Error bounds for molecular Hamiltonians inverted from experimental data

J. M. Geremia* and Herschel Rabitz
Department of Chemistry, Princeton University, Princeton, New Jersey 08544
(Received 30 August 2002; published 28 February 2003)

Inverting experimental data provides a powerful technique for obtaining information about molecular Hamiltonians. However, rigorously quantifying how laboratory error propagates through the inversion algorithm has always presented a challenge. In this paper, we develop an inversion algorithm that realistically treats experimental error. It propagates the distribution of observed laboratory measurements into a family of Hamiltonians that are statistically consistent with the distribution of the data. This algorithm is built upon the formalism of map-facilitated inversion to alleviate computational expense and permit the use of powerful nonlinear optimization algorithms. Its capabilities are demonstrated by identifying inversion families for the $X\,^1\Sigma_g^+$ and $a\,^3\Sigma_u^+$ states of Na_2 that are consistent with the laboratory data.

DOI: 10.1103/PhysRevA.67.022711 PACS number(s): 34.20.−b

I. INTRODUCTION

In order to predict molecular processes, it is first necessary to have quantitative knowledge of the Hamiltonian governing the system's behavior [1]. As such, constructing molecular potential energy surfaces has been a longstanding area of active research. Despite advances in *ab initio* techniques, many of the most powerful methods for determining molecular potentials extract them from laboratory measurements. Data inversion, however, is nontrivial for two reasons. First, the connection between the Hamiltonian and its associated observables is nonlinear and indirect—the Schrödinger equation is buried within it. Second, laboratory data contain experimental error.

Providing a realistic analysis of how laboratory uncertainty propagates through an inversion algorithm is challenging because the experimental error must be treated on the same footing as the data. Inversion is often computationally expensive and it is even more costly to map the experimental error back to a proper characterization of the Hamiltonian uncertainty. Furthermore, these issues are complicated by the nonlinear character of the relationship between the molecular Hamiltonian and the data. Consequently, traditional inversion methods have provided only approximate error analyses. For example, Ho *et al.* [2] addressed error in the $X\,^1\Sigma_g^+$ and $a\,^3\Sigma_u^+$ state potentials for Na_2 using a Monte Carlo treatment of laboratory noise, but resorted to linearization.

In order to rigorously treat experimental error, it is necessary to *invert the entire distribution of measured laboratory data*, not simply the mean values of the observables. From a laboratory perspective, data values that differ within the experimental precision are indistinguishable. The relative confidence in any particular value is described by a probability distribution (often Gaussian) that results from the experimental statistics. In order for data inversion to reflect these properties of the measurements, it should similarly provide a family of *indistinguishable* potentials whose confidence is described by a probability distribution. The important feature of the inversion family is that any potential from within it should reproduce the experimental data to within the laboratory precision. Similarly, potentials from outside the family should correspond to observables that lie outside the experimental tolerance. Furthermore, a potential from the inversion family that has a high inversion confidence should correspond to a value of the observable that has a high laboratory confidence. A family of potentials that meets these criteria will be referred to as *consistent with the data*.

This definition of *consistency* introduces a new means of expressing inversion that involves identifying a family that contains many distinct Hamiltonians. Obtaining an inversion family requires a global and nonlinear inversion algorithm because a Gaussian (or any other) distribution in the data need not correspond to a Gaussian (or similar) distribution of inverted potentials. This nonintuitive feature of Hamiltonian inversions results because the relationship between potential-energy space and its associated observables is nonlinear.

The need for global, nonlinear inversion capabilities places strict demands on the computational efficiency of the inversion algorithm. Resolving the inversion family can require extracting hundreds of distinct Hamiltonians from the data to reveal the distribution. It is therefore important to adopt map-facilitated inversion techniques [3–5] that have been specifically developed for finding solution families. Functional maps [6] alleviate the expense of repeatedly solving the Schrödinger equation, making it possible to identify the full, nonlinear inversion family.

In this paper, we extend the formalism of map-facilitated inversion to treat a distribution of laboratory data and obtain a full family of consistent potentials. To the best of our knowledge, this is the first general inversion algorithm to treat quantum-mechanical experimental data in such a manner. The inversion algorithm is presented in Sec. II, and a demonstration of the procedure is provided in Sec. III. We invert the rovibrational spectra of Babaky and Hussein [7] and Barrow *et al.* [8] for the $X\,^1\Sigma_g^+$ state potential of Na_2 and the spectrum of Li *et al.* [9] for the $a\,^3\Sigma_u^+$ state. The purpose of treating this collection of data is to demonstrate the concept of inverting a distribution, rather than just inverting the mean data. We also provide evidence that error

*Present address: Norman Bridge Laboratory of Physics, Mail Code 12-33, California Institute of Technology, Pasadena, CA 91125. Electronic address: jgeremia@Caltech.EDU

propagation can be nonlinear, even when the data are of high laboratory precision.

II. ALGORITHM

This section adapts a recently introduced map-facilitated data inversion method [3,4] in order to invert a full probability distribution of laboratory data. The procedure can be separated into two stages. First, a numerical potential $\rightarrow$observable map [6] is learned by solving the Schrödinger equation for a small, but judiciously chosen, collection of representative potentials. Then, the inversion is performed as an optimization over the map, which provides a highly accurate proxy for the Schrödinger equation. Evaluating the map is fast, and it permits the use of global, nonlinear search algorithms. The optimization result is the full family of potentials consistent with the distribution of the data.

A. Learning the forward map

The direct connection between the potential $V(r)$ and Φ, for example, the spectrum, can rarely be formulated in closed form. Mapping the relationship implicitly involves solving the Schrödinger equation and then computing the relevant observables,

$$[H_0+V(r)]|\psi\rangle=E|\psi\rangle, \tag{1a}$$

$$\Phi_m=\langle\psi|\hat{O}_m|\psi\rangle, \tag{1b}$$

where H_0 is the portion of the Hamiltonian that is assumed known and not subject to identification. The Hermitian operators $\hat{O}_m$ correspond to the laboratory observables (e.g., spectral lines) being inverted.

The composition of Eqs. (1a) and (1b) to form a mapping f between potential-energy space $\mathbf{V}$ and its corresponding observables,

$$f[V]:V\rightarrow\Phi, \quad V\in\mathbf{V}, \tag{2}$$

is a functional (indicated by the $[\cdot]$ notation) of the input potential-energy function V. Evaluating Eq. (2) implicitly requires solving the Schrödinger equation and computing the desired observables according to Eqs. (1a) and (1b).

In practice, numerical methods are employed to construct the map, and $f[V]$ must be approximated by replacing the function input $V(r)$ with a collection of variables, $\mathbf{v}\equiv\{v_1,\ldots,v_{N_v}\}$, that distinguish between different members of potential-energy space. This transformation results in an N_v-dimensional function $f(\mathbf{v})$ or a *map*, from potentials to observables. The particular meaning of the potential space variables v_i depends on the nature of the Hamiltonian being inverted. For example, in Sec. III, the variables are the values of the Na_2 potentials at fixed interatomic separations [e.g., Eq. (10)].

Constructing the map used in the inversion optimization is accomplished by tabulating $f(\mathbf{v})$ for a representative sample of potentials, $\{\mathbf{v}_1,\ldots\}$. Generally, it is impossible to numerically resolve f on a full N_v-dimensional grid (as is common for low dimensional functions) because representing potential space with sufficient flexibility often requires $N_v\gg 1$. Consequently, it is necessary to employ a more judicious procedure for sampling and interpolating $f(\mathbf{v})$, if linearization is to be avoided.

It has recently been demonstrated [6] that the nonlinear map can be constructed by reformulating $f(\mathbf{v})$ as a finite expansion,

$$f(\mathbf{v})=f_0+\sum_{i=1}^{N_v}f_i(v_i)+\sum_{i<j}^{N_v}f_{ij}(v_i,v_j)+\cdots +f_{1,\ldots,N_v}(v_1,\ldots,v_{N_v}), \tag{3}$$

where f_0 is a constant term, $\{f_i(v_i)\}$ are single-variable functions, $\{f_{ij}(v_i,v_j)\}$ are bivariate functions, etc. Expansions of the form in Eq. (3) are referred to as high-dimensional model representation. They form a family of multivariate representations used to capture the input$\rightarrow$output relationships of many high-dimensional physical systems [5,6,10–16]. It is important to note that Eq. (3) is not a Taylor series. It can be truncated without sacrificing nonlinearity because the low order terms, although subdimensional, can display arbitrary nonlinearity.

The hierarchy of terms in Eq. (3) can be assigned an ordering index L that groups the functions according to dimensionality. $L=0$ corresponds to the constant f_0, $L=1$ includes only the single-variable functions $\{f_i\}$, $L=2$ corresponds to the $\{f_{ij}\}$, etc. The physical interpretation of the expansion functions is that they represent the collective effects of the potential space variables in order of increasing intervariable cooperation. For example, the $f_i(v_i)$ terms describe how v_i, alone, affects the value of the physical observable. The bivariate functions, $f_{ij}(v_i,v_j)$, reflect the simultaneous influence of v_i, v_j, and so on. The final term, $f_{1,\ldots,N_v}$, captures any residual dependence of the observable on all of the variables acting in nonseparable cooperation.

It has been shown that Eq. (3) converges to low order, $L\ll N_v$ for quantum potential-observable relationships with a properly chosen form for the variables v_i [6]. A low order, converged map expansion can be truncated after its last significant order without sacrificing accuracy or nonlinearity. This dramatically reduces the computational labor of map construction. For example, 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 $O(SN),O(S^2N_v^2),\ldots$, for $L=1,2,\ldots$, respectively, resulting in only polynomial sampling complexity in N_v. Constructing the potential$\rightarrow$observable map can be achieved with feasible computational expense. Specifics for how the individual expansion functions should be evaluated can be found in previous papers [3–6].

B. Inverting the functional relationship

The objective of the inversion is to identify the family of potentials that are consistent with the laboratory probability distributions, $q^{(lab)}[\Phi]=\{q_i^{(lab)}[\Phi_i]\}$, $i=1,\ldots,M$, for each individual observable in the full data set, $\Phi^{(lab)}=\{\Phi_1^{(lab)},\ldots,\Phi_M^{(lab)}\}$. In many cases, the collection of dis-

tributions $q^{(lab)}[\Phi]$, are Gaussian and characterized by their standard deviations $\varepsilon_i^{(lab)}$. In other instances, $q^{(lab)}[\Phi]$ might be non-normal or possibly only characterized by hard bounds $\pm\varepsilon_i^{(lab)}$.

During the inversion search, rather than investigating individual surfaces, every trial solution is a full family of potentials, $\tilde{\mathbf{V}}=\{\tilde{\mathbf{v}}_1,\ldots,\tilde{\mathbf{v}}_{N_s}\}$ [17], along with its associated set of observables, $f(\tilde{\mathbf{V}})=\{f_i(\tilde{\mathbf{v}}_s)\}$, $i=1,\ldots,M$, and $s=1,\ldots,N_s$ [18]. For every trial family, the distributions of observables, $q_i[\Phi_i;f_i(\tilde{\mathbf{V}})]$, are computed by evaluating the map over the family members, $\{f_i(\tilde{\mathbf{v}}_1),\ldots,f_i(\tilde{\mathbf{v}}_{N_s})\}$, and accumulating the distribution by binning, or other suitable techniques. In practice, the number of family members, N_s, must be large enough to numerically resolve $q_i[\Phi_i;f_i(\tilde{\mathbf{V}})]$.

Minimizing the difference between the data distribution, $q_i^{(lab)}[\Phi_i]$, and map-computed distribution, $q_i[\Phi_i;f_i(\tilde{\mathbf{V}})]$, over the $i=1,\ldots,M$ distinct observables,

$$\mathcal{J}_0(\tilde{\mathbf{V}})=\frac{1}{M}\sum_{i=1}^{M}\int |q_i^{(lab)}[\Phi_i]-q_i[\Phi_i;f_i(\tilde{\mathbf{V}})]|^2\,d\Phi_i, \tag{4}$$

produces the inversion family by morphing the family until the distributions, $q_i^{(lab)}[\Phi_i]$ and $q_i[\Phi_i;f_i(\tilde{\mathbf{V}})]$, are the same (or as close as possible). However, without further constraints, Eq. (4) may insufficiently discriminate between realistic potentials (i.e., smooth, asymptotically correct, etc.) and others that, although consistent with the data, are physically unrealistic. It is often impossible to identify realistic potentials based on the data alone.

The choice of characteristics that distinguish a physically acceptable potential is a matter of judgment, and the inversion algorithm must provide flexible discrimination. Incorporating properties, such as smoothness, without restricting the functional representation of the potential can be accomplished using regularization methods [19,20]. Regularization introduces additional operators,

$$\mathcal{J}(\tilde{\mathbf{V}})=\mathcal{J}_0(\tilde{\mathbf{V}})+\beta\sum_{s=1}^{N_s}\|\hat{K}\tilde{\mathbf{v}}_s\|^2, \tag{5}$$

that reward favorable properties, such as smoothness. $\hat{K}$ can take on many forms, but is often a differential operator, particularly the second derivative of the potential with respect to its coordinates (cf., 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 small so that minimizing Eq. (5) first finds families consistent with the data, $\mathcal{J}_0(\tilde{\mathbf{v}})\to 0$. After the first term has vanished, the regularization component becomes significant and forces unrealistic functions that coincidentally reproduce the data to be abandoned.

Taking full advantage of the high-speed, nonlinear nature of $f(\mathbf{v})$ allows Eq. (5) to be globally minimized. Genetic algorithms (GAs) [21] are attractive because they simultaneously propagate multiple families and are relatively insensitive to starting values. Unlike local searching methods, which can be prone to divergence unless the initial guess is fortunate, GAs 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. The inverse solution family

Using a GA to minimize Eq. (5) produces a population of N_p optimized families, $\{\mathbf{V}_1^*,\ldots,\mathbf{V}_{N_p}^*\}$. Each contains N_s potentials, $\mathbf{V}_p^*=\{\mathbf{v}_{p,1}^*,\ldots,\mathbf{v}_{p,N_s}^*\}$, $p=1,\ldots,N_p$, with associated distributions, $q_i[\Phi_i;f_i(\mathbf{V}_p^*)]$ [22]. Normally, the most fit member of the GA population, $\mathbf{V}_1^*$, provides the final, optimal inversion family, denoted $\mathbf{V}^*$, and the remainder of the population, $p=2,\ldots,N_p$ is ignored. Section II D will discuss the algorithm used to decide which GA population member best provides $\mathbf{V}^*$.

Once the family $\mathbf{V}^*$ has been identified, it is useful to define another set of distributions, $p_i[v_i^*]$, that describe the probability of values for each potential space variable v_i^* over the inversion family. These new distributions, accumulated by binning, are implicitly related to $q^{(lab)}[\Phi]$ and $q[\Phi;f(\mathbf{V}^*)]$ through the inversion. The $p_i[v_i^*]$ reflect the nonlinearity of the potential-observable relationship and are rarely predictable even when the observable distributions, $q_i^{(lab)}[\Phi_i]$, are simple, such as Gaussian or uniform [cf., Figs. 2(b) and 2(c) and Figs. 3(b) and 3(c)]. The combined effect of nonlinearity between the potential and the observable and the fact that the inversion must simultaneously be consistent with all of the members of the data set often produces complex structure in the distribution. Using linear statistics, such as the variance, is generally inappropriate for characterizing the $p_i[v_i^*]$.

The distributions, $p_i[v_i^*]$, can be used to identify the best estimate of each potential space variable v_i by computing the expectation values,

$$\langle v_i^*\rangle=\frac{1}{N_s}\sum_{s=1}^{N_s}\mathbf{v}_{s,i}^*p_i[v_i^*],\quad i=1,\ldots,N_v \tag{6}$$

over $\mathbf{V}^*$. The potential $\langle\mathbf{v}^*\rangle$ produced by all of the expectation values in Eq. (6) provides the best estimate of the actual molecular potential. Bounds on the inversion family can be defined as the confidence limits,

$${}^{<}v_i^*\equiv\inf_v\left\{v\left|\int_v^{\langle v_i^*\rangle}p_i[v_i^*]dv_i\leqslant\frac{\epsilon}{2}\right.\right\}, \tag{7a}$$

$${}^{>}v_i^*\equiv\sup_v\left\{v\left|\int_{\langle v_i^*\rangle}^{v}p_i[v_i^*]dv_i\leqslant\frac{\epsilon}{2}\right.\right\}, \tag{7b}$$

where ϵ is the desired confidence. Equations (7a) and (7b) define the upper and lower bounds, ${}^{<}\mathbf{v}^*$ and ${}^{>}\mathbf{v}^*$, of the inversion family, and the probability of finding the actual potential between ${}^{<}\mathbf{v}^*$ and ${}^{>}\mathbf{v}^*$ is ε. The uncertainty in the

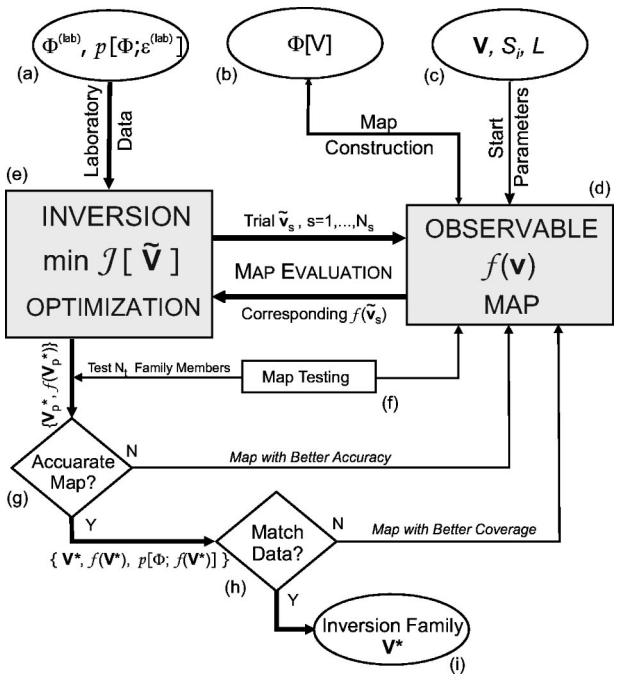

FIG. 1. A schematic representation of the nonlinear inversion algorithm that is capable of identifying the full family of potentials that reproduce the observed data to within its uncertainty. A detailed description of the algorithm is given in Sec. II. The procedure makes use of high-speed, nonlinear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

inversion family as a function of its coordinates r is given by the distance between its corresponding upper and lower family bounds,

$$\Delta v_i^* = {}^{>}v_i^* - {}^{<}v_i^* , \tag{8}$$

where each Δv_i^* is associated with its corresponding coordinate r_i.

D. Full nonlinear inversion algorithm

Combining the concepts in Secs. II A and II B produces the full map-facilitated inversion algorithm in Fig. 1, and the labels (a)–(i) described below refer to this figure. As input, the algorithm requires (a) the data members $\Phi_i^{(lab)}$ and their laboratory distributions $q_i^{(lab)}[\Phi_i]$, (b) a computational means (software treated as a black box by the algorithm) for solving the forward problem, $\Phi[V]$, and (c) the map construction details including the expected domain $\mathbf{V}$, the number of samples per map variable, S_i, and the map order L. The starting parameters affect the efficiency of the inversion in two competing ways and must be chosen appropriately. If the domain $\mathbf{V}$ is too small, it is possible that potentials that satisfy the data will remain undetected because they lie outside the map. Similarly, the inversion might fail to reproduce the data and would need to be repeated over a new region of potential space. In the alternate extreme, maintaining accuracy over too large a map 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 $\Phi[V]$ over its domain (in some cases, a previously generated map might be reused), the inversion optimizer (e) minimizes Eq. (5) using $f(\mathbf{v})$ to rapidly compute the distribution of observables for each trial family of potentials $\tilde{\mathbf{V}}$ it considers. The family size N_s (an algorithmic parameter) is normally on the order of 10^3 and must be large to ensure that all the distributions, $q_i[\Phi_i; f_i(\tilde{\mathbf{V}})]$, can be resolved. Other GA parameters such as the number of distinct families evolved by the GA (i.e., the population size) N_p and the mutation and crossover rates must be set to provide good exploration of Eq. (5).

The GA gene construction and population operators, such as mutation and crossover, should account for the fact that each GA individual is a full family of trial Hamiltonians. An appropriate GA genome simply concatenates the $\{\tilde{\mathbf{v}}_{p,1}, \ldots, \tilde{\mathbf{v}}_{p,N_s}\}$ Hamiltonian vectors (for the pth individual) into an $N_s \times N_v$ dimensional augmented vector. Since it is the fitness of the entire Hamiltonian family, rather than the behavior of individual Hamiltonians, that is being optimized, mutation and crossover should not mix trial Hamiltonians within a given trial family. The genetic operators should only mix corresponding Hamiltonians, i.e., $\tilde{\mathbf{v}}_{p,s}$ with $\tilde{\mathbf{v}}_{p,s'}$, $s' \neq s$. Nominal values for the mutation and crossover rates are given by the ranges, $r_m = [5\%, 20\%]$ and $r_c = [60\%, 80\%]$, respectively. A population size of $N_p \sim 50$–100 individuals, each containing a family of N_s trial potentials, is normally sufficient to provide a global optimization.

The best member of the optimized population, $\mathbf{V}^* \equiv \mathbf{V}_1^*$, is the family most consistent with the distribution of laboratory observables. However, optimization alone does not guarantee that this family represents a successful inversion, and two postoptimization tests must be performed. First, $N_t \leqslant N_s$ (an algorithmic parameter) individual potentials from the family, $\mathbf{V}^*$, are randomly chosen and their associated spectra are tested (f) for map error by comparing them with explicit solutions of Eqs. (1a) and (1b). Any optimized potential where the map error exceeds a specified tolerance ϵ_t fails the map quality test. If the number of family members that fail the testing process is greater than N_f (an algorithmic parameter), that member of the GA population is deemed faulty and abandoned. The map quality test is repeated for the second member of the GA population $\mathbf{V}_2^*$ and so on, until finding a member of the optimized family with sufficient accuracy. If none of the N_p families qualify, $f(\mathbf{v})$ is deemed faulty (g) and is regenerated over some 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 $\Phi[V]$, based on the argument that $f(\mathbf{v})$ is generally more accurate for smaller domains [6]. A large GA population N_p and family N_s size helps minimize the effect of isolated, poor individuals if the map displays high overall accuracy.

Once map testing is successful, the result is an optimized family $\mathbf{V}^*$ that hopefully produces an observable distribution consistent with that of the laboratory data. However, optimization alone will likely only achieve $q_i[\Phi_i;f_i(\mathbf{V}^*)] = q_i^{(lab)}[\Phi_i]$, if the map domain $\mathbf{V}$ contains the system's real potential. If the map domain is inadequate, the optimization will fail to minimize Eq. (5) below a specified threshold ϵ_J. If this test is not satisfied, the inversion failed and the entire map construction-optimization process is repeated by recentering the domain around the individual potential, $\mathbf{v}_s^* \in \mathbf{V}^*$, that best matches the data. Otherwise, the inversion algorithm is complete.

III. THE $X\,^1\Sigma_g^+$ AND $a\,^3\Sigma_u^+$ STATES OF Na_2

Na_2 has received recent attention in both high-precision experiments [7–9,23,24] and theoretical [2,25–30] studies. Quantitative knowledge of the Na_2 interaction is necessary for understanding its Bose-Einstein condensation behavior [31–34] and there has been considerable interest in constructing high-precision potentials for the $X\,^1\Sigma_g^+$ and $a\,^3\Sigma_u^+$ states. Toward this end, both *ab initio* and inversion techniques have been applied to Na_2. Here, we consider this well studied system in order to demonstrate the concept of inverting a laboratory data probability distribution.

1. Background of Na_2 Potentials

Rydberg-Klein-Rees (RKR) [35–37] potentials for both the $X\,^1\Sigma_g^+$ and $a\,^3\Sigma_u^+$ states have been obtained by Zemke and Stwalley [38,39]. The RKR surfaces were constructed using the $X\,^1\Sigma_g^+$ state spectroscopic constants of Babaky and Hussein [7] and measurements of several rovibrational terms by Li *et al.* [9] for the $a\,^3\Sigma_u^+$ state. Côté and Dalgarno [23] have obtained potentials for both $X\,^1\Sigma_g^+$ and $a\,^3\Sigma_u^+$ Na_2 extending the Zemke and Stwalley surfaces by incorporating the theoretical dispersion coefficients of Marinescu *et al.* [29] and an asymptotic exchange term of the form derived by Smirnov and Chibisov [40].

The potential curves of Na_2 were recently re-examined both experimentally and theoretically. Färbert and Demtröder [41] have reported the fine and hyperfine structures of the triplet states, $a\,^3\Sigma_u^+$ and $a\,^3\Sigma_g^+$, which were also theoretically investigated by Stoof *et al.* [42] and Mies *et al.* [43]. Quality *ab initio* calculations for the triplet state have been performed by Gutowski [28] at a higher level of theory than the older potentials of Konowalow [25,26]. Gutowski's calculations, which utilized fourth-order Møller-Plesset [MP2-MP4] and coupled cluster theory [CCSD, CCSD(T)] with single, double, and noniterative triple excitations, lead him to suggest that the *ab initio* potential might be more accurate than those derived from experiment [28]. Ho, Rabitz, and Scoles [2] inverted the rovibrational spectra of Babaky and Hussein [7], Barrow *et al.* [8], and Li *et al.* [9] using a reproducing kernel Hilbert space (RKHS) procedure. The RKHS results were in excellent agreement with the data and the existing RKR potentials, and helped to resolve the discrepancy between the experimental and *ab initio* potentials, but did not agree on the equilibrium bond length.

A. Na_2 inversion

The algorithm presented in Sec. II was employed to recover the $X\,^1\Sigma_g^+$ and $a\,^3\Sigma_u^+$ potentials of Na_2 from rotationally resolved laser induced fluorescence spectra. In addition to the rovibrational data, theoretical dispersion coefficients were included to ensure proper long-range behavior, and regularization was utilized to produce smooth surfaces. The inversion results and the nonlinear error analysis were compared to other proposed potentials for both states in relation to the error bounds obtained using the inversion algorithm.

In the context of rovibrational spectra, Eqs. (1a) and (1b) specifically take the form,

$$[H_0+V(r)]\psi_{v,J}=E_{v,J}\psi_{v,J}, \tag{9a}$$

$$\Phi=\{T_{v,J}\}, \tag{9b}$$

where $V(r)$ is the Born-Oppenheimer potential being inverted and the rovibrational terms, $T_{v,J}=E_{v,J}+D_e^0$, are the laboratory observables. The conventional notation for the quantum numbers has been adopted and D_e^0 is the molecular dissociation energy.

Constructing a map such that Eq. (3) converges to low order requires a proper representation of the potential-energy space. It has been demonstrated [6] that an appropriate choice of variables $\mathbf{v}$ for diatomic potentials is given by discretizing $V(r)$. The values of $V(r)$ at a collection of N_v points,

$$v_i=V(r_i), \quad i=1,\ldots,N_v \tag{10}$$

provide the potential space variables for constructing the map. The value of the potential between these points is found by interpolating a smooth surface through the v_i. Thus, the variables are the interpolation points used to construct the potential. Since N_v can be made as large as necessary to prevent interpolation errors, $f(\mathbf{v})$ can well approximate the underlying functional.

B. The $X\,^1\Sigma_g^+$ state potential

The inversion data for $X\,^1\Sigma_g^+$ Na_2 consisted of the spectral terms, $T_{v,J}$ for $0\leqslant v\leqslant 62$ (with the exception of $v=58$, which was not measured) and $J=13,15$. Accurate spectroscopic constants, $G(v)$ and $B(v)$, for the lower and intermediate, $v=0,\ldots,44$, vibrational levels of the $X\,^1\Sigma_g^+$ state have been determined by Babaky and Hussein [7] and reliable centrifugal distortion constants $D(v)$ have been determined by Ho *et al.* [2]. These constants were used to obtain the rovibrational terms,

$$T_{v,J}=G(v)+B(v)J(J+1)+D(v)[J(J+1)]^2 \tag{11}$$

for $v\leqslant 44$. Higher vibrational terms, $v=45,\ldots,57,59,\ldots 62$, have been measured by Barrow *et al.* [8] for $J=13$ and 15. The theoretical values of Marinescu *et al.* [29] for the long-range dispersion coefficients: $C_6=5.709\,413\times10^6$ cm^{-1} Å^6, $C_8=1.509\,853\times10^8$ cm^{-1} Å^8, and $C_{10}=4.181\,653\times10^9$ cm^{-1} Å^{10} were adopted for the potential beyond its LeRoy radius of 9.5 Å [44]. Com-

bining the spectral terms and the theoretical dispersion coefficients, provided the $M=129$ members of the inversion data,

$$\Phi_i^{(lab)}=T_{v,J}^{(lab)}, \quad i=1,\ldots,126, \tag{12a}$$

$$\Phi_i^{(lab)}=\{C_6,C_8,C_{10}\}, \quad i=127,128,129, \tag{12b}$$

where $v=0,\ldots,57,59,\ldots,62$, and $J=13,15$. The error distributions, $q_i^{(lab)}[\Phi_i]$, for the rovibrational terms were Gaussian with a standard deviation of $\varepsilon^{(lab)}=0.01$ cm^{-1} [8], and the dispersion coefficients were assumed error-free. The value of D_e^0 obtained by Barrow *et al.* [8], was used to compute the term values in Eq. (1b) during map construction.

The algorithm in Fig. 1 was started by providing the data members $\Phi_i^{(lab)}$, their error distributions $q_i^{(lab)}[\Phi_i]$, software for computing the rovibrational spectral quantities, and the initial map parameters. Map construction utilized the Fourier grid Hamiltonian method of Marston and Balint-Kurti [45], with 650 grid points, to solve the radial Schrödinger equation for the necessary rovibrational eigenvalues. The map domain covered a region of potential-energy space ± 500 cm^{-1} in the well and ± 5000 cm^{-1} in the repulsive wall, centered around the RKHS surface [2]. Each first-order map function was resolved at $S_i=5$ points, and the potential was discretized at 100 points, r_i, evenly spaced between $r_1=1.0$ Å and $r_{100}=16.0$ Å. Cubic-spline interpolation was used to evaluate the potential between its discretized values.

A regularization operator,

$$\hat{K}\mathbf{v}=\sum_{j=2}^{N_v-1}\left\|\frac{v_{j+1}-2v_j+v_{j-1}}{\Delta r^2}\right\|^2 \tag{13}$$

involving the three-point second derivative over the radial coordinate was utilized to ensure smooth potentials, and the regularization parameter β was set to 1×10^{-10}. Equation (5) was minimized using a steady-state GA [21] with a population of $N_p=50$ distributions, each containing $N_s=1000$ trial potentials. A mutation rate of $r_m=15\%$ and a crossover rate of $r_c=70\%$ were adopted and the testing parameters were chosen as $N_t=500$, $\epsilon_t=0.01$ cm^{-1}, $N_f=5$, and $\epsilon_J=10^{-3}$.

The map-facilitated algorithm produced the best inversion family $\mathbf{V}^*$ containing $N_s=1000$ potentials consistent with the Gaussian distributed data. Statistics for the term values produced by $\mathbf{V}^*$ are listed in Table I, where in all cases, the recovered distributions were Gaussian, and show excellent agreement with the experimental quantities. For brevity, only $v\leqslant 36$ is shown in Table I; however, the inversion reproduced the $v>36$ distributions with comparable accuracy.

Since the dispersion coefficients were treated as data, all of the potentials in the resulting inversion family were also consistent with the correct long-range behavior of the potential. Trial potentials that did not agree with the dispersion coefficients were not retained in the solution family. In this manner, it was possible to construct potentials that were consistent with data that reflected both the long- and short-range features of the interaction.

The inversion results for the $X\,^1\Sigma_g^+$ potential are depicted in Fig. 2(a) where the $\varepsilon=90\%$ confidence family $\mathbf{V}^*$ (shaded region), the RKR potential of Zemke and Stwalley (points), and the RKHS potential of Ho *et al.* (dark line) are in good agreement. Konowalow's *ab initio* potential is also plotted (dotted line) for comparison. On the scale of Fig. 2(a), including the blowup of the well region, the best estimate of the potential $\langle\mathbf{v}^*\rangle$ defined in Eq. (6), and the RKHS curve are essentially indistinguishable; however, both differ from the RKR potential. A comparison of the dissociation energy, associated with the well depth D_e and the equilibrium bond length r_e predicted by the various $X\,^1\Sigma_g^+$ state potential determinations is provided by Table II. Although there is no overlap with the older *ab initio* potential, there is remarkable agreement between the inverted surfaces. The well depth and bond length associated with $\langle\mathbf{v}^*\rangle$ are 6022.0249 cm^{-1} and 3.0791 Å, respectively.

The uncertainty in the inverted potential was measured using the metrics Δv_i^* defined in Eq. (8), for a confidence limit of $\epsilon=90\%$. Although larger in the repulsive wall, Δv_i^* was confined to $\lesssim 8\times10^{-2}$ cm^{-1} in the well and tail regions. The long-range behavior is confined by the dispersion coefficients that were accurately reproduced by all of the members of $\mathbf{V}^*$. A nonlinear relationship between the potential and spectral lines persisted even over the small area of potential-energy space described by the inversion family. The presence of nonlinear effects was confirmed by both investigating the expansion functions in Eq. (3) and by inspecting the the distribution of potentials in $\mathbf{V}^*$. Figures 2(b) and (c) depict the distribution of potentials, $p_i(v_i^*)$, and the expected value, $\langle v_i^*\rangle$, at two arbitrary points, $r_{19}=3.08$ Å and $r_{22}=3.32$ Å. The non-Gaussian nature of these distributions arises from the underlying nonlinear potential-spectrum relationship.

C. The $a\,^3\Sigma_u^+$ state potential

The $a\,^3\Sigma_u^+$ state rovibrational terms $T_{v,J}$ reported by Li *et al.* [9] for $v=0,\ldots,12$ and $J=14$ and 16, as well as the theoretical dispersion coefficients of Marinescu [29] were used to provide the inversion data set,

$$\Phi_i^{(lab)}=T_{v,J}^{(lab)}, \quad i=1,\ldots,26, \tag{14a}$$

$$\Phi_i^{(lab)}=\{C_6,C_8,C_{10}\}, \quad i=27,28,29, \tag{14b}$$

containing $M=29$ members. The error distributions $q_i^{(lab)}[\Phi_i]$ in the rovibrational terms were Gaussian with a standard deviation of $\varepsilon^{(lab)}=0.32$ cm^{-1} [9] in the rovibrational levels. The dispersion coefficients, $C_6=5.709\,413\times10^6$ cm^{-1} Å^6, $C_8=1.509\,853\times10^8$ cm^{-1} Å^8, and $C_{10}=4.181\,653\times10^9$ cm^{-1} Å^{10} were assumed error-free. Again, the value of D_e^0, obtained by Barrow *et al.* [8] was used to compute the term values in Eq. (1b) during map construction.

The algorithm in Fig. 1 was executed using similar operating parameters as for the $X\,^1\Sigma_g^+$ state inversion. The map was constructed to first, $L=1$, order using $S_i=6$ points in each dimension over a domain that spanned ± 150 cm^{-1}

TABLE I. Comparison of the distribution of rovibrational terms produced by the solution space of $X\ ^1\Sigma_g^+$ potentials depicted in Fig. 2 and the experimental values, $\Phi^{(lab)}$, from Babaky and Hussein [7] with a standard deviation (Std. Dev.) of $\varepsilon_i^{(lab)}=0.01\ \mathrm{cm}^{-1}$, for $J=13,15$ and $v=0,\ldots,36$.

	$J=13$	$q_i[\Phi_i;f_i(\mathbf{V}^*)]/\mathrm{cm}^{-1}$			$J=15$	$q_i[\Phi_i;f_i(\mathbf{V}^*)]/\mathrm{cm}^{-1}$	
v	$\Phi_i^{(lab)}/\mathrm{cm}^{-1}$	Avg.	Std. Dev.	v	$\Phi_i^{(lab)}/\mathrm{cm}^{-1}$	Avg.	Std. Dev.
0	107.4124	107.4126	0.012	0	116.3657	116.3613	0.011
1	264.9050	264.9072	0.009	1	273.8050	273.8065	0.010
2	420.9406	420.9462	0.010	2	429.7874	429.7844	0.013
3	575.5158	575.5151	0.010	3	584.3111	584.3146	0.010
4	728.6102	728.6181	0.010	4	737.3528	737.3558	0.011
5	880.2126	880.2156	0.011	5	888.9019	888.9019	0.010
6	1030.3076	1030.3092	0.010	6	1038.9426	1038.9436	0.011
7	1178.8820	1178.8827	0.014	7	1187.4614	1187.4612	0.011
8	1325.9218	1325.9258	0.010	8	1334.4445	1334.4453	0.010
9	1471.4157	1471.4173	0.012	9	1479.8816	1479.8877	0.009
10	1615.3421	1615.3438	0.012	10	1623.7496	1623.7497	0.009
11	1757.6877	1757.6827	0.010	11	1766.0361	1766.0361	0.010
12	1898.4330	1898.4378	0.011	12	1906.7213	1906.7272	0.013
13	2037.5616	2037.5626	0.010	13	2045.7891	2045.7897	0.010
14	2175.0518	2175.0572	0.009	14	2183.2174	2183.2174	0.012
15	2310.8845	2310.8852	0.010	15	2318.9876	2318.9856	0.010
16	2445.0336	2445.0333	0.011	16	2453.0725	2453.0725	0.015
17	2577.4776	2566.4762	0.010	17	2585.4511	2585.4571	0.010
18	2708.1937	2708.1937	0.011	18	2716.1012	2716.1015	0.012
19	2837.1507	2837.1536	0.010	19	2844.9899	2844.9859	0.011
20	2964.3270	2964.3270	0.012	20	2972.0974	2972.0974	0.010
21	3089.6885	3089.6888	0.012	21	3097.3877	3097.3837	0.011
22	3213.2067	3213.2087	0.011	22	3220.8331	3220.8335	0.011
23	3334.8516	3334.8513	0.010	23	3342.4040	3342.3922	0.013
24	3454.5862	3454.5842	0.010	24	3462.0623	3462.0540	0.011
25	3572.3771	3572.3723	0.011	25	3579.7751	3579.7602	0.012
26	3688.1852	3688.1822	0.010	26	3695.5029	3695.5022	0.010
27	3801.9731	3801.9741	0.011	27	3809.2086	3809.2032	0.013
28	3913.6954	3913.6884	0.010	28	3920.8460	3920.8509	0.010
29	4023.3122	4023.3182	0.010	29	4030.3761	4030.3664	0.011
30	4130.7757	4130.7757	0.012	30	4137.7507	4137.7432	0.010
31	4236.0362	4236.0334	0.012	31	4242.9193	4242.9142	0.010
32	4339.0452	4339.0435	0.010	32	4345.8342	4345.8197	0.011
33	4439.7482	4439.7482	0.011	33	4446.4402	4446.4500	0.010
34	4538.0920	4538.0834	0.010	34	4544.6847	4544.6732	0.011
35	4634.0131	4634.0131	0.011	35	4640.5025	4640.4975	0.010
36	4727.4566	4727.4564	0.011	36	4733.8403	4733.8355	0.011

around the *ab initio* potential of Gutowski [28]. 100 radial points, r_i, evenly spaced between $r_1=1.0$ Å and $r_{100}=16.0$ Å were used to resolve the $a\ ^3\Sigma_u^+$ surface and 212 grid points were utilized in the discrete Fourier-grid Hamiltonian. The same regularization operator and value of β used in the singlet state inversion were adopted.

As in the $X\ ^1\Sigma_g^+$ state inversion, optimization produced an inversion family containing $N_s=1000$ members with observable distributions consistent with the experimental results. The statistics of the term values corresponding to the inversion family are compared with the experimental values in Table III where there is excellent agreement. A comparison of the dissociation energy, associated with the well depth D_e and the equilibrium bond length r_e predicted by the various proposed surfaces is shown in Table II and the bond length and well depth for $\langle \mathbf{v}^* \rangle$ are 5.118 Å and 175.28 cm^{-1}, respectively. Although there is no direct agreement with the *ab initio* dissociation energy, there is overlap for the bond length.

The $\epsilon=90\%$ confidence limits for the $a\ ^3\Sigma_u^+$ state potential are depicted by the shaded region in Fig. 3(a). The best estimate of the potential $\langle \mathbf{v}^* \rangle$ is shown by the dotted line in Fig. 3(a) and is very similar to the RKHS potential (dark solid line). The inset plot of the well region in Fig. 3(a)

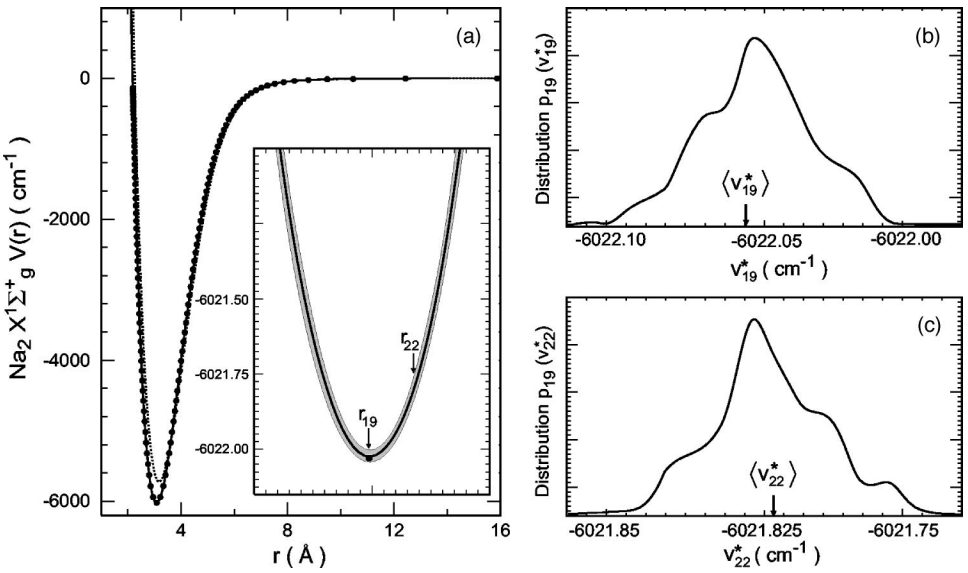

FIG. 2. (a) Global inversion results for the $X\,^1\Sigma_g^+$ state potential (shaded region of the inset) compared with the RKR results of Zemke and Stwalley [24] (points), RKHS results of Ho, Rabitz, and Scoles [2] (solid line); and the *ab initio* results of Konowalow *et al.* [25,26] (dotted line). The shaded region represents the potential-energy space subdomain in which there is a 90% probability of finding the true molecular potential, based on inverting the data of Babaky and Hussein [7] and Barrow *et al.* [8] using the algorithm in this paper. The best estimate of the potential from the present inversion is indistinguishable from the RKHS surface. The distributions of potential space variables, $q_{19}(v^*_{19})$ and $q_{22}(v^*_{22})$, including the best estimates, $\langle v_i^* \rangle$, for $r_{19}=3.08$ Å and $r_{22}=3.32$ Å are shown in (b) and (c).

compares the different proposed $a\,^3\Sigma_u^+$ state potentials. Again, **V*** (shaded region) and the RKHS potential (dark solid line) overlap; however, they do not coincide with the RKR surface (points) of Zemke and Stwalley (note that the RKR point shown is not the well minimum of that potential). The expanded region of the plot in Fig. 3(a) shows that both *ab initio* surfaces lie outside of the 90% confidence interval.

Figures 3(b) and (c) show the distribution of potentials, $p_{11}(v^*_{11})$ and $p_{13}(v^*_{13})$, at two arbitrary radial points, $r_{11}=5.1$ Å and $r_{13}=5.54$ Å. Again, the relationship between the potential and its spectral lines is nonlinear over the range of uncertainty. Except for the tail, where the dispersion coefficients confine the width of the solution family, the inversion is of lower precision than for the $X\,^1\Sigma_g^+$ state. Over most of the well, the uncertainty in the $a\,^3\Sigma_u^+$ state surface, measured by the 90% confidence interval, Δv_i^*, is ~ 1 cm^{-1}.

IV. CONCLUSION

In this paper, we introduced a laboratory data inversion procedure capable of rigorously treating experimental error. The most significant feature of this new procedure is that it inverts the full probability distribution of laboratory data. The global, map-facilitated algorithm adopted here utilized no dynamical models and employed highly accurate maps to identify the full family of potentials consistent with the statistics of the data. We also demonstrated that a linear treatment of inversion error is inadequate. This is because the nonlinear nature of the potential-observable relationship distorts the input probability distribution of the laboratory data. A Gaussian distribution of observables generally does not produce a Gaussian distribution of inverted Hamiltonians.

The capabilities of the inversion algorithm were demonstrated by extracting potentials for the $X\,^1\Sigma_g^+$ and $a\,^3\Sigma_u^+$ states of Na_2 from laboratory data. The potentials and error

TABLE II. Comparison of the well depth D_e and equilibrium bond length r_e for various $X\,^1\Sigma_g^+$ and $a\,^3\Sigma_u^+$ state potentials.

Potential	$X\,^1\Sigma_g^+$		$a\,^3\Sigma_u^+$	
	r_e (Å)	D_e (cm^{-1})	r_e (Å)	D_e (cm^{-1})
Ab initio[a]	3.124	5725.04	5.192	177.7
RKR[b]	3.0795	6022.0255	5.0911	174.45
RKHS[c]	3.0796±0.0010	6022.0246±0.0494	5.0893±0.0623	174.959±1.177
V*[d]	[3.0691,3.0799]	[6021.9981,6022.1011]	[5.017,5.201]	[174.533,175.910]

[a] $X\,^1\Sigma_g^+$ state results from Konowalow *et al.* [25,26] and $a\,^3\Sigma_u^+$ state results from Gutowski [28].
[b] $X\,^1\Sigma_g^+$ and $a\,^3\Sigma_u^+$ state results from Zemke and Stwalley [38,39].
[c] $X\,^1\Sigma_g^+$ and $a\,^3\Sigma_u^+$ state results from Ho, Rabitz, and Scoles [2].
[d] 90% confidence limits for r_e and D_e obtained from the families in Figs. 2 and 3 using 10^8 Monte Carlo sample potentials.

TABLE III. Comparison of the distribution of rovibrational terms produced by the $a\,^3\Sigma_u^+$ state inversion family in Fig. 3 and the experimental values, $\Phi_i^{(lab)}$, determined by Li *et al.* [9] with a standard deviation (Std. Dev.) of $\varepsilon_i^{(lab)}=0.32\ \mathrm{cm}^{-1}$, for $J=14,16$.

	$J=14$	$q_i[\Phi_i;f_i(\mathbf{V}^*)]$ / cm^{-1}			$J=16$	$q_i[\Phi_i;f_i(\mathbf{V}^*)]$ / cm^{-1}	
v	$\Phi_i^{(lab)}/\mathrm{cm}^{-1}$	Avg.	Std. Dev.	v	$\Phi_i^{(lab)}/\mathrm{cm}^{-1}$	Avg.	Std. Dev.
0	5871.7430	5871.7533	0.31	0	5875.2336	5875.2381	0.31
1	5894.3630	5894.3519	0.33	1	5897.5436	5897.4234	0.35
2	5914.9550	5914.9638	0.31	2	5918.0240	5917.9381	0.32
3	5934.1040	5933.9913	0.31	3	5937.0428	5936.9662	0.31
4	5951.4570	5951.4654	0.31	4	5954.3524	5954.4187	0.31
5	5966.9000	5966.8724	0.32	5	5969.5040	5969.4232	0.33
6	5980.4260	5980.4231	0.31	6	5982.8192	5982.8447	0.32
7	5991.7570	5991.7652	0.34	7	5993.8464	5993.8947	0.31
8	6001.5230	6001.5293	0.32	8	6003.7116	6003.6723	0.33
9	6009.6810	6009.6841	0.31	9	6011.4852	6011.4669	0.30
10	6015.8400	6015.7843	0.31	10	6017.4520	6017.4031	0.33
11	6020.3360	6020.2402	0.32	11	6021.6132	6021.5723	0.31
12	6022.9800	6022.8671	0.33	12	6024.0340	6023.9384	0.33

analyses performed here were compared with the recent RKHS curves of Ho *et al.*, the RKR inverted potentials of Zemke and Stwalley, and the various *ab initio* potentials of Gutowski and Konowalow *et al.*

The precision $X\,^1\Sigma_g^+$ state experimental data produced a high-quality inversion with a tight family of potentials consistent with the data. On average, the uncertainty in the extracted potential was $\lesssim 0.1\ \mathrm{cm}^{-1}$. Good agreement was found between the previously determined RKHS and RKR potentials and the full inversion family identified here. The data for the rovibrational levels of the $a\,^3\Sigma_u^+$ state were less precise, producing a family of potentials with $\sim 1\ \mathrm{cm}^{-1}$ uncertainty around the well. All members of the family exactly reproduced the theoretical dispersion coefficients beyond the LeRoy radius of 9.5 Å.

An essential point of our analysis, which was based upon

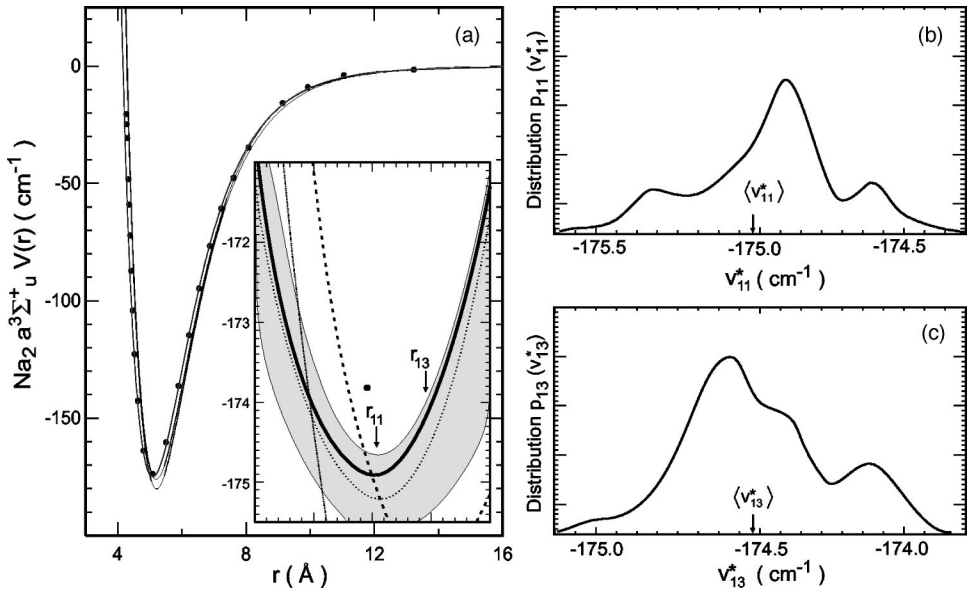

FIG. 3. (a) Global inversion results, $\mathbf{V}^*$ (shaded region), for the $a\,^3\Sigma_u^+$ state potential compared with the RKHS results of Ho, Rabitz, and Scoles [2] (dark solid line), the RKR results of Zemke and Stwalley [38] (points), and the *ab initio* results of Gutowski [28] (dark broken line), and Konowalow *et al.* [25,26] (thin line). The shaded region represents the portion of potential-energy space in which there is a 90% probability of finding the true potential, based on inverting the data of Li *et al.* [9], and the dotted line through $\mathbf{V}^*$ is the best estimate of the potential. The distribution of potentials, $q_{11}(v_{11}^*)$ (b) and $q_{13}(v_{13}^*)$ (c), within the inversion family is shown at the points $r_{11}=5.1$ Å and $r_{13}=5.54$ Å. The values, $\langle v_i^*\rangle$, in (b) and (c) show the corresponding best estimate of the potential at their associated points, r_i. The distributions reflect the presence of significant nonlinearities between the full inversion family and its consequent spectral observables and demonstrate the importance of using nonlinear error analysis.

the concept of *consistency* between the data and the family of inverted molecular potentials, is that no one feature of the surface can be used for comparison. For example, simply comparing a well depth or equilibrium bond length is a simple means for qualitatively comparing to potentials. However, in order to rigorously declare that two potentials are different, the two surfaces must produce observables that differ more than the laboratory uncertainty. Furthermore, the entire potential must be used to make such a comparison—it is insufficient to compare only several characteristic points. In this sense, the full laboratory data distribution provides the appropriate test of the potential.

Although a diatomic system was considered here, map-facilitated inversions have been performed on more complex systems such as Ar-HCl [4]. The algorithm described in Sec. II D could also be used to treat more difficult to calculate observables, such as the fine and hyperfine structures. This extension, using a map-facilitated algorithm for Na_2, would require approximately a 100–150 coupled-channel computation, which is quite practical. The power of map facilitation is that it permits sophisticated inversion techniques to be applied to computationally expensive molecular systems and observables.

ACKNOWLEDGMENTS

The authors acknowledge Dr. Tak-San Ho for beneficial conversations and for providing previously tabulated data for the RKHS and *ab initio* potentials used here for comparison. This work was supported by the DOE.

[1] R. Levine and R.B. Bernstein, *Molecular Reaction Dynamics and Chemical Reactivity* (Oxford University Press, New York, 1987).
[2] T.-S. Ho, H. Rabitz, and G. Scoles, J. Chem. Phys. **112**, 6218 (2000).
[3] J. Geremia and H. Rabitz, Phys. Rev. A **64**, 022710 (2001).
[4] J. Geremia and H. Rabitz, J. Chem. Phys. **115**, 8899 (2001).
[5] J. Geremia, E. Weiss, and H. Rabitz, Chem. Phys. **267**, 209 (2001).
[6] J. Geremia, C. Rosenthal, and H. Rabitz, J. Chem. Phys. **114**, 9325 (2001).
[7] O. Babaky and K. Hussein, Can. J. Phys. **67**, 912 (1988).
[8] R.F. Barrow, J. Verges, C. Effantin, K. Hussein, and J.D. Dincan, Chem. Phys. Lett. **104**, 179 (1984).
[9] L. Li, S.F. Rice, and R.W. Field, J. Chem. Phys. **82**, 1178 (1985).
[10] O. Alis and H. Rabitz, Trans. J. Math. Chem. **25**, 197 (1999).
[11] H. Rabitz, O. Alis, J. Shorter, and K. Shim, Comput. Phys. Commun. **11**, 117 (1999).
[12] J.I. Shorter and H. Rabitz, Geophys. Res. Lett. **27**, 3485 (2000).
[13] J. Shorter, P. Ip, and H. Rabitz, J. Phys. Chem. A **36**, 7192 (1999).
[14] H. Rabitz and K. Shim, J. Chem. Phys. **111**, 10640 (1999).
[15] K. Shim and H. Rabitz, Phys. Rev. B **58**, 1940 (1998).
[16] H. Rabitz and K. Shim, J. Chem. Phys. **111**, 10640 (1999).
[17] The notation, $\tilde{\mathbf{v}}$, represents an *individual* trial potential with corresponding map-predicted observables, $f(\tilde{\mathbf{v}})=\{f_i(\tilde{\mathbf{v}})\}$, $i=1,\ldots,M$, for M members of the data set. $\tilde{\mathbf{V}}$ represents a *family* of trial potentials, $\tilde{\mathbf{V}}\equiv\{\tilde{\mathbf{v}}_1,\ldots,\tilde{\mathbf{v}}_{N_s}\}$, and $f(\tilde{\mathbf{V}})=\{f(\tilde{\mathbf{v}}_s)\}$, $s=1,\ldots,N_s$, is the associated *family* of observables produced by the potentials in $\tilde{\mathbf{V}}$. $f_i(\tilde{\mathbf{V}})$ is the map predicted *family* of $\Phi_i^{(lab)}$, and displays a distribution, $q_i[\Phi_i;f_i(\tilde{\mathbf{V}})]$. Every $\tilde{\mathbf{v}}$ and $\tilde{\mathbf{V}}$ is a member of the map domain, $\mathbf{V}$.
[18] Note that an N_v dimensional family member, $f_i(\tilde{\mathbf{v}}_s)$, should not be confused with the first-order map functions, $f_i(v_i)$.
[19] A. Tihkonov and V. Arsenin, *Solutions of Ill-posed Problems* (Winston, Washington/Wiley, New York, 1977).
[20] K.J. Miller, Math. Anal. **1**, 52 (1970).
[21] D. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning* (Addison-Wesley, Reading, MA, 1989).
[22] $f_i(\mathbf{v}^*)$ is the map-predicted value of Φ_i for a single optimized potential $\mathbf{v}^*$, and $f_i(\mathbf{V}^*)$ is the map-predicted family of observables with a different value of Φ_i for each $\mathbf{v}_s^*$ in $\mathbf{V}^*$. $q_i[\Phi_i;f_i(\mathbf{V}^*)]$ is the distribution of observables produced by the optimal family, $\mathbf{V}^*$.
[23] R. Côté and A. Dalgarno, Phys. Rev. A **50**, 4827 (1994).
[24] K.M. Jones, S. Maleki, S. Size, P.D. Lett, C.J. Williams, H. Richling, H. Knockel, E. Tiemann, P.L.G.H. Wang, and W.C. Stwalley, Phys. Rev. A **54**, R1006 (1996).
[25] D.D. Konowalow, M.E. Rosenkrantz, and M.L. Olson, J. Chem. Phys. **72**, 2612 (1980).
[26] D.D. Konowalow and M.E. Rosenkrantz, J. Phys. Chem. **86**, 1099 (1982).
[27] S. Magnier, P. Millie, O. Dulieu, and F. Masnou-Seeuws, J. Phys. Chem. **98**, 7113 (1993).
[28] M. Gutowski, J. Phys. Chem. **110**, 4695 (1999).
[29] M. Marinescu, H.R. Sadeghpour, and A. Dalgarno, Phys. Rev. A **49**, 982 (1994).
[30] G. Hadinger, S. Magnier, and M. Aubert-Frecon, J. Mol. Spectrosc. **175**, 441 (1996).
[31] M.H. Anderson, J.R. Ensher, M.R. Matthews, C.E. Wieman, and E.A. Cornell, Science **269**, 198 (1995).
[32] C.C. Bradley, C.A. Sackett, J.J. Tollett, and R.G. Hulet, Phys. Rev. Lett. **75**, 1687 (1995).
[33] C.C. Bradley, C.A. Sackett, and R.G. Hulet, Phys. Rev. Lett. **78**, 985 (1997).
[34] K.B. Davis, M.-O. Mrwes, M.R. Andrews, N.J. van Druten, D.S. Durfee, D.M. Kurn, and W. Ketterle, Phys. Rev. Lett. **75**, 3969 (1995).
[35] R. Rydberg, Z. Phys. **73**, 376 (1931).
[36] O. Klein, Z. Phys. **76**, 226 (1932).
[37] A.L.G. Rees, Proc. Phys. Soc. London **59**, 998 (1947).
[38] W.T. Zemke and W.C. Stwalley, J. Chem. Phys. **100**, 2661 (1994).
[39] W.T. Zemke and W.C. Stwalley, J. Chem. Phys. **111**, 4962 (1999).

[40] B.M. Smirnov and M.I. Chibisov, Sov. Phys. JETP **21**, 624 (1965).
[41] A. Färbert and W. Demtröder, Chem. Phys. Lett. **264**, 255 (1997).
[42] H. Stoof, J. Koelman, and B.J. Verhaar, Phys. Rev. B **38**, 4688 (1988).
[43] F.H. Mies, E. Tiesinga, and P. Julienne, Phys. Rev. A **61**, 022721 (2000).
[44] R.J. LeRoy, *Specialist Periodical Reports, Molecular Spectroscopy* (Chemical Society, London, 1973), Vol. 1.
[45] C.C. Marston and G. Balint-Kurti, J. Chem. Phys. **91**, 3571 (1989).