}^{\infty} |a_{ik}|^2 \rightarrow 0, \quad \text{as} \quad \mu \rightarrow \infty.$$

This also implies that as  $\mu \rightarrow \infty$ , for  $i = 1, \dots, N$ ,

$$(3.12) \quad a_{ik} \rightarrow a'_{ik}, \quad \text{where } 1 \leq k \leq N, \quad \text{and} \quad a_{ik} \rightarrow 0, \quad \text{where } k > N.$$

Since  $\sum_{i=1}^N \langle \phi_i, \hat{H} \phi_i \rangle = \sum_{i=1}^N \langle \varphi_i, \hat{H} \varphi_i \rangle$ , from the proof of theorem 3.3 (i.e. equation (3.9)), we conclude that

$$\sum_{i=1}^N \langle f_i, \hat{H} f_i \rangle \rightarrow \sum_{i=1}^N \langle \varphi_i, \hat{H} \varphi_i \rangle \quad \text{as} \quad \mu \rightarrow \infty.$$

Rewriting the above expression we have

$$\sum_{i=1}^N \sum_{k=1}^{\infty} \lambda_k |a_{ik}|^2 \rightarrow \sum_{i=1}^N \sum_{k=1}^N \lambda_k |a'_{ik}|^2 \quad \text{as} \quad \mu \rightarrow \infty.$$

From (3.12) we conclude that for any finite  $M > N$ ,  $\sum_{i=1}^N \sum_{k=1}^M \lambda_k |a_{ik}|^2$  converges to  $\sum_{i=1}^N \sum_{k=1}^N \lambda_k |a'_{ik}|^2$  as  $\mu \rightarrow \infty$ . Hence,

$$(3.13) \quad \sum_{i=1}^N \sum_{k=M}^{\infty} \lambda_k |a_{ik}|^2 \rightarrow 0 \quad \mu \rightarrow \infty.$$

Now for  $i, j = 1, \dots, N$ , consider

$$(3.14) \quad \begin{aligned} \langle f_i, \hat{H} f_j \rangle - \langle \varphi_i, \hat{H} \varphi_j \rangle &= \sum_{k=1}^{\infty} \lambda_k a_{ik}^* a_{jk} - \sum_{k=1}^N \lambda_k a_{ik}^* a'_{jk} \\ &= \sum_{k=1}^N \lambda_k (a_{ik}^* a_{jk} - a_{ik}^* a'_{jk}) + \sum_{k=N+1}^{\infty} \lambda_k a_{ik}^* a_{jk} \end{aligned}$$

The first summation in line (3.14) goes to zero as  $\mu \rightarrow \infty$ , because of (3.12). If all  $\lambda_k$ 's are negative (and therefore bounded, as we have arranged them in nondecreasing order), then a simple application of Cauchy-Schwarz and using (3.11), implies that the second summation in line (3.14) also goes to zeros as  $\mu \rightarrow \infty$ . Otherwise, there exists a finite  $M$  greater than  $N$  such that for  $k > M$ ,  $\lambda_k$  is positive. Now we write the second summation in line (3.14) as

$$\sum_{k=N+1}^M \lambda_k a_{ik}^* a_{jk} + \sum_{k=M+1}^{\infty} \lambda_k a_{ik}^* a_{jk}.$$



Again the first summation in the above expression goes to zero as  $\mu \rightarrow \infty$  because of (3.12). For the second summation in the above expression, note that by Cauchy-Schwarz

$$\left| \sum_{k=M+1}^{\infty} \lambda_k a_{ik}^* a_{jk} \right| \leq \left( \sum_{k=M+1}^{\infty} \lambda_k |a_{ik}|^2 \right)^{1/2} \left( \sum_{k=M+1}^{\infty} \lambda_k |a_{jk}|^2 \right)^{1/2}.$$

The reason that we use  $\lambda_k$  instead of  $|\lambda_k|$  on the RHS is due to the assumption that  $\lambda_k$ 's are positive for  $k > M$ . Equation (3.13), in particular, yields that the RHS of the above expression goes to zero as  $\mu \rightarrow \infty$ . Thus, we have shown that the expression on line (3.14) goes to 0 as  $\mu \rightarrow \infty$ . This completes the proof.  $\square$

**REMARK 3.6.** *Observe that theorem 3.5 does not immediately follow from the result of theorem 3.3. This is due to the fact that the Hamiltonian  $\hat{H}$  is not generally a bounded operator on  $L^2$  functions.*

**THEOREM 3.7.** *Assume  $\lambda_{N+1} > \lambda_N$ . Suppose  $\{\mu_t\}_{t=1}^{\infty}$  is an increasing positive sequence of parameters with  $\mu_t \rightarrow \infty$ . Let  $\{f_i^t\}_{i=1}^N$  denote the solutions to the regularized optimization problem (1.3) with  $\mu = \mu_t$ . For every  $t$ , there exist a unitary transformation of  $\phi_1, \dots, \phi_N$ , denoted by  $\{\varphi_i^t\}_{i=1}^N$ , such that for  $i = 1, \dots, N$ ,*

$$\|f_i^t - \varphi_i^t\|_{H^1} \rightarrow 0 \quad \text{as} \quad t \rightarrow \infty.$$

**Proof:** It suffices to show the result for a fixed index  $i$ . To ease the notation we drop index  $i$ . Theorem 3.3 implies that there exists a sequence  $\{\varphi^t\}_{t=1}^{\infty}$  such that

$$(3.15) \quad \|f^t - \varphi^t\|_2 \rightarrow 0 \quad \text{as} \quad t \rightarrow \infty.$$

To show the result, it suffices to show that

$$\|\nabla f^t - \nabla \varphi^t\|_2 \rightarrow 0 \quad \text{as} \quad t \rightarrow \infty.$$

From the proof of theorem 3.5 (equation (3.10) with  $j = i$ ), we know that

$$(3.16) \quad \langle f^t, \hat{H} f^t \rangle \rightarrow \langle \varphi^t, \hat{H} \varphi^t \rangle \quad \text{as} \quad t \rightarrow \infty.$$

Applying integration by parts to (3.16) (recall that  $\Omega$  has periodic boundary) and using the assumption  $\|V(\mathbf{x})\|_{\infty} < \infty$  and equation (3.15), yield that

$$(3.17) \quad \int_{\Omega} |\nabla f^t|^2 \, d\mathbf{x} \rightarrow \int_{\Omega} |\nabla \varphi^t|^2 \, d\mathbf{x} \quad \text{as} \quad t \rightarrow \infty.$$

Now, observe that

$$\begin{aligned} & \int |\nabla f^t - \nabla \varphi^t|^2 \, d\mathbf{x} \\ &= \int |\nabla f^t|^2 \, d\mathbf{x} + \int |\nabla \varphi^t|^2 \, d\mathbf{x} - 2 \int \nabla f^t \cdot \nabla \varphi^t \, d\mathbf{x} \\ &= \int |\nabla f^t|^2 \, d\mathbf{x} + \int |\nabla \varphi^t|^2 \, d\mathbf{x} - 2 \int (\nabla f^t - \nabla \varphi^t) \cdot \nabla \varphi^t \, d\mathbf{x} - 2 \int |\nabla \varphi^t|^2 \, d\mathbf{x} \\ (3.18) \quad &= \left( \int |\nabla f^t|^2 \, d\mathbf{x} - \int |\nabla \varphi^t|^2 \, d\mathbf{x} \right) + 2 \int (f^t - \varphi^t) \Delta \varphi^t \, d\mathbf{x} \end{aligned}$$

As  $t \rightarrow \infty$ , the first term in (3.18) converges to 0 because of equation (3.17). On the other hand, because  $\varphi^t$  is formed from a unitary transformation of  $\phi_1, \dots, \phi_N$  and  $\|V(\mathbf{x})\|_\infty$  is bounded, there exists a constant  $C$  (independent of  $t$  but depending on  $\Omega$ ,  $N$  and  $\|V(\mathbf{x})\|_\infty$ ) such that

$$(3.19) \quad \|\Delta\varphi^t\|_2 < C \quad \forall t.$$

Applying Cauchy-Schwarz and using equations (3.15) and (3.19) imply that the second term in (3.18) converges to 0 as  $t \rightarrow \infty$ . This completes the proof.  $\square$

**4. Conclusion.** In this paper we prove consistency results for compressed modes introduced in [6]. Although we show that the results hold in a more general setting, the most important application of the results of this paper is for compressed modes.

In [6, 7], the authors pioneered the use of  $L^1$  regularization to compute modes that are spatially localized. We proved that, under some necessary assumptions on the spectrum of the Hamiltonian, as the regularization term in the non-convex optimization problem (1.2) vanishes, the compressed modes indeed converge to a sequence of Wannier-like functions. We also provided an affirmative proof for a conjecture in [6].

**Acknowledgments.** I am indebted to Professors Russel Caffisch and Stan Osher for invaluable guidance and helpful discussions and comments. I am thankful of the anonymous referees for their insightful remarks and helpful comments.

#### REFERENCES

- [1] E. J. CANDÈS, Y. C. ELДАР, T. STROHMER, AND V. VORONINSKI, *Phase retrieval via matrix completion*, SIAM Journal on Imaging Sciences, 6 (2013), pp. 199–225.
- [2] E. J. CANDÈS, J. ROMBERG, AND T. TAO, *Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information*, Information Theory, IEEE Transactions on, 52 (2006), pp. 489–509.
- [3] D. L. DONOHO, *Compressed sensing*, Information Theory, IEEE Transactions on, 52 (2006), pp. 1289–1306.
- [4] N. MARZARI AND D. VANDERBILT, *Maximally localized generalized Wannier functions for composite energy bands*, Physical Review B, 56 (1997), pp. 12847–12865.
- [5] L. J. NELSON, G. HART, F. ZHOU, AND V. OZOLIŃŠ, *Compressive sensing as a paradigm for building physics models*, Physical Review B, 87,035125 (2013).
- [6] V. OZOLIŃŠ, R. LAI, R. E. CAFLISCH, AND S. J. OSHER, *Compressed Modes for Variational Problems in Mathematics and Physics* Proceedings of the National Academy of Sciences, 110 (2013), pp. 18368–18373.
- [7] V. OZOLIŃŠ, R. LAI, R. E. CAFLISCH, AND S. J. OSHER, *Compressed plane waves—compactly supported multiresolution basis for the Laplace operator*, Proceedings of the National Academy of Sciences, 111 (2014), pp. 1691–1696.
- [8] B. RECHT, M. FAZEL, AND P. PARRILO, *Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization*, SIAM Review, 52 (2010), pp. 471–501.
- [9] H. SCHAEFFER, R. E. CAFLISCH, C. D. HAUCK, AND S. J. OSHER, *Sparse dynamics for partial differential equations*, Proceedings of the National Academy of Sciences, 110 (2013), pp. 6634–6639.
- [10] G. H. WANNIER, *The structure of electronic excitation levels in insulating crystals*, Physical Review, 52 (1937), pp. 0191–0197.