Global, nonlinear algorithm for inverting quantum-mechanical observations

J. M. Geremia and Herschel Rabitz
Department of Chemistry, Princeton University, Princeton, New Jersey 08544
(Received 12 March 2001; published 9 July 2001)

Inverting laboratory measurements of quantum-mechanical observables to recover the underlying molecular potential typically produces nonunique solutions. Without quantifying the full family of potentials consistent with the measurements, it is impossible to fully determine how experimental error and limited data affect the inversion, or to assess the quality of the recovered potential. Here, we present a global, nonlinear algorithm for extracting molecular potentials from measurements of quantum-mechanical observables. The method utilizes a mapping technique to learn the relationship between a broad domain of potentials and their resulting observables to facilitate the inversion. Once constructed, the maps reduce the arduous task of repeatedly solving the Schrödinger equation for each trial potential tested during the inversion and permit the use of normally expensive, global optimization procedures to thoroughly explore the distribution of potentials consistent with the data. As a demonstration, the new algorithm is applied to quantum collision cross sections to illustrate the effect of experimental error and finite resolution of the scattering observables on the recovered potential. A series of simulated inversions were performed to examine these issues along with the inversion of laboratory differential cross-section data for He+Ne scattering. These illustrations show that laboratory errors can have a nonlinear effect on the family of extracted potentials. Furthermore, the examples provide a benchmark for the capabilities of the proposed algorithm to stably reveal the full distribution of potentials consistent with the data. The algorithm may be applied to other observables and molecular systems with more spatial coordinates.

DOI: 10.1103/PhysRevA.64.022710 PACS number(s): 34.20.−b, 32.90.+a, 34.90.+q

I. INTRODUCTION

Quantitative knowledge of molecular potential-energy surfaces is essential for understanding many chemical and physical processes [1] as well as for accurately calculating their outcome. Obtaining quantitatively reliable potentials has typically been accomplished using either of two general approaches: (1) through *ab initio* calculations, or (2) by extraction from laboratory measurements of quantum-mechanical observables. In both cases, flexible surface representation techniques can maximize the outcome of these efforts. Although *ab initio* potentials are sometimes unable to account for the full level of experimental detail obtained with modern high-resolution spectroscopy and scattering techniques, quantum-chemistry calculations have reached a sophisticated level of development. The practice of inverting laboratory data, however, has not yet attained that degree of maturity.

Quantum-mechanical measurements typically depend upon the potential via subtle, nonlinear functional relationships that are system and observable dependent. As a result of these complications, and because different observables require specific machinery for both measuring and calculating them, no truly nonlinear, practical inversion procedure has been available. Instead, observable specific algorithms that often incorporate dynamical approximations or restrict the functional form of the potential have adapted. These algorithms can unwittingly be misguided by incorrect *a priori* information. Available inversion techniques include the Rydberg-Klein-Rees (RKR) method [2–4] for diatomic spectral problems, semiclassical [5,6] and quantum [6–9] methods for certain scattering processes, the exponential distorted wave [10] and sudden [11] approximations for rotationally inelastic atom-diatom scattering problems, etc. Functional sensitivity analysis [12,13] based algorithms [9], while applicable to arbitrary observables, are local in scope when searching for the potential. Linearization methods require a good initial guess, lest they become attracted to suboptimal extrema, or possibly diverge.

The inversion of laboratory measurements must consider various experimental factors, including noise and incomplete data, as well as the generally nonlinear impact of these factors upon the extracted potential. The route back to the potential from the data is not unique, producing an ill-posed problem [14], and there is typically a *family* of potentials that adequately reproduce the data to within its error. A further consequence of nonlinearity is that a normal distribution in the data does not invert to a normal, or even uniform, distribution of recovered potentials. Local, linear methods typically recover only one surface with, at best, an idealized treatment of error propagation. A *global* inversion should strive to identify the full family of solutions consistent with the data without restricting functional flexibility.

A general approach towards global inversion involves repeatedly solving the forward problem by calculating the observable Φ produced by trial potentials V until identifying one (or more) that reproduce the experimental data $\Phi^{(lab)}$. The objective is often expressed as an optimization process that minimizes the difference between the observed $\Phi^{(lab)}$ and the calculated $\Phi[V]$ data sets,

$$\min_{V} \|\Phi^{(lab)} - \Phi[V]\|^2 \tag{1}$$

over potential energy space. Every trial potential requires a new evaluation of Eq. (1), which nominally entails solving the Schrödinger equation.

The latter brute-force approach of repeatedly solving the forward problem works in principle— the family of potentials that satisfy the data will minimize Eq. (1), assuming that there are no significant systematic experimental errors. However, repeatedly solving the forward problem is often not practical; calculating the observable is frequently expensive and the net cost of performing global searches over large domains of potential energy space can be extremely prohibitive. Previously, the only practical fallback approaches required either restricting the functional form of the potential or linearizing Eq. (1) and searching it locally using gradient methods.

Overcoming the drawbacks of linearization necessitates nonlocal optimization of Eq. (1), but employing global search algorithms is too expensive if the observable must be explicitly computed for each trial potential. A high-speed bypass is needed to circumvent the arduous task of performing the repeated forward computations, while sacrificing neither accuracy nor nonlinearity in the potential-observable relationship. These goals can be accomplished using functional mapping methods recently developed for representing potential-observable relationships [15]. Maps, learned off line from a relatively small sample of intelligently chosen, representative potentials, can be used to quantitatively relate a candidate potential, as well as any other variables necessary to describe the system (e.g., mass, energy, scattering angle, etc.), to the observables of interest. The maps, generated by explicitly calculating the observables from the representative potentials, can encompass a broad domain (i.e., arbitrary potential shapes without resort to constrained forms) and retain the full nonlinearity of the underlying relationship.

Once constructed, high-quality forward maps of the underlying relationship can be rapidly interpolated to accurately provide the observable for any system in its domain. Map evaluation requires no further solutions of the Schrödinger equation and can be extremely fast compared to explicitly calculating the observable. For the examples in Sec. III, map evaluation is approximately 10^4 times faster than explicity solving the Schrödinger equation. Even more favorable scaling is anticipated for larger molecular systems. Global, nonlinear optimization algorithms can then be applied to Eq. (1) to provide a special capability for identifying the full family of potentials that reproduce the laboratory data. The combined nonlinearity of the mapping procedure and the use of global optimization methods all but eliminate the shortcomings of local inversions and provide the best family of inverted potentials consistent with the quantity and quality of the provided laboratory data.

Section II introduces the map-facilitated inversion algorithm by describing the map construction, the map inversion, and the machinery that connects them. Section III employs the algorithm to recover scattering potentials from simulated collision cross-section data. The study investigates solution multiplicity and the quality of the potential that can be recovered from incomplete data using model atomic scattering systems. As a second example, we invert actual laboratory measurements of elastic He+Ne differential scattering cross sections. These examples provide a perspective on scattering inversion and convenient test cases for assessing the performance of the nonlinear algorithm. Extension of the concept to more complex systems and spectral data is conceptually straightforward and will be presented in following works.

II. ALGORITHM

The nonlinear inversion algorithm can be broken down into two stages: (1) learning the potential-observable relationship by map construction, and (2) recovering the set of potentials that reproduce the data by map exploration. The first step requires calculating observables for a representative set of potentials that are then employed in a special interpolator to produce the full map. The second step requires a global optimization algorithm to find the full family of surfaces consistent with the observed data. The interplay between the two stages can be intricate and requires a communication architecture to interconnect the components. If the map domain does not contain any potentials that reproduce the laboratory data, the optimization component will fail. The algorithm must then expand the map domain to treat a broader, or different, region of potential energy space for exploration. Similarly, the inversion algorithm must verify that the map is accurate over its domain by occasionally testing it against the *truth* (i.e., an explicit computation of the observable) for quality control.

A. Learning the forward map

Learning the potential→observable map to quantitative accuracy is equivalent to deducing the underlying physical relationship. While an explicit connection between the observable Φ and potential, $V(\mathbf{r})$, is generally not known, it is an implicit composition of the Schrödinger equation and the observable's expectation value,

$$i\hbar \frac{\partial \psi}{\partial t} = [H_0 + V(\mathbf{r})]\psi, \tag{2a}$$

$$\Phi = \langle \psi | O | \psi \rangle, \tag{2b}$$

where H_0 is the portion of the Hamiltonian that is assumed known and not subject to identification, and O is the quantum-mechanical operator associated with the observed data to be inverted.

Although both Eqs. (2a) and (2b) are well understood, their composition to form a map,

$$\Phi = \Phi([V], \mathbf{e}) \tag{3}$$

is rarely known. $\Phi([V], \mathbf{e})$ is the potential-observable relationship expressed as a functional of the potential and a function of the extra variables, $\mathbf{e} \equiv \{e_1, \ldots, e_{N_e}\}$, also needed to specify the Hamiltonian or other necessary conditions. For example, $\mathbf{e}$ might include the mass of the system, the energy, the transition dipole matrix elements, or any other quantities that affect the observable besides the potential. The dual variables $(V, \mathbf{e})$ completely describe the quantum system, and $\Phi([V], \mathbf{e})$ provides a unique map to the observable.

Constructing the map requires sampling $\Phi([V],\mathbf{e})$ over $[V]$ and $\mathbf{e}$ with sufficient resolution to permit accurate interpolation. However, in its raw form, $\Phi([V],\mathbf{e})$ is not amenable to easy map construction since it contains a functional argument, $[V]$. In practice, the potential-observable relationship must be learned using numerical techniques that necessitates approximating $\Phi([V],\mathbf{e})$, with an ordinary function f,

$$([V],\mathbf{e}) \overset{f}{\mapsto} \Phi \tag{4}$$

that can be learned by discrete sampling. The transformation from $\Phi([V],\mathbf{e})$ to $f(\mathbf{v},\mathbf{e})$ is accomplished by replacing the functional argument $[V]$ with a collection of variables $\mathbf{v} \equiv \{v_1, \ldots, v_{N_v}\}$ that distinguish between the different members of the map's potential space domain. The precise definition of $\mathbf{v}$ is problem dependent, however, ensuring sufficient functional flexibility in $V(\mathbf{r})$ generally requires many potential space variables, $N_v \gg 1$, making f high dimensional. The number of extra variables N_e is typically small compared to N_v.

While there are physical arguments for maintaining the distinct identity of the potential and extra variables $\mathbf{v}$ and $\mathbf{e}$, respectively, there is no mathematical distinction between their role in the map function, $f(\mathbf{v},\mathbf{e})$. Defining a single input vector $\mathbf{x}$, that contains all the variables that affect the observable,

$$f(\mathbf{x}) \equiv f(\mathbf{v},\mathbf{e}), \quad \mathbf{x} \equiv \mathbf{v} \cup \mathbf{e}, \tag{5}$$

enables a more notationally compact formulation. $f(\mathbf{x})$ has the overall dimension $N = N_v + N_e$ and is defined over a specified domain $\mathbf{v} \in \mathbf{V}$ and $\mathbf{e} \in \mathbf{E}$, where $\mathbf{X} = \mathbf{V} \cup \mathbf{E}$.

In principle, the map can be learned by sampling $f(\mathbf{x})$ over its N variables and then interpolating between the sample points. However, in practice, full numerical resolution is impractical in light of the map function's high dimensionality. A regular mesh with S discrete points in each x_i would call for S^N sample points. For small S and N, exponential scaling is tolerable; however, representing molecular potential spaces can easily require $N \sim 10^2, 10^3, \ldots$, and $S \sim 10$ or more. Solving the Schrödinger equation $> 10^{100}$ times is not a practical approach to map construction. It is generally recognized that the exponential sampling complexity in N, referred to as the *curse of dimensionality*, renders the direct interpolation of N-dimensional functions impossible for systems of even modest dimension.

Exponential scaling in the dimension is traditionally handled by foregoing global coverage and resorting to local linearization; however, the map domains of interest here extend well beyond the local, perturbative regime, and it is critical to retain all the nonlinearities of $\Phi([V],\mathbf{e})$. Recently, it was demonstrated that nonlinear map construction is possible provided that both $V(\mathbf{r})$ and $f(\mathbf{x})$ are properly reformulated [15]. For expressing the potential, it is first beneficial to subdivide $V(\mathbf{r})$ into a large number N_v of component functions,

$$V(\mathbf{r}) = \sum_{i=1}^{N_v} V_i(\mathbf{r}), \tag{6}$$

where each $V_i(r)$ is a nonzero constant, v_i, only in a small Δ-neighborhood centered around its corresponding spatial point $\mathbf{r}_i$,

$$V_i(\mathbf{r}) = \begin{cases} v_i = V(\mathbf{r}_i): & |\mathbf{r} - \mathbf{r}_i| \leqslant \Delta \\ 0: & |\mathbf{r} - \mathbf{r}_i| > \Delta \end{cases}. \tag{7}$$

N_v can be made as large as necessary for the problem at hand, and in that fashion, $f(\mathbf{x})$ can well approximate the underlying functional since $\Delta \propto N_v^{-1}$. The only approximation results from assuming that $V_i(\mathbf{r})$ is constant, with the value v_i, over its nonzero Δ neighborhood. Defining the potential in this manner can also be extended to treating the v_i as interpolation points. Then, the potential is given by fitting a smooth surface through the v_i.

It is further beneficial to exactly reformulate $f(\mathbf{v},\mathbf{e})$ according to the following finite expansion ranging in dimension from 0 to N,

$$f(\mathbf{x}) = f_0 + \sum_{i=1}^{N} f_i(x_i) + \sum_{i<j} f_{ij}(x_i, x_j) + \cdots + f_{1\ldots N}(x_1, \ldots, x_N), \tag{8}$$

where f_0 is the nominal value of f over its domain, the $f_i(x_i)$ terms are single-variable functions that describe the individual effect of their respective variable x_i on the observable, the $f_{ij}(x_i, x_j)$ terms are bivariate functions that describe the *cooperative* effect of their variables (x_i, x_j) on the observable, etc. The final term, $f_{1\ldots N}$, captures any residual dependence of the observable on all of the variables acting in nonseparable cooperation with one another. Expansions of the form in Eq. (8) are referred to as high-dimensional model representation (HDMR), forming a family of multivariate representations used to capture the input→output relationships of high-dimensional physical systems [16–18].

The hierarchy of terms in Eq. (8) is assigned an ordering index L, that groups the functions according to dimensionality. $L=0$ corresponds to the constant term, f_0, $L=1$ corresponds to the single-variable functions, $\{f_i\}$, $L=2$ corresponds to $\{f_{ij}\}$, etc. If Eq. (8) must be taken to Nth order, then HDMR provides no relief from exponential scaling. It has been shown that Eq. (8) is expected to converge to low order, $L \ll N$ for quantum potential-observable relationships involving spatial potentials that have been reformulated according to Eq. (7) [15]. There is also extensive evidence that Eq. (8) converges to low order, for many other well-posed input→output relationships, particularly those derived from physical systems such as solar radiation driven atmospheric heating due to trace gas absorbers [19], chemical kinetics concentrations mapped over large time steps [20], semiconductor materials properties [17,21–23], and laboratory quantum control outcomes [24].

A low-order, converged map expansion can be truncated after its last significant order terms without sacrificing accuracy and can dramatically reduce the computational labor of map construction. The complexity of constructing an Lth-order map is given by that of attaining the Lth-order terms. Therefore, the total number of sample points scales as $O(SN), O(S^2N^2), \ldots$, for $L=1,2,\ldots$, respectively, resulting in only polynomic sampling complexity in N. The low order terms, although subdimensional, are not confined to be low-degree polynomials or have any particular form. Equation (8) is not analogous to a Taylor series in that regard, and can display all the relevant nonlinearity dictated by the particular problem.

The expansion functions are not unique and there are multiple, equivalent formulations of HDMR [16–18]. The method that is employed for the system of interest should be selected on the grounds of sampling convenience, and for the illustrations in Sec. III, the cut-center formulation of HDMR is utilized. In cut-center HDMR the expansion functions are constructed according to,

$$f_0 = f(\overline{x_1}, \overline{\ldots}, \overline{x_N}), \tag{9a}$$

$$f_i(x_i) = f(\overline{x_1}, \overline{\ldots}, x_i, \overline{\ldots}, \overline{x_N}) - f_0, \tag{9b}$$

$$f_{ij}(x_i, x_j) = f(\overline{x_1}, \overline{\ldots}, x_i, \overline{\ldots}, x_j, \overline{\ldots}, \overline{x_N}) - f_i - f_j - f_0, \tag{9c}$$

$$\vdots,$$

around an N-dimensional reference point $\overline{\mathbf{x}}$ referred to as the cut-center. $f(\mathbf{x})$ is invariant to the choice of $\overline{\mathbf{x}}$ provided that the expansion is taken to convergence. The notation, $\overline{\ldots}$, indicates that the corresponding arguments are taken as their reference value. Cut-center HDMR is often algorithmically efficient for problems of moderate dimension, $N \lesssim 10^2$. In very high dimensions it may be prudent to use an alternative formulation for determining the functions in Eq. (8) such as random sample (RS) HDMR that is based on Monte Carlo methods. In RS-HDMR, the evaluation count for sampling $\Phi([V],\mathbf{e})$ during map construction is independent of dimension [17].

Cooperativity between $\mathbf{v}$ and $\mathbf{e}$ is expected to remain; otherwise, the observable would display the same energy, mass, etc., behavior regardless of the underlying potential. Fortunately, the number of extra variables N_e is typically small and the sampling complexity is effectively determined by N_v. Even if second- or third-order terms arise between the extra variables $\mathbf{e}$ and their interaction with $\mathbf{v}$, map generation is computationally practical because it has been recast as a collection of lower-dimensional problems. The example in the next section provides an illustration where second-order terms between $\mathbf{v}$ and $\mathbf{e}$ remain significant when $N_v \gg 1$. In practice, it is sometimes advantageous to truncate Eq. (8) just prior to full convergence because the reduction in sampling effort far outweighs the introduction of tolerable map error and a practical algorithm operating with a set of maps can overcome the error in any single map.

B. Inverting the functional relationship

Identifying the family of potentials that reproduce the laboratory data using maps is accomplished by replacing explicit solutions of the forward problem in Eq. (1) with the map. This process nominally involves minimizing $\|\Phi^{(lab)} - f(\mathbf{x})\|^2$ over the potential space variables, $\mathbf{v} \in \mathbf{V} \subset \mathbf{X}$. During inversion, the extra variables $\mathbf{e}$ are not generally optimized; instead, $\mathbf{e}$ may describe other Hamiltonian variables, external experimental conditions, or how the data was collected (e.g., the reduced mass, energy, scattering angle, pressure, etc.). These conditions accompany the laboratory data, as will be discussed in Sec. II C, and it is useful to think of the laboratory data as being dependent on the extra variables, i.e., $\Phi^{(lab)} \equiv \Phi^{(lab)}(\mathbf{e})$.

Minimizing $\|\Phi^{(lab)}(\mathbf{e}) - f(\mathbf{v},\mathbf{e})\|^2$ over $\mathbf{v}$ provides the core of the inversion procedure, however, it does not properly treat experimental error or properly normalize the individual data members $\Phi_i^{(lab)}$ that can span orders of magnitude. Addressing the former issue, the quantity, $\|\Phi^{(lab)}(\mathbf{e}) - f(\mathbf{v},\mathbf{e})\|^2$ is zero only when the observable for the trial potential $f(\mathbf{v},\mathbf{e})$ exactly matches the laboratory data. Instead, any trial potential that reproduces $\Phi^{(lab)}$ to within its experimental error is equally valid. The best means of treating experimental error depends on the level of detail available about the data. Ideally, an error distribution function, $p[\Phi^{(lab)}]$, would be available and the inversion family would be the collection of potentials whose associated observable distribution matched that of the laboratory measurements. The result is a distribution of potentials with the most probable members around that confidence limits can be drawn.

In many cases, only sparing information about the nature of the error distribution is available as estimated error bars, $\varepsilon^{(lab)}$, without any qualification of how the error is distributed (i.e., normally distributed, or otherwise). When this is the case, it is often appropriate to assume a uniform error distribution between the hard limits, $\pm\varepsilon^{(lab)}$. The resulting inversion family is the hard bounded collection of potentials with associated observables that lie within $\Phi^{(lab)} \pm \varepsilon^{(lab)}$. In this case, it is still possible to identify the most probable member of the inversion family, even when detailed information about the laboratory error distribution is unavailable. However, due to the uniformity of the data error distribution, the other members of the potential family will all reproduce the data, as well as the most probable potential.

For the illustrations here, the laboratory uncertainty will be addressed by adopting a comparison function that ranks the trial potential by treating $\pm\varepsilon^{(lab)}$ as hard bounds with a uniform distribution between the limits, $\Phi^{(lab)} \pm \varepsilon^{(lab)}$. An example of an inversion where a Gaussian error distribution is utilized is presented in a separate work [25].

A suitable cost function that accounts for data error and the possibly large differences in magnitude between the members of the data set, is given by,

$$C(\mathbf{v})=\frac{1}{M}\sum_{i=1}^{M}\begin{cases}0: & |\Phi_i^{(lab)}(\mathbf{e}_i)-f_i(\mathbf{v},\mathbf{e}_i)|\leqslant\varepsilon_i^{(lab)}\\ \left\|\dfrac{\Phi_i^{(lab)}(\mathbf{e}_i)-f_i(\mathbf{v},\mathbf{e}_i)}{\Phi_i^{(lab)}(\mathbf{e}_i)}\right\|^2: & |\Phi_i^{(lab)}(\mathbf{e}_i)-f_i(\mathbf{v},\mathbf{e}_i)|>\varepsilon_i^{(lab)},\end{cases} \tag{10}$$

where $\varepsilon_i^{(lab)}$ is the uncertainty in the ith member in the data set $\Phi_i^{(lab)}\in\Phi^{(lab)}$ of size M. $f_i(\mathbf{v},\mathbf{e}_i)$ is the map predicted value for the ith data member [not to be confused with the first-order expansion function, $f_i(x_i)$]. Hidden in the notation is the fact that there can also be laboratory error in the measurement conditions $\mathbf{e}_i$, such as when $\mathbf{e}_i$ represents a collision energy, laser intensity, sample pressure, etc. In practice, $\Phi_i^{(lab)}$ is often an average or convolution over replicate measurements, and to reflect this, the quantities in Eq. (10) can be replaced by $\langle\Phi_i^{(lab)}\rangle_{\mathbf{e}}$ and $\langle f_i(\mathbf{v},\mathbf{e}_i)\rangle_{\mathbf{e}}$. The latter quantity is calculated by appropriately averaging or convolving over the extra map variables **e**. In this fashion, any errors in **e** are folded into the inversion process.

Equation (10) may accept potential functions that, although correctly reproducing the laboratory measurements, are not physically realistic. It is generally rarely possible to distinguish between real potentials (i.e., continuous, smooth, displaying proper asymptotic behavior, etc.) and other functions that coincidentally reproduce $\Phi^{(lab)}$ based on the data alone, and the inversion algorithm must explicitly deal with the possibility of unrealistic minima in Eq. (10). Naturally, the choice of characteristics that distinguish real potentials is a matter of physical judgment.

Physically unrealistic members of the solution family can be identified and removed using regularization methods [14,26] to incorporate properties, such as smoothness, into the recovered potentials without further restricting the functional flexibility of the inversion. Realistic potentials can be found by minimizing a cost function similar to

$$\mathcal{J}(\mathbf{v})=C(\mathbf{v})+\beta\|\hat{K}\mathbf{v}\|^2 \tag{11}$$

over **v**. $\hat{K}$ is a regularizing operator used to reward properties, such as smoothness, of the inverse solution and can take on many forms. Typically, $\hat{K}$ is a differential operator, particularly the second derivative of the potential with respect to its coordinates as this emphasizes smoothness (c.f., Sec. III). β is a constant that weights the importance of regularization relative to that of reproducing the data. Generally, it is best to make β extremely small so that the optimization of Eq. (11) first finds solutions that reproduce the data, $C(\mathbf{v})\rightarrow 0$. After the first term has vanished, the regularization component becomes significant forcing unrealistic functions that coincidentally reproduce the data to be abandoned.

Finally, it is of paramount importance that Eq. (11) be globally minimized. Local optimizers converge to a single, local potential and fail to take full advantage of the high-speed, nonlinear nature of $f(\mathbf{x})$. Here, genetic algorithms (GA's) [27] are employed to identify the full family of possibly very different, inverse solutions. GA's simultaneously identify multiple minima and are insensitive to starting values when operated under practical conditions. Unlike local searching methods, which are prone to divergence unless the initial guess is fortunate, GA's are stable with excellent convergence behavior because of elitism and population overlap making them ideal for complex landscapes such as those that arise when inverting quantum-mechanical data.

C. Full nonlinear inversion algorithm

Combining the concepts in Secs. II A and II B produces the full map-facilitated inversion algorithm, which is given in Fig. 1, and the labels (a)–(i) described below refer to this figure. As input, the algorithm requires (a) the data $\Phi^{(lab)}$, its laboratory error $\varepsilon^{(lab)}$, and its experimental conditions described by **e**, (b) a computational means (software treated as a black box by the algorithm) for evaluating $\Phi([V],\mathbf{e})$, and (c) the map construction details including the expected domain **X**, the number of samples per variable S_i, and the map order L. The latter starting parameters affect the efficiency of the inversion in two competing ways and must be chosen

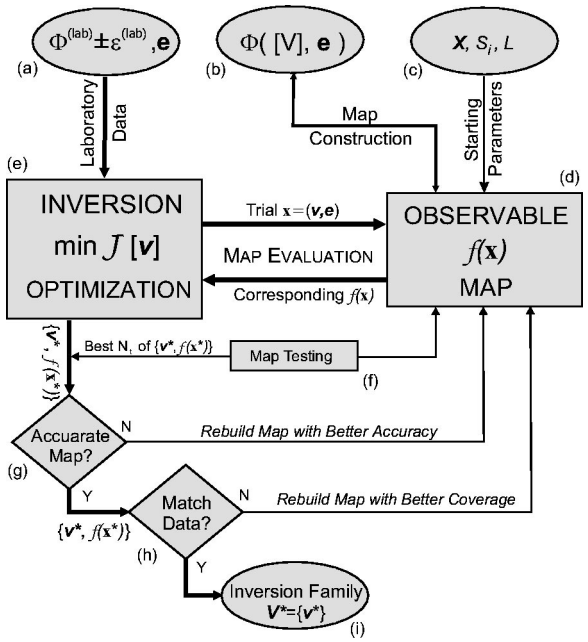

FIG. 1. A schematic of the nonlinear inversion algorithm capable of identifying the full family of potentials that reproduce the observed data to within its uncertainty. A complete description of the algorithm's components and operation is given in Sec. II. The procedure makes use of high-speed, nonlinear maps to replace the explicit potential-observable relationship in the inversion optimization. The efficiency of the maps permit the use of global searching methods such as genetic algorithms.

appropriately. If the domain **X** is too small, it is possible that potentials that satisfy the data will remain undetected because they lie outside the map. The inversion might fail to reproduce the data and would have to be repeated over a new region of potential space. In the alternate extreme, retaining accuracy over too large a domain might require increasing the order L, dramatically adding to the cost of learning the map. A balance that uses the largest possible domain for the lowest acceptable map order generally provides the best operating conditions.

After the map (d) is constructed by sampling $\Phi([V],\mathbf{e})$ over its domain (in some cases, a previously generated map might be reused), the inversion optimizer (e) minimizes Eq. (11) using $f(\mathbf{x})$ to rapidly compute the observable for each trial potential it considers. The GA population size N_p (an algorithmic parameter) is normally on the order of 10^3 and must be large to ensure that the full family of consistent potentials is identified. Other GA parameters such as the mutation and cross-over rates must be set to provide good exploration of Eq. (11), and are nominally given by the ranges, $r_m=[5\%,20\%]$ and $r_c=[60\%,80\%]$, respectively.

The GA minimization produces a collection of N_p optimized potential-energy functions; however, optimization alone does not guarantee potentials that reproduce the data. Two post-optimization tests must be performed to determine whether the inversion was successful. First, N_t (an algorithmic parameter) of the optimized potentials $\{\mathbf{v}_i^*\}$ are chosen and their associated observables, $f(\mathbf{v}_i^*,\mathbf{e}_i)$, are tested (f) for map error by comparing them to the *truth* by explicitly evaluating $\Phi([V],\mathbf{e})$. Any optimized potential for which $f(\mathbf{v}_i^*,\mathbf{e}_i)$ deviates from $\Phi([V],\mathbf{e})$ by more than $\varepsilon^{(lab)}$ is abandoned because it represents a false minimum. If the number of individuals that fail the testing process is greater than N_f (an algorithmic parameter), the map is deemed faulty (g) and is regenerated over a fraction N_r (an algorithmic parameter, typically a half) of its initial domain. This voting process (g) provides quality control and allows the algorithm to adaptively refine its map to ensure an accurate representation of $\Phi([V],\mathbf{e})$. The approach is based on the argument that $f(\mathbf{v},\mathbf{e})$ is generally more accurate for smaller domains [15]. A large GA population size N_p helps minimize the effect of finding isolated poor individuals if the map displays high overall accuracy.

If the GA output passes map testing, the observables for the corresponding potentials are compared to the laboratory data (b) in a second test (h). Those that match $\Phi^{(lab)}$ to within the experimental error are retained in the solution set $\{\mathbf{v}_i^*\}$ and those that do not are discarded. Although an optimized potential that violates any single member of the data set should be rejected (as is the case in Sec. III), this requirement can be relaxed, as desired. If none of the optimized potentials reproduce the data, the inversion failed and the entire map construction/optimization process is repeated by recentering the domain around the closest failed GA result (h). Additionally, if too few of the optimal potentials fail to match the data, the potential family might be poorly resolved. The domain recentering process should also be repeated if less than some fraction N^* (an algorithmic parameter, typically one half) of the optimized potentials are consistent with $\Phi^{(lab)}$. Inversion is completed once a satisfactory fraction of the optimized potentials $\{\mathbf{v}_i^*\}$ successfully reproduce the data.

D. The inverse solution family

The output of the inversion algorithm, Fig. 1(i), is a set containing the N_s potentials $\{\mathbf{v}_1^*,\ldots,\mathbf{v}_{N_s}^*\}$ that accurately reproduce $\Phi^{(lab)}$ to within its error, $\varepsilon^{(lab)}$. This optimal set is a discrete estimate of the full inverse family $\mathbf{V}^*$ or the subdomain of the map that is consistent with the laboratory data. $\mathbf{V}^*$ can be defined using the upper and lower bounds of each inverted variable,

$$^{<}\mathbf{v}^* = \min_m\{v_{m,i}^*\}, \tag{12a}$$

$$^{>}\mathbf{v}^* = \max_m\{v_{m,i}^*\}, \tag{12b}$$

where $v_{m,i}^*$ is the ith potential space variable from the mth member of $\mathbf{V}^*$. Provided that enough optimized potentials were found to completely resolve the inverse domain (which can be reasonably assured with $N_p \gg 1$), the true molecular potential should lie between the bounds $^{<}\mathbf{v}^*$ and $^{>}\mathbf{v}^*$. Considering the comments above Eq. (10) regarding data error propagation, the full interpretation of the bounds $^{<}\mathbf{v}^*$ and $^{>}\mathbf{v}^*$ depends on the information available about the data error distribution function, $p[\Phi^{(lab)}]$. Here, $\varepsilon^{(lab)}$ is considered a hard bound implying that $^{<}\mathbf{v}^*$ and $^{>}\mathbf{v}^*$ are the associated bounds within which the true potential will lie.

The potential-observable relationship is likely to remain nonlinear over the inverse region, and $\mathbf{V}^*$ will rarely display a normal (or *a priori* predictable) distribution of potentials consistent with the data even when the laboratory error distribution is simple, such as Gaussian or uniform [c.f., Figs. 8(C) and 8(D)]. The combined effect of nonlinearity between the potential and the observables and the fact that the inversion must simultaneously satisfy all of the members of the data set will typically produce a nontrivial distribution of potentials in the inversion family that might be irregular over its range. Normal, linear statistics will rarely be useful for analyzing the solution region over the distribution of recovered potentials, possibly even in the special case where the inversion family is small enough for the relationship between the potential and observables to remain linear.

The inversion quality, measured as the uncertainty in each potential space variable v_i^* can be assessed by considering its distribution $p_i(v_i^*)$ over the inversion family $\mathbf{V}^*$. The uncertainty in the recovered potential can be assessed by measuring the tightness of the solution domain as a function of the spatial coordinate $\mathbf{r}$ and the inversion is of overall better quality when the solution spread is small. When treating $\pm\varepsilon^{(lab)}$ as hard bounds, the most rigorous and conservative measure is given by,

$$\Delta v_i^* = {}^{>}\mathbf{v}^* - {}^{<}\mathbf{v}^*, \tag{13}$$

where each Δv_i^* is associated with its corresponding radial point $\mathbf{r}_i$ and the only firm conclusion is that the true potential lies between $^<\mathbf{v}^*$ and $^>\mathbf{v}^*$. The illustrations in Sec. III use Eq. (13) to assess the quality of the inverted potential; however, for experimental data where there is better knowledge of its error distribution, it might prove beneficial to utilize an alternative definition of Δv_i^*.

Finally, having a single metric to measure the quality of the inversion is convenient for comparing different solutions. It is useful to define an inversion quality figure of merit,

$$F_m = \sqrt{\sum_{i=1}^{N_v} \left(\frac{2\Delta v_i^*}{^>v_i^* + ^<v_i^*} \right)^2}, \tag{14}$$

where F_m is small for a good inversion and large for a poor one.

III. ILLUSTRATION: QUANTUM-MECHANICAL ELASTIC SCATTERING

As a demonstration of the nonlinear inversion algorithm, we considered elastic quantum-mechanical atom-atom scattering [28,29]. Elastic scattering provides a convenient benchmark and proof of concept for map-facilitated algorithms since even for atomic collisions, the potential $\rightarrow$observable map is highly nonlinear and the map is high dimensional [15]. The questions of solution multiplicity and the effect of data noise and incompleteness on atomic scattering inversions have never been substantially addressed.

Section III B considers simulated scattering data over which we vary the amount of data error, the extent and resolution of the data, and errors in **e** are neglected. The objective is to investigate the family of acceptable potentials in relation to errors in the data and the amount of data used in the inversion. To address these questions, three groups of inversions are performed with: (1) integral cross sections while varying the data errors and the scattering energy range, (2) differential cross sections while varying the data errors and the scattering angle range, and (3) three integrated (integral, differential, and viscosity) cross sections simultaneously. In Sec. III C we demonstrate an inversion using actual laboratory differential elastic cross-section data for He+Ne and identify the full range of potentials consistent with the data and show that a high-precision potential can be extracted.

A. Nonlinear scattering maps

We have recently demonstrated that global, nonlinear maps can be constructed for differential $\sigma(\theta)$, integral σ_T, diffusion σ_d, and viscosity σ_v elastic scattering cross sections over a wide range of scattering energies and angles. In these maps, the cross sections are written as functions of the reduced potential and reduced scattering energy,

$$v_i = U(r_i) = \frac{2\mu}{\hbar^2} V(r_i), \quad i = 1, \ldots, N_v \tag{15}$$

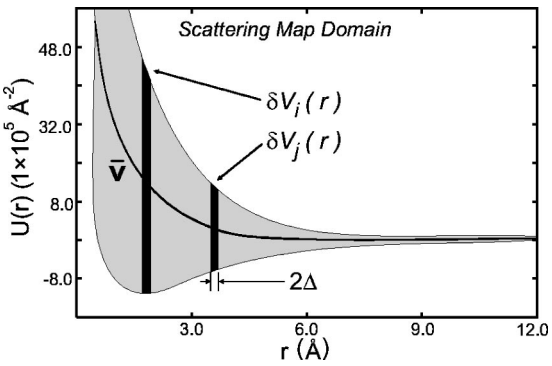

FIG. 2. The shaded region corresponds to the potential space domain for which elastic, integral, differential, effective diffusion, and effective viscosity cross sections were learned using global functional maps. Each map can accurately provide its corresponding cross section for *any* potential (smooth or irregular) in the shaded domain. The bars labeled $\delta U_i(r_i)$ and $\delta U_j(r_j)$ denote the dynamic range of the map variables, v_i and v_j at locations, r_i and r_j. The curve, $\bar{\mathbf{v}}$ denotes the reference potential utilized in the map generation process, as described in Sec. II.

$$e_1 = k^2 = \frac{2\mu}{\hbar^2} E \tag{16}$$

where the first $N-1$ map variables are used to specify the reduced potential, $x_i = v_i = V(r_i)$ for the fixed interatomic radial points, $\{r_1, \ldots, r_{N_v}\}$. The final variable, $x_N = e_1$, is the reduced energy k^2.

Specifics on how quantum scattering can be formulated in terms of Eqs. (15) and (16), a detailed analysis of map accuracy, and a physical argument for the fast convergence of the map expansion, Eq. (8), are provided in a preceding paper on the forward mapping process [15]. As a brief overview, differential $\sigma(\theta)$, integral σ_T, diffusion σ_d, and viscosity σ_v elastic scattering cross sections were mapped to high accuracy for all possible potential curves that can be drawn through the shaded region shown in Fig. 2. The bars, labeled $\delta U_i(r)$ and $\delta U_j(r)$ depict the extent of the potential space domain produced by varying each v_i and v_j over its respective window centered at r_i and r_j. The full domain is composed of N_v such bars and includes potentials with well depths ranging from those describing chemical bonding interactions, to Van der Waals interactions, and out to purely repulsive encounters. Scattering systems with $(2.4\times10^1 \leqslant k^2 \leqslant 2.8\times10^5)$ Å^{-2} and potential wells as deep as 8.7×10^5 Å^{-2} are described by a single global map. Based on currently known diatomic potentials, this domain likely covers all pairwise potentials involving main block elements of the Periodic Table. [30] However, if a potential outside the domain in Fig. 2 is needed, the maps can be expanded to include it with little extra effort.

For each observable, a global scattering map was constructed using 501 variables. $N_v = 500$ nodes provides excellent spatial resolution of $U(r)$, with $\Delta r = 0.024$ Å. In constructing the map, each x_i was sampled at 20 discrete points, and the differential cross section was resolved at 1° increments over $10° \leqslant \theta \leqslant 120°$; a separate map was constructed

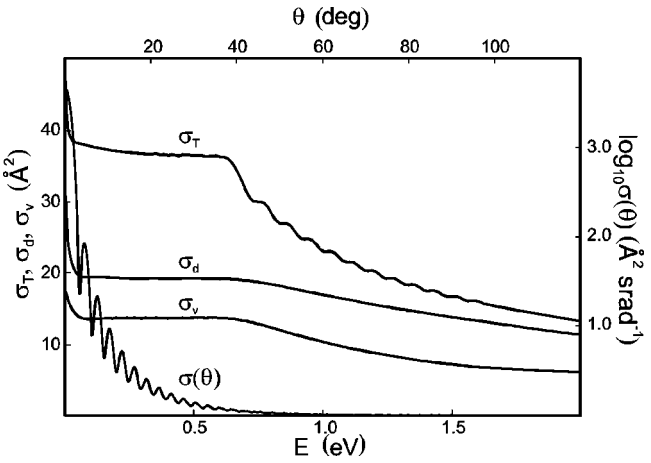

FIG. 3. Simulated cross-section data for the scattering system described in Sec. III B used to perform the inversion illustrations.

at each specified angle. The reference potential for the map [c.f., Eqs. (9a)–(9c)] was chosen as the geometric center of the domain, i.e., each $\overline{x_i}$ was selected as the center of its range, and the set of points $\overline{\mathbf{v}}$ defines the cut-center potential $\overline{V}(r)$ depicted in Fig. 2.

The maps were demonstrated accurate with prediction error $\lesssim 0.3\%$ for integral, effective diffusion, and effective viscosity cross sections (averaged over reduced energy) and to $\lesssim 1\%$ for differential cross sections (averaged over reduced energy and solid angle). The errors in these cross sections are smaller than the experimental uncertainty generally attributed to their laboratory determination [31].

B. Simulated nonlinear inversions

As a first illustration, consider a model Leonard-Jones (13),(16) interatomic potential of the form

$$V_{LJ}(r)=4\epsilon_{LJ}\left[\left(\frac{r}{r_0}\right)^{-12}-\left(\frac{r}{r_0}\right)^{-6}\right], \tag{17}$$

where ϵ_{LJ} is the well depth and r_0 designates the distance at which the potential changes from repulsive to attractive. The location of the well minimum is given by, $r_m=2^{1/6}r_0$. For the purpose of this example, ϵ_{LJ} was taken as 100.0 meV, and r_0 as 2.80 Å . The reduced mass μ was chosen to be 4.00 amu. The model potential, with its well minimum at 3.1 Å , is considerably different from the reference potential $\overline{\mathbf{v}}$ in Fig. 2.

Experimental data for the simulated inversions was generated by calculating the various cross sections produced by the potential in Eq. (17). These error-free differential, integral, viscosity, and diffusion cross-section data sets are depicted in Fig. 3 and were calculated from the model potential using standard methods [28,29,32]. Simulated noise was added to the error-free cross sections according to

$$\Phi^{(lab)}=\sigma^{LJ}(1+\rho\epsilon_\sigma), \tag{18}$$

where σ^{LJ} is the cross section of interest, $\{\sigma(\theta),\sigma_T,\sigma_d,\sigma_v\}$, ϵ_σ is the desired relative error, and ρ_i is a uniform random variable over $[-1,1]$.

1. Integral cross-section inversion

The first set of inversions are performed using integral cross sections as a function of energy $\sigma_T(E)$. Generating $f(\mathbf{x})$, produced a map accurate to $\lesssim 0.3\%$ error over the energy range, $20\leqslant E\leqslant 2000$ meV for the 4.00 amu scattering system as described above. The optimization step was performed by minimizing the cost functional in Eq. (11) with the regularization term

$$\hat{K}\mathbf{v}=\sum_{j=2}^{N_v-1}\left\|\frac{v_{j+1}-2v_j+v_{j-1}}{\Delta r^2}\right\|^2 \tag{19}$$

chosen to minimize the three-point discrete second derivative for smoothness. In case, the amount of data M was chosen so that the cross-section σ_T was sampled every 18 meV ($M=110$ for the full energy range, $20\leqslant E\leqslant 2000$ meV). The regularization constant β Å ^{4}/meV^2, was set to 1×10^{-15} and optimization was performed via an adaptive steady-state genetic algorithm [27] with a population size $N_p=500$ over 2000 generations. The map quality algorithmic parameters, (c.f., Sec. II C), were set to: $N_t=100$ and $N_f=5$, although map testing was not critical given the expanse and accuracy of the initial map. An acceptable inverse potential family was set at $N^*=1/2$, or 250 members of the GA population. In all of the present inversions, including that of He+Ne, the initial map was of sufficient quality and breadth that it was global. No additional maps were required.

The results from eleven inversions involving the integral cross section are depicted in Fig. 4. The shaded regions correspond to the subspaces of the map domain that adequately reproduce the cross-section data. They depict the limits $^{<}v_i^*$ and $^{>}v_i^*$ defined in Eqs. (12a) and (12b) for each radial point, r_i. In Fig. 4, and in the analogous Figs. 5–7, the true potential in Eq. (17) lies within the family of identified potentials.

Plots A1–A5 depict the inverse families consistent with data sets containing 0.1%, 1%, 2%, 3%, and 4% relative error, respectively. In each of the plots, A1–A5, the full energy range, 20–2000 meV (with $M=110$) was used for the inversion and the data error was simulated by setting ϵ_{σ_T} to 0.001, 0.01, 0.02, 0.3, and 0.04 in Eq. (18), respectively. A different random number ρ was used in simulating each data member.

Inversion quality naturally drops with increasing data error as it becomes easier to satisfy the less demanding data set; however, the error does not rise linearly with the data uncertainty. The figure of merit Eq. (14), metric, F_m, for the plots A1–A5 are $F_m=3.1\times10^{-2}$, 1.4×10^{-1}, 6.9×10^{-1}, 1.0×10^{1}, and 9.7×10^{1}, respectively, showing a steep decrease in inversion quality as the data uncertainty rises. This behavior provides an example of where the fully nonlinear algorithm in Fig. 1 is needed to properly perform the inver-

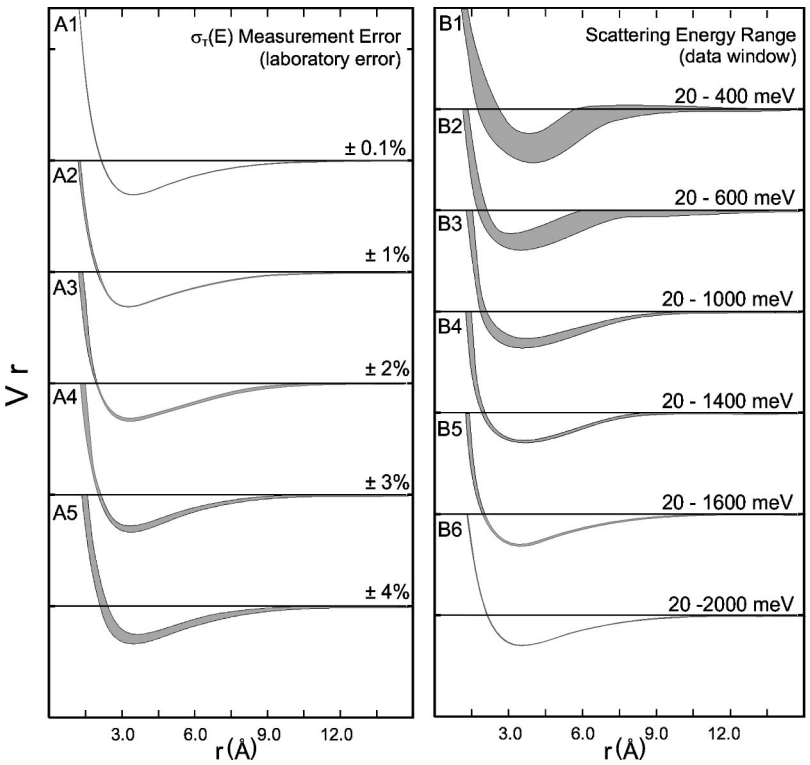

FIG. 4. The family of potentials recovered from finite, error contaminated integral cross-section data. For each case, any curve drawn through the shaded region reproduces all of the corresponding data to within its error. Panel A: the family of potentials recovered from data containing various degrees of relative error. Panel B: the family of potentials recovered from different amounts of data with $\pm 0.1\%$ error over different scattering energy ranges. Each potential is plotted in a -500 to 500 meV window.

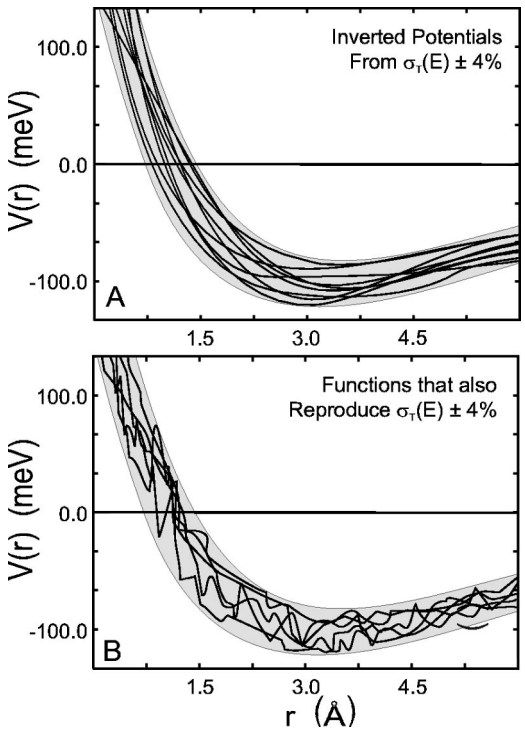

FIG. 5. Magnified view of the well region of the potential recovered from integral cross sections over the energy range, $20<E<1000$ meV with 4% relative error in the data [Fig. 4(A5)]. Both panels, (A) and (B), show potential curves that reproduce the scattering data to within its uncertainty. In (A) smoothness of the potential was required by employing regularization, while in (B) smoothness was not required.

sion and assess its quality. The quality of actual laboratory measurements of integral cross sections generally correspond to those in plots A2–A3 [31].

Plots B1–B6 depict the inverse families consistent with data sets measured over [20,400], [20,600], [20,1000], [20,1400], [20,1600], and [20,2000] meV, respectively. In each case, the inversion was performed using cross sections sampled every 18 meV, setting M accordingly. The data was simulated using the small error of 0.1% to focus on the impact of the energy range. Again, the solution regions were found by superimposing the inverse sets, $\{\mathbf{v}_i^*\}$, and finding the maximum and minimum value at each r_i.

The inversion quality rises sharply with an increase in the amount of data and the figures of merit corresponding to plots B1–B6 are $F_m = 2.1\times10^4$, 2.5×10^3, 3.0×10^2, 9.9×10^{-1}, 2.7×10^{-1}, and 3.3×10^{-2}, respectively (F_m is not exactly identical for A1 and B6 because of the nondeterministic nature of GA's, although the difference is entirely negligible). Of particular interest is the fact that the inversion quality increases suddenly as the energy range, $1\leqslant E\leqslant 2$ eV is covered in the inversion. Comparison with Fig. 3 shows that this energy range contains the majority of the structure in the σ_T plot. This observation supports the generally accepted qualitative notion that information content increases with the amount of structure in the data.

Finally, the role of regularization in Eq. (19) deserves some attention. Even though physically acceptable potentials are known to be smooth functions that display ubiquitous qualitative behavior, e.g. $V(r)\to 0$ as $r\to\infty$ and $V(r)\to\infty$ as $r\to 0$, etc., it is illustrative to see whether other functions, not necessarily resembling typical potentials, can also satisfy

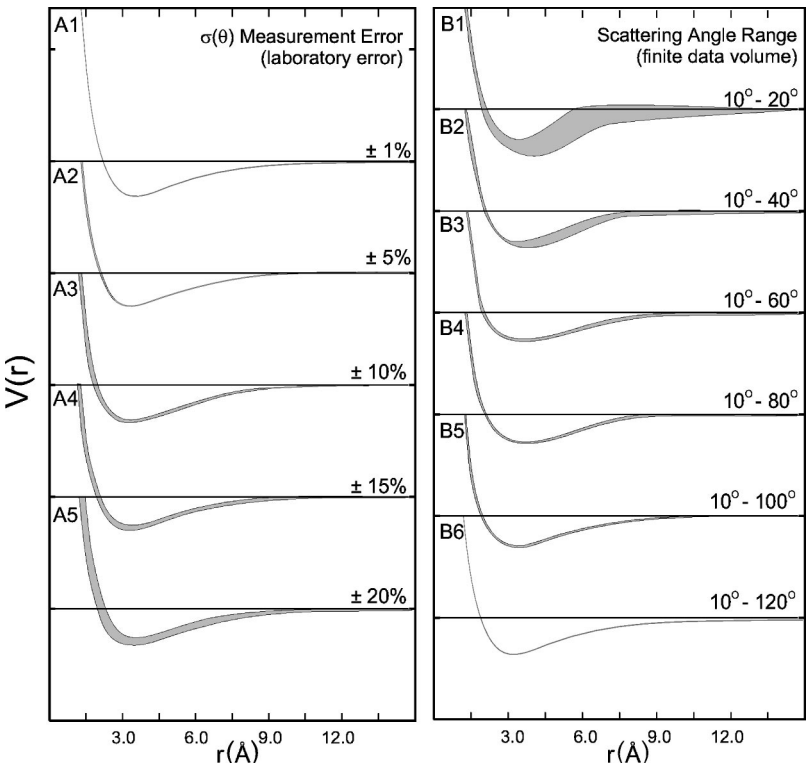

FIG. 6. The family of potentials recovered from finite, error-contaminated differential cross-section data. For each case, any curve drawn through the shaded region reproduces all of the corresponding data to within its error. Panel (A): the family of potentials recovered from data containing various degrees of relative error. Panel (B): the family of potentials recovered from different amounts of data containing $\pm 1\%$ error. Each potential is plotted in a -500 to 500 meV window.

the observed data. Such a search can be accomplished by turning off regularization in the optimization cost function, $\mathcal{J}$ by setting $\beta=0$ in Eq. (11).

Figure 5(A) shows a magnified view of the inverse family $\mathbf{V}^*=\{\mathbf{v}_i^*\}$ from the well portion of the solution region found in Fig. 4(A5) where regularization was performed. Although the sample of potentials plotted inside the region appear to differ at this level of detail, they are all smooth functions that, overall, behave as expected and produce cross sections consistent with the data. In contrast, Fig. 5(B), which depicts the same region of the potential, shows the inversion results when regularization was not performed. These curves, which are highly nonsmooth and quite unusual as potentials, do reproduce the data to within its precision. Of course, while the functions in Fig. 5(B) are unrealistic, the exercise in Fig. 5(A) and Fig. 5(B) emphasizes an important reminder that it is dangerous to assume too much *a priori* information when dealing with potentials; pseudo *a priori* guidance could lead to what appears to be a unique surface, but in fact is one of many possible acceptable potentials.

2. Differential cross-section inversion

A set of eleven inversions, analogous to those in Sec. III B 1 were performed using differential cross sections as a function of angle, $\sigma(\theta)$, at a fixed energy, $E=500$ meV. The differential cross-section maps (one for each sampled angle) were shown to be accurate to $\lesssim 1\%$ error over the angles $10°\leqslant\theta\leqslant 120°$ sampled at 1° increments, or $M=110$ cross-section samples over the full range [15]. The cost functional in Eq. (11) also employed the regularization in Eq. (5). All algorithmic conditions were identical to those in Sec. III B 1.

The results from eleven sample inversions involving the differential cross section are depicted in Fig. 6. Plots A1–A5 depict the inverse families consistent with data sets containing 1%, 5%, 10%, 15%, and 20% relative error, respectively. In each of A1–A5, the full angle range, $10°\leqslant\theta\leqslant 120°$ (with $M=110$) was used for the inversion and the data error was simulated by setting ϵ_{σ_T} accordingly in Eq. (18) and using a different ρ for each data member.

As with the integral cross section data, the inversion quality drops as the data error increases. Again, it does not scale linearly in the data uncertainty. The corresponding figure of merit, Eq. (14), F_m, for the plots A1–A5 are $F_m=8.3\times10^{-2}$, 3.0×10^{-1}, 1.0×10^{0}, 5.5×10^{0}, and 2.8×10^{1}, respectively, showing a steep drop in inversion quality as the data uncertainty rises. Actual laboratory measurements of differential cross sections generally correspond to data errors of $\sim 10-20\%$ (c.f., plots A3–A5) and often higher error [31].

Plots B1–B6 depict the families of potentials recovered from data sets with $\pm 1\%$ error measured over scattering angles of [10°,20°], [10°,40°], [10°,60°], [10°,80°], [10°,100°], and [10°,120°], respectively. The inversion was performed using cross sections sampled every 1°, setting M accordingly in $\mathcal{J}(\mathbf{v})$. The figures of merit corresponding to plots B1–B6 are $F_m=8.1\times10^{4}$, 2.7×10^{2}, 1.9×10^{0}, 5.7×10^{0}, 9.9×10^{-2}, and 8.1×10^{-2}, respectively.

Compared to the integral cross section in Sec. III B 1, the information content of $\sigma(\theta)$ appears more uniform over its independent variable, θ. There is no sudden increase in inversion quality over the data window, and this observation is consistent with the differential cross section plot in Fig. 3 where the structure is well distributed. Inversions using $\sigma(\theta)$ also appear more tolerant to experimental uncertainty than their integral counterpart [compare Fig. 4(A5) to Fig. 6(A5)]. This tolerance may be attributable to the complex structure

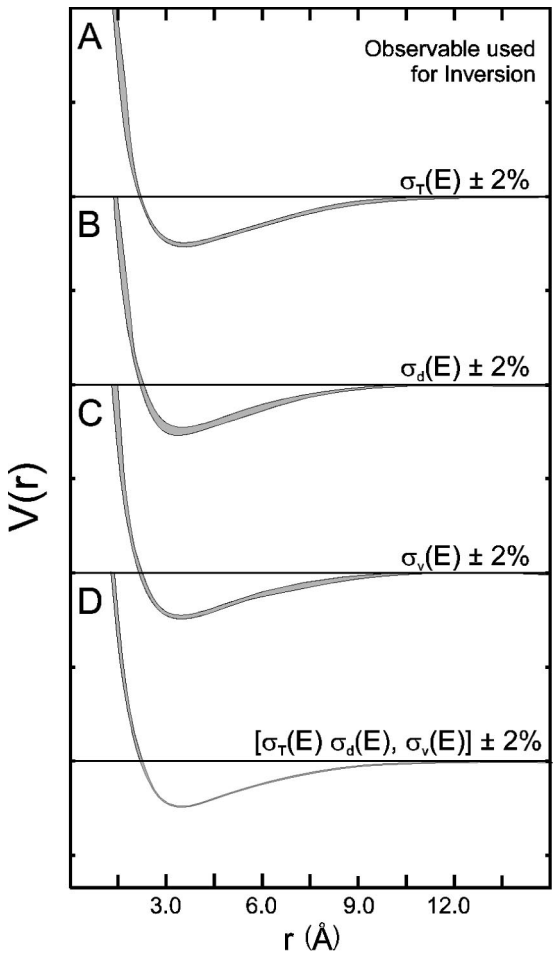

FIG. 7. Family of potentials recovered using different combinations of scattering cross section data, all accurate to $\pm 2\%$ relative error: (A) integral cross section only, (B) diffusion cross section only, (C) viscosity cross section only, (D) all three cross sections simultaneously. In each case, any curve drawn through the shaded regions reproduces the data to within its error. Using the combined data produces a higher quality inversion result.

found in the differential cross-section data as a function of angle and is consistent with the observation that $\sigma(\theta)$ is more sensitive to the potential [15].

3. Combined observable inversion

In the final illustration, the objective was to invert a data set containing multiple simultaneous cross sections, $\Phi^{(lab)} = \{\sigma_T, \sigma_d, \sigma_v\}$, over their full range. A cost functional identical to that in Eq. (19) was adopted. All algorithmic conditions were the same as in Secs. III B 1 and III B 2. Plots A, B, and C in Fig. 7 demonstrate the inverse families obtained using the integral σ_T, diffusion σ_d, and viscosity σ_v, cross sections individually. In each case, the data sets contained 110 cross sections accurate to 2% relative error over the energy range, $20 \leqslant E \leqslant 2000$ meV producing a full data set with $M = 330$ measurements. The figures of merit for the individual inversions performed in A, B, and C are $F_m = 6.9 \times 10^{-1}$, 1.8×10^{0}, and 4.2×10^{-1}, respectively. Figure 7(D) depicts the result of the combined data inversion where the figure of merit was found to be $F_m = 1.6 \times 10^{-1}$. Here, combining data sets containing 2% relative error enabled an inversion of comparable quality to the 1% integral cross section data set in Fig. 4(A2), supporting the notion that diverse data, even of limited quality, generally enables a better inversion.

C. He-Ne inversion

The nonlinear algorithm in Fig. 1 is specifically designed to address the issues that arise when dealing with actual laboratory experiments. As an example of an inversion using real data, we considered the low-energy elastic differential scattering of He+Ne in a crossed-beam experiment by Kiel *et al.* [33]. The scattering cross section $\Phi_i^{(lab)} = \sigma(\theta_i; E)$ is available for a finite set of $M = 88$ angles $\{\theta_i\}$ at a relative collision energy of $E = 29.2$ meV. The standard deviations (i.e., treated as bounds here), used to define $\varepsilon_i^{(lab)}$, are also available from the measurements. As was the case with previous inversions using this data set [9], errors introduced by the experiment's angular apparatus function were considered negligible.

The inversion was performed using the same global scattering map employed in the previous section without additional solutions of the Schrödinger equation, except for quality testing, c.f., Fig. 1(f). The large expanse of the map's potential space domain enabled a truly global inversion. A cost functional identical to Eqs. (10) and (5) was used, with $M = 88$ data members at angles corresponding to the points in Fig. 8(B). All algorithmic parameters and the value of the regularization constant were the same as those used above. The full He+Ne inversion, using the already learned global map, required 8.1s on an 800 MHz $2 \times i686$ PIII machine.

The results of the He+Ne inversion are depicted in Figs. 8(A)–(D). Any potential drawn through the shaded region of Fig. 8(A) reproduces every member of the measured data set to within its precision. The inset plot in Fig. 8(A) provides a detailed view of the well, including 10 members $\{\mathbf{v}_1^*, \ldots, \mathbf{v}_{10}^*\}$ selected from the full inverse family containing 480 members as the output from the algorithm in Fig. 1. As suggested by the ten sample potentials, the inversion family distribution is not uniform. Plots of the full distribution over $\mathbf{V}^*$ for two of the potential space variables, v_{10} and v_{15} corresponding to the radial points, $r_{10} = 2.72$ Å and $r_{15} = 3.12$ Å are shown in Figs. 8(C) and 8(D). The distributions, $p_{10}(v_{10}^*)$ and $p_{15}(v_{15}^*)$, reflect the nonlinearity of the potential-observable relationship over the members of the inversion family, and demonstrate that linear statistics can be very misleading.

Figure 8(B) depicts the differential cross sections produced by the family of recovered potentials (shaded region) and the error bars correspond to those of the actual laboratory data points [33]. The algorithm in Fig. 1 guided the inversion such that the width of the calculated observable distribution did not violate any of these error bars. In this regard, the data errors around the oscillation extrema and the small angle data provide the primary information to constrain on the inversion family. The inset plot provides a closer look at the differential cross section over $25° \leqslant \theta$

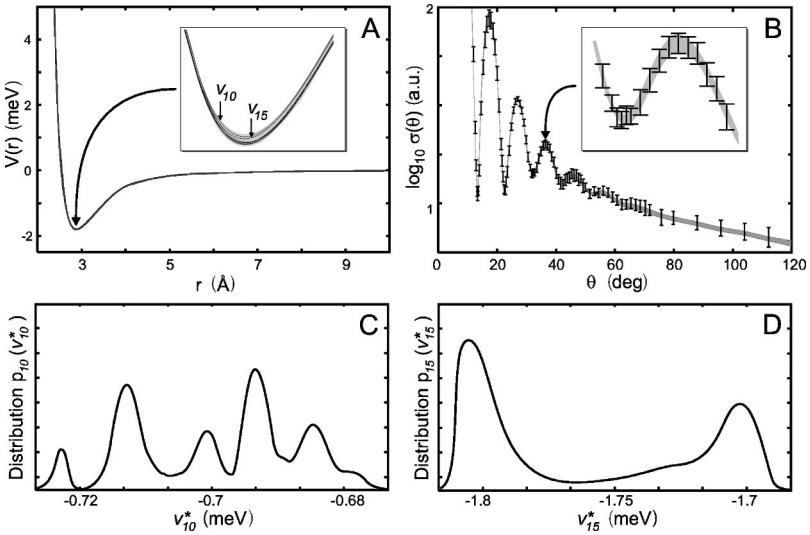

FIG. 8. Results for a global inversion using the laboratory He+Ne elastic scattering data of Keil *et al.* [33]. Any potential drawn through the inversion family (shaded region) in (A) reproduces the experimental data to within its uncertainty, and the expanded view of the well region depicts 10 members from the full family. (B) illustrates the range of cross sections produced by the inversion family in (A) as well as the 88 laboratory data points, with their reported uncertainties. (C) and (D) depict the distribution of values, $p_{10}(v^*_{10})$ and $p_{15}(v^*_{15})$, for the extracted potentials, v^*_{10} and v^*_{15} at the radial points, $r_{10}=2.72$ Å and $r_{15}=3.12$ Å [c.f., the inset in (A)], over the inversion family.

$\leqslant 35°$ where the cross sections produced by the inversion family completely fill the experimental error bars. This behavior should be contrasted with the large angle scattering, where the distribution is narrower than the experimental uncertainty. Requiring that all the data members must be satisfied constrains the inversion family more than the error in any individual measurement might imply.

IV. CONCLUSION

This paper presented a general, nonlinear map-facilitated procedure for inverting laboratory measurements of quantum observables to recover the family of potentials consistent with the data. Handling the nonlinear nature of most potential-observable relationships, finite amounts of data, and the presence of experimental error requires an algorithm capable of incorporating nonuniqueness in the inversion. The procedure presented above is capable of identifying a large (if not the full) solution space of potentials consistent with the data.

The illustrations demonstrated that a linear treatment can be incomplete, and the newly presented algorithm overcame the shortcomings of local techniques by adopting a global search. Both simulated quantum elastic scattering cross-section data and actual He+Ne elastic differential cross sections were successfully inverted. For the simulations, the quality of the inverted potential was investigated in terms of the amount of data (i.e., energy range for the integral, diffusion, and viscosity cross sections and angular range for the differential cross section) and the reliability of the data (i.e., its experimental error). It was also shown that some regions of the data (i.e., specific energy ranges, in particular) can contain significantly more detailed information about the potential than others. The simultaneous inversion of multiple data sets corresponding to different observables (i.e., for integral, diffusion and viscosity cross sections) was shown to be beneficial. Finally, a global inversion was performed on the He+Ne data.

The nonlinear procedure introduced here has a general structure and may be applied to other inversion and optimization processes. Future applications include extracting potentials from spectral data and the inversion of more complex scattering systems. Extending the mapping procedure to temporal potentials for applications in quantum control is also being investigated [24]. Although global optimization has previously been too expensive to perform on general problems, these works demonstrate how to replace the arduous task of repeatedly solving the associated forward problem with an accurate and efficient forward map. In the examples of Sec. III, the repeated use of the same map for inverting different sets of data points out that suitably developed maps may serve an archival role making them available for multiple applications with little extra expense associated with solving the Schrödinger equation. This overall map facilitated inversion concept is anticipated to have broad applicability to many problems.

ACKNOWLEDGMENTS

The authors acknowledge support from the NSF and the DoD.

[1] R. Levine and R. B. Bernstein, *Molecular Reaction Dynamics and Chemical Reactivity* (Oxford University Press, New York, 1987).
[2] R. Rydberg, Z. Phys. **73**, 376 (1931).
[3] O. Klein, Z. Phys. **76**, 226 (1932).
[4] A. L. G. Rees, Proc. Phys. Soc. Jpn. **59**, 998 (1947).
[5] W. Miller, J. Chem. Phys. **59**, 3631 (1969).
[6] U. Buck, Rev. Mod. Phys. **46**, 369 (1974).
[7] R. B. Gerber, M. Shapiro, U. Buck, and J. Schleusene, Phys. Rev. Lett. **41**, 236 (1968).
[8] T. -S. Ho and H. Rabitz, J. Chem. Phys. **90**, 1519 (1989).
[9] T. -S. Ho and H. Rabitz, J. Phys. Chem. **97**, 13 449 (1993).

[10] R. B. Gerber, V. Buch, U. Buck, J. Maneke, and G. Schleusene, Phys. Rev. Lett. **44**, 1937 (1980).
[11] R. Schinke, J. Chem. Phys. **73**, 6117 (1980).
[12] P. Frank, *Introduction to System Sensitivity Theory* (Academic, New York, 1978).
[13] V. Volterra, *Theory of Functionals* (Dover, New York, 1959).
[14] A. Tihkonov and V. Arsenin, *Solutions of Ill-posed Problems* (Winston/Wiley, Washington, 1977).
[15] J. Geremia, C. Rosenthal, and H. Rabitz, J. Chem. Phys. **114**, 9825 (2001).
[16] O. Alis and H. Rabitz, J. Math. Chem. **25**, 197 (1999).
[17] H. Rabitz, O. Alis, J. Shorter, and K. Shim, Comput. Phys. Commun. **11**, 117 (1999).
[18] *Mathematical Methods for Sensitivity Analysis*, edited by H. Rabitz, A. Alis, O. in Satelli, K. Chan, and M. Scott (Wiley, New York, 2000).
[19] J. I. Shorter and H. Rabitz, Geophys. Res. Lett. **27**, 3485 (2000).
[20] J. Shorter, P. Ip, and H. Rabitz, J. Phys. Chem. A **36**, 7192 (1999).
[21] H. Rabitz and K. Shim, J. Chem. Phys. **111**, 10 640 (1998).
[22] K. Shim and H. Rabitz Phys. Rev. B **57**, 12 874 (1998).
[23] H. Rabitz and K. Shim, J. Chem. Phys. **111**, 10 640 (1999).
[24] J. Geremia, E. Weiss, and H. Rabitz, Chem. Phys. **267**, 209 (2001).
[25] J. Geremia and H. Rabitz (unpublished).
[26] K. J. Miller, Math. Ann. **52**, 1 (1970).
[27] D. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning* (Addison-Wesley, Reading, MA, 1989).
[28] M. Child, *Molecular Collision Theory* (Academic, London, 1974).
[29] R. B. Bernstein, *Atom-Molecule Collision Theory* (Plenum Press, New York, 1979).
[30] K. Huber and G. Hertzberg, *Constants of Diatomic Molecules* (Van Nostrand, Princeton, 1979).
[31] *Atomic and Molecular Beam Methods 1 and 2*, edited by G. Scoles (Oxford University Press, New York, 1992).
[32] L. I. Schiff, *Quantum Mechanics: Third Edition* (McGraw-Hill, New York, 1968).
[33] M. Keil, L. Danielson, U. Buck, J. Schleusener, F. Huisken, and T. Dingle, J. Chem. Phys. **89**, 2866 (1988).