REPORT DOCUMENTATION PAGE

Form Approved OMB NO. 0704-0188

The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggesstions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA, 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any oenalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.
PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From - To)
26-06-2008	Final Report	15-Jul-2004 - 14-Jan-2008

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Closed loop adaptive refinement of dynamical models for complex chemical reactions	W911NF-04-1-0341
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER 611102
6. AUTHORS Herschel Rabitz	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAMES AND ADDRESSES	8. PERFORMING ORGANIZATION REPORT NUMBER
Princeton University Office Of Research & Project Administration The Trustees of Princeton University Princeton, NJ 08544 -0036	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S) ARO
U.S. Army Research Office P.O. Box 12211 Research Triangle Park, NC 27709-2211	11. SPONSOR/MONITOR'S REPORT NUMBER(S) 47038-CH-DRI.1

12. DISTRIBUTION AVAILIBILITY STATEMENT

Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

The views, opinions and/or findings contained in this report are those of the author(s) and should not contrued as an official Department of the Army position, policy or decision, unless so designated by other documentation.

14. ABSTRACT

This research is concerned with the development of a systematic method for efficiently performing molecular dynamics (MD) simulations of complex chemical reactions and optimizing the underlying potential energy surfaces (PESs), ultimately using suitable laboratory data in a closed loop fashion. Two main objectives of the research are to (a) identify key parameters of each PES based on the global non-linear input-output Random Sampling High Dimensional Model Representation (RS-HDMR) mapping technique [1-7] and (b) use the RS-HDMR maps to efficiently capture the PES observable relationships [8-10]. The RS-HDMR analysis in turn provides essential information for subsequent full implementation of PES optimization

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	15. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Herschel Rabitz
U	U	U	SAR		19b. TELEPHONE NUMBER 609-258-3917

Standard Form 298 (Rev 8/98)
Prescribed by ANSI Std. Z39.18

Report Title

Closed loop adaptive refinement of dynamical models for complex chemical reactions

ABSTRACT

This research is concerned with the development of a systematic method for efficiently performing molecular dynamics (MD) simulations of complex chemical reactions and optimizing the underlying potential energy surfaces (PESs), ultimately using suitable laboratory data in a closed loop fashion. Two main objectives of the research are to (a) identify key parameters of each PES based on the global non-linear input-output Random Sampling High Dimensional Model Representation (RS-HDMR) mapping technique [1-7] and (b) use the RS-HDMR maps to efficiently capture the PES observable relationships [8-10]. The RS-HDMR analysis in turn provides essential information for subsequent full implementation of PES optimization within the proposed adaptive closed-loop learning algorithm in conjunction with laboratory feedback. In this project we have (1) formulated a fully equivalent operational model (FEOM) based on RS-HDMR, in place of the time-consuming Newton equations of motion for performing multi-dimensional MD simulations, and (2)performed detailed studies on intermolecular energy transfer for the model systems.

List of papers submitted or published that acknowledge ARO support during this reporting period. List the papers, including journal references, in the following categories:

(a) Papers published in peer-reviewed journals (N/A for none)

Extraction of parameters and their error distributions from cyclic voltammograms using bootstrap re-sampling enhanced by solution maps: a computational study, L. Bieniasz, H. Rabitz, Anal. Chem., 78, 8430 (2006).

Number of Papers published in peer-reviewed journals: 1.00

(b) Papers published in non-peer-reviewed journals or in conference proceedings (N/A for none)

Number of Papers published in non peer-reviewed journals: 0.00

(c) Presentations

Number of Presentations: 0.00

Non Peer-Reviewed Conference Proceeding publications (other than abstracts):

Number of Non Peer-Reviewed Conference Proceeding publications (other than abstracts): 0

Peer-Reviewed Conference Proceeding publications (other than abstracts):

Number of Peer-Reviewed Conference Proceeding publications (other than abstracts):

(d) Manuscripts

Global sensitivity analysis by random sampling – high dimensional model representation (RS-HDMR), G. Li, H. Rabitz, P. Yelvington, O. Oluwole, F. Bacon, C. Kolb, Z. Chen, Y. Ju, S.-W. Wang, P. Georgopoulos and J. Schoendorf, Technometrics.

Random sampling-high dimensional model representation (RS-HDMR) with application on molecular dynamics simulations, Li, G., Ho, T. -S., Rabitz, H. (2008)manuscript in preparation.

Number of Manuscripts: 2.00

Number of Inventions:

Graduate Students

NAME	PERCENT_SUPPORTED
FTE Equivalent:	
Total Number:	

Names of Post Doctorates

NAME	PERCENT_SUPPORTED
xiao-jiang Feng	
FTE Equivalent:	
Total Number:	1

Names of Faculty Supported

NAME	PERCENT_SUPPORTED	National Academy Member
Herschel Rabitz		No
FTE Equivalent:		
Total Number:	1	

Names of Under Graduate students supported

NAME	PERCENT_SUPPORTED
FTE Equivalent:	
Total Number:	

Student Metrics

This section only applies to graduating undergraduates supported by this agreement in this reporting period

The number of undergraduates funded by this agreement who graduated during this period: .....

The number of undergraduates funded by this agreement who graduated during this period with a degree in science, mathematics, engineering, or technology fields:.....

The number of undergraduates funded by your agreement who graduated during this period and will continue to pursue a graduate or Ph.D. degree in science, mathematics, engineering, or technology fields:.....

Number of graduating undergraduates who achieved a 3.5 GPA to 4.0 (4.0 max scale):.....

Number of graduating undergraduates funded by a DoD funded Center of Excellence grant for Education, Research and Engineering:.....

The number of undergraduates funded by your agreement who graduated during this period and intend to work for the Department of Defense .....

The number of undergraduates funded by your agreement who graduated during this period and will receive scholarships or fellowships for further studies in science, mathematics, engineering or technology fields:.....

Names of Personnel receiving masters degrees

NAME
Total Number:

Names of personnel receiving PHDs

NAME	
xiao-jiang Feng	
Total Number:	**1**

Names of other research staff

NAME	PERCENT_SUPPORTED
FTE Equivalent:	
Total Number:	

Sub Contractors (DD882)

Inventions (DD882)

Closed-Loop Adaptive Refinement of Dynamical Models for Complex Chemical Reactions: Application of RS-HDMR to Molecular Dynamics Simulations

Herschel Rabitz, Princeton University
Frick Laboratory, Princeton, NJ 08540, hrabitz@princeton.edu

(Final report 2005 – 2008)

1 Research Objectives

This research is concerned with the development of a systematic method for efficiently performing molecular dynamics (MD) simulations of complex chemical reactions and optimizing the underlying potential energy surfaces (PESs), ultimately using suitable laboratory data in a closed loop fashion. Two main objectives of the research are to (a) identify key parameters of each PES based on the global non-linear input-output Random Sampling-High Dimensional Model Representation (RS-HDMR) mapping technique [1-7] and (b) use the RS-HDMR maps to efficiently capture the PES → observable relationships [8-10]. The RS-HDMR analysis in turn provides essential information for subsequent full implementation of PES optimization within the proposed adaptive closed-loop learning algorithm in conjunction with laboratory feedback. In this project we have (1) formulated a fully equivalent operational model (FEOM) based on RS-HDMR, in place of the time-consuming Newton equations of motion for performing multi-dimensional MD simulations, and (2) performed detailed studies on intermolecular energy transfer for the model systems Cl + CH_3Br [11-19], Ar +H_2O [20-22], and Ar + CH_4 [20-25]. They are summarized below, along with a plan for future research. As HDMR is of broad applicability in chemistry, we have also carried out a parallel application of the same principle in an electrochemical study. In the latter case the parameters are rate constants, charge transfer coefficients and reference potentials, and the HDMR maps were for cyclic voltammograms. The test shows how the HDMR maps could be used in an adaptive fashion to refine the model parameters in analogy with the chemical dynamics goal.

2 Formulation of full equivalent operational models (FEOM) for MD simulations

MD simulations necessitate solving Newton's equations for all atomic positions and velocities, thus entailing a large number of trajectories calculated with different initial conditions. Averaging the results of all the trajectories over the initial conditions produces the desired MD simulations. The computational effort of MD simulations is often formidable. Instead of integrating a large number of Newton equations of motion, the RS-HDMR input-output mappings make possible simple algebraic evaluations from a modest sample of molecular dynamics simulation data (outputs) given the underlying PES parameters (inputs). To this end, we have found that by treating both the PES parameters and initial conditions as inputs in the classical trajectory calculations, effective RS-HDMR input-output maps may

be constructed using only PES parameters, while neglecting the differences in the initial conditions. The general machinery of HDMR consists of a set of quantitative model assessment and analysis tools for capturing high dimensional input-output system behavior. In particular, the RS-HDMR serves as a powerful and practical technique of the HDMR methods incorporating efficient random sampling procedures [4-7]. In general, the HDMR of any n-variate function $f(\mathbf{x})$ is construed as a finite hierarchical correlated function expansion of the input variables $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ [1,2]:

$$\begin{aligned} f(\mathbf{x}) &= f_0 + \sum_{i=1}^{n} f_i(x_i) + \sum_{i<j=1}^{n} f_{ij}(x_i, x_j) + \ldots \\ &\quad + \ldots + f_{12\ldots n}(x_1, x_2, \ldots, x_n), \end{aligned} \tag{1}$$

where the component functions f_0, $f_i(x_i)$, $f_{ij}(x_i, x_j)$,... can be systematically optimized.

During the MD simulations, for a given set of inputs $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, such as the PES parameters and initial condition parameters of the reactants, and any output function $f(\mathbf{x})$, such as the internal energy E_{int}, vibrational energy E_{vib}, rotational energy E_{rot}, bond length, or bond angle of the products, the corresponding RS-HDMR component functions, cf. eq. (**??**), can be constructed from a set (often a few thousand) solutions of Newton equations of motion, with randomly sampled initial conditions. The resultant RS-HDMR formula can in turn be used to interpolate $f(\mathbf{x})$ for any given initial condition without further solving Newton equations. Accordingly, a set of random data with simultaneous variation of the PES and initial condition parameters can be generated. Finally, the relevant RS-HDMR formulas are constructed only for the PES parameters. We have proved mathematically that freezing the initial conditions to construct partial RS-HDMR maps effectively averages over the random initial conditions in the limit that the sample size is sufficiently large. As a result, given a set of PES parameters within their bounds of variation, the resultant input-output RS-HDMR maps can accurately predict the MD simulation outputs in accordance with (i.e., the typical RS-HDMR prediction errors were $\sim 1\%$) realistic laboratory situations. The expected computational savings can be crucial to the subsequent implementation with laboratory feedback.

The resultant RS-HDMR formula can be used not only as a FEOM to reduce the computational effort, but also as a tool to perform sensitivity analysis and identify the important PES parameters, i.e., the important region of the potential energy surface. Eq. (**??**) may be written in an abbreviated form

$$f(\mathbf{x}) = f_0 + \sum_{l=1}^{n_p} g_l, \tag{2}$$

where g_l denotes any component function in eq. (**??**) and n_p is the total number of significant component functions used for approximation. The individual RS-HDMR component functions have a direct statistical correlation interpretation. This relation permits the total variance σ^2 of the output $f(\mathbf{x})$ to be decomposed into the independent and correlated contributions of the inputs [7]:

$$\sigma^2 \approx \sum_{l=1}^{n_p} \text{Cov}(f(\mathbf{x}), g_l(\mathbf{x})), \tag{3}$$

where $\mathrm{Cov}(\cdot,\cdot)$ denotes the covariance. The covariance $\mathrm{Cov}(f(\mathbf{x}), g_l(\mathbf{x}))$ is composed of two terms

$$\mathrm{Cov}(f(\mathbf{x}), g_l(\mathbf{x})) = (g_l(\mathbf{x}), g_l(\mathbf{x})) + (\sum_{\substack{k=1\\k\neq l}}^{n_p} g_k(\mathbf{x}), g_l(\mathbf{x})), \tag{4}$$

where $(g_k(\mathbf{x}), g_l(\mathbf{x}))$ denotes the inner product of $g_k(\mathbf{x})$ and $g_l(\mathbf{x})$). The first term in eq. (**??**) measures the deterministic contributions of g_l, and the second term in eq. (**??**) measures the stochastic contribution of g_l with other component functions caused by a correlated probability density function (pdf). Both provide useful information of inputs and their sampling relationship.

The sensitivity indexes $S_l, S_l^a, S_l^b (l = 1, 2, \ldots, n_p)$ are defined as

$$S_l = \mathrm{Cov}(g_l(\mathbf{x}), f(\mathbf{x}))/\sigma^2, \tag{5}$$

$$S_l^a = (g_l(\mathbf{x}), g_l(\mathbf{x}))/\sigma^2, \tag{6}$$

$$S_l^b = S_l - S_l^a. \tag{7}$$

If eq. (**??**) includes all significant component functions g_l, we have

$$\sum_{l=1}^{n_p} S_l = \sum_{l=1}^{n_p} \mathrm{Cov}(g_l(\mathbf{x}), f(\mathbf{x}))/\sigma^2 \approx 1. \tag{8}$$

The magnitudes of S_l, S_l^a, S_l^b can be used to quantitatively determine the importance order of inputs, pairs of inputs, etc. For systems with independent input sampling, $S_l^b = 0$ and we may simply use S_l. For systems with a correlated pdf S_l^a may be used to identify their importance, and S_l^b is used to identify input sampling correlations. The sum of all S_l may be used as a measure of the analysis quality. If $\sum_l S_l \approx 1$, the sensitivity analysis is complete and reliable.

3 RS-HDMR MD Simulation Results

The RS-HDMR component functions have been constructed for intermolecular energy transfer in Cl + CH_3Br [11-19], Ar + H_2O [20-22], and Ar + CH_4 [20-25] collisions, each using a large set classical trajectory calculations with randomly chosen initial conditions. The resultant RS-HDMR maps (i.e., serving as FEOM's) were used to interpolate the output data as functions of the PES parameters for given initial conditions without explicitly executing further solutions of the underlying Newton equations. Similar results were obtained for Ar + H_2O and Ar + CH_4. The results for Cl^- + CH_3Br, Ar + H_2O are given below.

3.1 Cl^- + CH_3Br Reaction

A global RS-HDMR input/output map for the S_N2 reaction

$$Cl^- + CH_3Br \rightarrow ClCH_3 + Br^- \tag{9}$$

was constructed. Here the input is the ϕ-angle bending potential function V_ϕ^{MC} for CH_3Cl molecule [11],

$$V_\phi^{\mathrm{MC}} = f_\phi^{\mathrm{MC}} \sum_{i=1}^{3} (\phi_i - \phi_{\mathrm{MC}})^2/2 + a_\phi^{\mathrm{MC}} \sum_{i=1}^{3} (\phi_i - \phi_{\mathrm{MC}})^3, \tag{10}$$

where $\phi_{\mathrm{MC}} = 108.4528°$, and the output is the molecular bond lengths for CCl, CBr, CH and bound angles for ClCBr, ClCH, HCH. Specifically, the force constants f_ϕ^{MC} and a_ϕ^{MC} were variated about 10% around the nominal values 0.8424 and -0.230506. One thousand data for these two parameters were randomly generated and their reaction trajectories were calculated using the code VENUS. The output data were recorded after the reaction has taken place. A global RS-HDMR map was constructed using $f_\phi^{\mathrm{MC}}, a_\phi^{\mathrm{MC}}$ and V_ϕ^{MC} chosen as three inputs represented as x_1, x_2, x_3. The CCl, CBr, CH bond lengths and ClCBr, ClCH, HCH bond angles were chosen as the outputs.

3.1.1 Prediction accuracy of RS-HDMR

The 3rd order RS-HDMRs, whose component functions were approximated by the 3rd order optimal orthonormal polynomials, were constructed for 11 outputs (CCl, CBr, three CH bond lengths, ClCBr, three ClCH, two HCH bond angles) at different fixed reaction time. The comparison between the model values and different order RS-HDMR approximations at $t = 270 \times 10^{-14}$ second (fs) is given in Table 1. The comparison between the model values and the 3rd order RS-HDMR approximations for CCl bond length at $t = 270$ fs is given in Fig. 1.

Since only three input variables were used, the input - output relation is quite simple, and the accuracy of different order RS-HDMR is satisfactory. The largest relative error is only 0.612%.

Table 1. The average and maximum relative errors of the 1st through 3rd order RS-HDMR for CCl, CBr bond lengths and ClCBr bond angle constructed from the 1000 random data at $t = 270$ fs

Output	1st order		2nd order		3rd order	
	Average	Maximum	Average	Maximum	Average	Maximum
CCl	0.00015	0.00053	0.00005	0.00020	0.00004	0.00021
CBr	0.00170	0.00092	0.00559	0.00559	0.00072	0.00484
ClCBr	0.00250	0.00612	0.00093	0.00375	0.00083	0.00354

3.1.2 Global Sensitivity Analysis

The global sensitivity analysis given by the 2nd order RS-HDMR was performed for the same data set. The results for S_i are given in Table 2, which shows that the input x_1 (f_ϕ^{MC}) is dominant. Since $\sum S_i > 0.95$ for three outputs, all significant input contributions have been counted. It also shows that the S_{ij}'s are small.

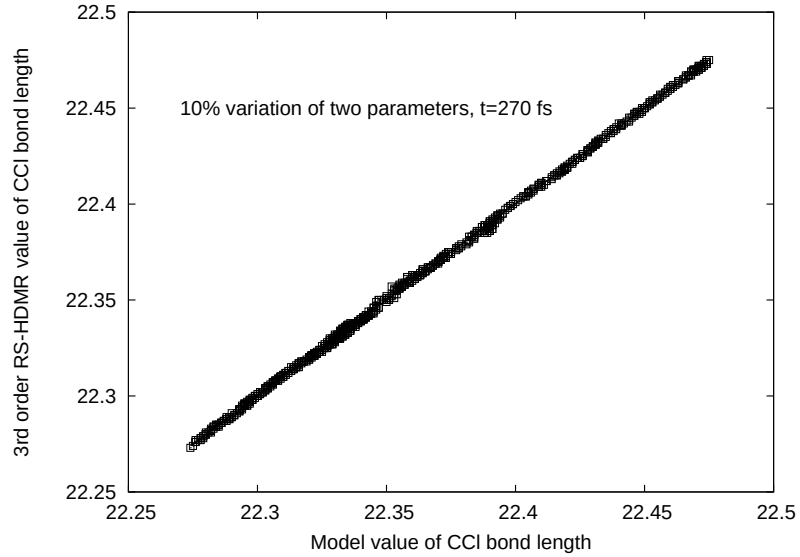

Figure 1: The truth plot of the 3rd order RS-HDMR approximation for CCl bond length.

Table 2. The resultant S_i

	CCl	CBr	ClCBr
S_1	0.9913	0.9409	0.9930
S_2	0.0019	0.0086	0.0008
S_3	0.0013	0.0048	0.0013
$\sum S_i$	0.9945	0.9543	0.9951

3.2 Ar + H_2O System

The potential energy surface of Ar + H_2O contains two parts: an intermolecular interaction V_{inter} (between Ar and H_2O) and an intramolecular interaction V_{intra} (for H_2O). For illustration, the intermolecular interaction is a combination of Lennard-Jones (12,6) diatomic potential functions describing both short- and long-range van der Waals interactions between Ar and constituent atoms of H_2O, while the intramolecular interaction is a sum of stretches (Morse oscillators) of bond-lengths and bending (Harmonic oscillators) of bond-angles. As a result, Ar + H_2O potential energy surface can be considered as a function of nine potential parameters, four for the intermolecular part V_{inter} and five for the intramolecular part V_{intra}. In our simulations, the potential energy function (PES) of the Ar + H_2O is modeled as [20]

$$V = V_{\text{inter}} + V_{\text{intra}}, \tag{11}$$

where the intermolecular interaction V_{inter} between Ar and H_2O is a sum of the generalized two-body Lennard-Jones terms

$$V_{\text{inter}} = \sum_{i=1}^{2} V_{\text{Ar}-\text{H}^i} + V_{\text{Ar}-\text{O}},$$

$$
\begin{aligned}
V_{\mathrm{Ar-X}^i} &= \frac{a_i}{R_i^{12}} + \frac{b_i}{R_i^6}, \\
\mathrm{X}^i &= \mathrm{H}^i,\ i = 1, 2,\ \mathrm{X}^3 = \mathrm{O}, \\
R_i &= R(\mathrm{Ar} - \mathrm{H}^i),\ i = 1, 2,\ R_3 = R(\mathrm{Ar} - \mathrm{O}),
\end{aligned} \tag{12}
$$

and the intramolecular interaction V_{intra} of H_2O is a sum of stretches (Morse oscillators) and bends (Harmonic oscillators)

$$
\begin{aligned}
V_{\mathrm{intra}} &= D_r \times \sum_{i=1}^{2} \Big[1 - \exp\Big(-\beta_r \times (r_i - r_0)\Big)\Big]^2 + \frac{1}{2} \times f_\theta \times (\theta - \theta_0)^2, \\
& \qquad r_i = r(\mathrm{O} - \mathrm{H}^i),\ i = 1, 2 \\
& \qquad \theta = \theta(\angle \mathrm{H}^1\mathrm{OH}^2).
\end{aligned} \tag{13}
$$

Here the nine nominal values of the PES parameters are $a_1 = a_2 = 178047.57$ kcalÅ^{12}/mol, $a_3 = 368860.80$ kcalÅ^{12}/mol, $b_1 = b_2 = -203.98076$ kcalÅ^6/mol, $b_3 = -453.03085$ kcalÅ^6/mol, $D_r = 125.6$ kcal/mol, $\beta_r = 2.19367$ Å^{-1} , $r_0 = 0.9572$ Å, $f_\theta = 0.688$ mdyn-Å/rad^2, and $\theta_0 = 104.52^o$

For simplicity, initially the molecules H_2O are assumed to be in an excited vibrational state, say $E_{\mathrm{vib}} = 25.0$ kcal/mol for H_2O and possibly with a rotational energy. Classical trajectories are calculated for an initial relative translational energy, say $E_{\mathrm{rel}} = 100$ kcal/mol (in general, E_{rel} can also be considered as a RS-HDMR input, thus subject to random samples) at arbitrarily sampled impact parameters b and orientations (of H_2O with respect to the impact direction). The output data of interest from the trajectory calculations are changes in (1) the relative translational energy ΔE_{rel}, (2) the H_2O rotational energy ΔE_{rot}, and (3) the H_2O vibrational energy ΔE_{vib}. The usually strong Coriolis coupling, for example, between H_2O rotational and vibrational modes will make the resulting rotational and vibrational energy fluctuate in time. As a result, a time-averaged rotational energy (thus a time-averaged vibrational energy) is usually considered after the collision. On the other hand, the corresponding internal energy $E_{\mathrm{int}} = E_{\mathrm{vib}} + E_{\mathrm{rot}}$ will remain constant after the collision.

3.2.1 Prediction accuracy of RS-HDMR

Seven (due to $a_1 = a_2$ and $b_1 = b_2$) potential parameters given above and four initial condition parameters were used as inputs. The potential parameters vary within 10% of their nominal values. E_{int}, E_{vib} and E_{rot} are used as outputs. 25000 random points of all inputs were generated and their corresponding outputs were calculated by Newton's equations of motion. The 3rd order RS-HDMR approximations with only potential parameters as inputs were constructed from the data. Ten sets of testing data were generated. Each set corresponds to a fixed potential with 10000 different initial conditions. The 1st to 3rd order RS-HDMR expansions were used to test the averaged values of E_{int} obtained from the ten sets of data. The results are given in Table 3.

Table 3. The relative error of prediction for the averaged value of E_{int}

Data set	RS-HDMR approximation		
	1st order	2nd order	3rd order
1	0.0038	0.0087	0.0086
2	0.0061	0.0030	0.0006
3	-0.0003	0.0008	-0.0027
4	0.0052	0.0042	0.0060
5	0.0044	0.0024	0.0034
6	0.0024	0.0010	0.0016
7	0.0023	0.0034	0.0066
8	-0.0002	0.0001	0.0006
9	0.0048	0.0048	0.0040
10	0.0009	0.0040	0.0047

The prediction error for averaged values of E_{int} is less than 1%, which is satisfactory. Similar results were obtained for E_{vib}. The error for E_{rot} is larger, but considering that the magnitude of E_{rot} is less than 1, the prediction accuracy for E_{rot} is acceptable.

3.2.2 Global Sensitivity Analysis

In our research, the sensitivity indexes can be used to identify the important regions of potential energy surface. For this purpose *only the seven potential parameters* for Ar + H_2O system were used as inputs. E_{int}, E_{vib} and E_{rot} were used as outputs. 25000 random points were generated. Different sample sizes (denoted as "used data") were used to construct the RS-HDMR component functions. The remained data (denoted as "test data") were used for testing. The results for E_{int} are given below. Figures 2 and 3 give the truth plot of the 3rd order RS-HDMR approximation, and Table 4 gives the averaged relative error of different order RS-HDMR approximations for the used and test data. Both are only a few percent.

Table 4. The average relative error of different order RS-HDMR constructed from 3000 points for E_{int}

Sample size	Used data			Test data		
	1st order	2nd order	3rd order	1st order	2nd order	3rd order
3000	0.0681	0.0545	0.0438	0.0676	0.0589	0.0552

The sensitivity indexes were calculated from the resultant RS-HDMR approximation. The sensitivity indexes are given in Tables 5-7.

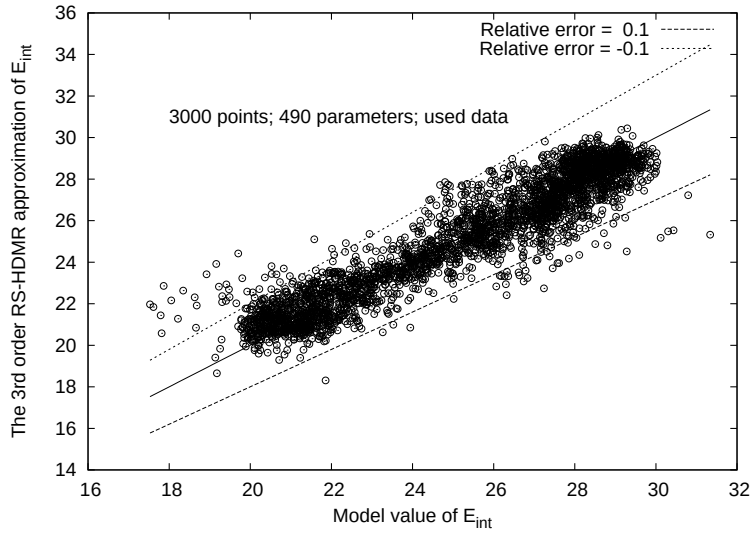

Figure 2: The truth plot for used data with E_{int}.

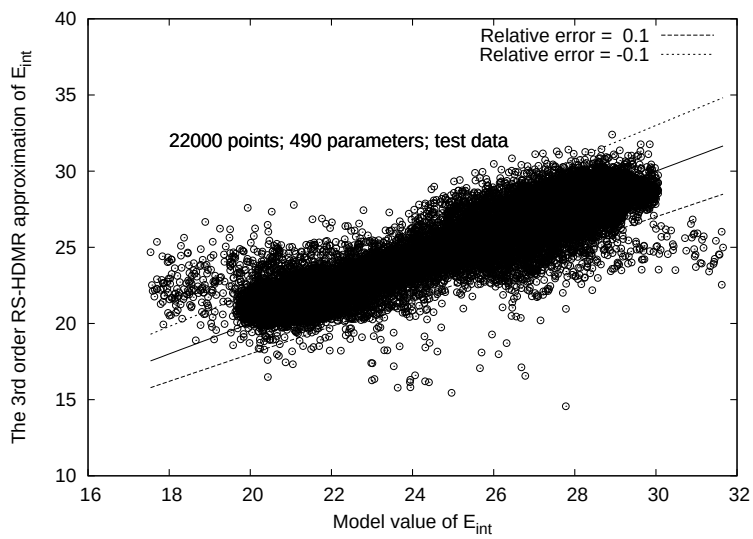

Figure 3: The truth plot for test data with E_{int}.

Table 5. The 1st order sensitivity indexes for E_{int}

Potential parameter	Import. Order	Sensitivity index		
		S_i^a	S_i^b	S_i
D_r	1	0.7036	0.0010	0.7046
a_3	2	0.0042	0.0009	0.0051
b_1	3	0.0011	0.0001	0.0012
$\sum S_i$				0.7108

Table 6. The 2nd order sensitivity indexes for E_{int}

Potential parameter	Import. Order	Sensitivity index		
		S^a_{ij}	S^b_{ij}	S_{ij}
(a_3, D_r)	1	0.0551	-0.0019	0.0532
(a_3, b_1)	2	0.0255	-0.0013	0.0242
(b_1, D_r)	3	0.0118	0.0007	0.0124
$\sum S_i + \sum S_{ij}$				0.8006

Table 7. The 3rd order sensitivity indexes for E_{int}

Potential parameter	Sensitivity index		
	S^a_{ijk}	S^b_{ijk}	S_{ijk}
(a_3, b_1, D_r)	0.0591	0.0036	0.0628
$\sum S_i + \sum S_{ij} + S_{235}$			0.8634

The sensitivity indexes in Tables 5-7 show that the most important potential parameters are D_r, then a_3, $b_1(b_2)$. The magnitudes of the sensitivity indexes quantitatively identify the importance of different potential regions. As the potential parameters are sampled independently, $S^b_i, S^b_{ij}, S^b_{ijk}$ should be zero. Their small values indicate reliable RS-HDMR mapping. Therefore, there is no significant difference for using $S^a_i, S^a_{ij}, S^a_{ijk}$ or S_i, S_{ij}, S_{ijk}. The total sum of sensitivity indexes is close to unity at 0.8635, which implies that the sensitivity analysis is meaningful. However, the contribution of other inputs ($\sim 14\%$) could not be identified. The reason is that the large contribution of D_r, a_3 and b_1 covered their small contribution. This circumstance needs further analysis.

4 RS-HDMR-based Inversion of cyclic voltammogram data

The molecular dynamics studies have the goal of creating highly accurate and efficient HDMR maps to facilitate laboratory data inversions for refinement of dynamical models. This task is generic with parallel applications arising in many areas of chemistry. While the particular aspects of the dynamical RS-HDMR were being developed (see Sections I and II above), we took the opportunity to test the basic RS-HDMR algorithm data inversion capabilities in an analogous application in complex electrochemistry. In this case the potential parameters of dynamics now become kinetic rate constants, charge transfer coefficients and reference potentials. The dynamical outputs now become the observed current over time. The resultant HDMR's were of very high accuracy and a full simulation was performed for inverting voltammographic data, including a statistical analysis of the identified parameters [26]. We expect that analogous behavior will occur in the case of molecular dynamics data inversions.

5 Conclusion and Future Work

For the current research work, the RS-HDMR component functions were approximated by a third-order polynomial expansion to reduce the computational cost. However, the present

study shows that the distinctions in the averaged output values for sampled PES parameters are often too small to fully resolve these differences using the third-order RS-HDMR approximation. The shortcoming of any polynomial expansion scheme is that using high-order polynomials may produce undesired highly oscillatory behavior, particularly on the sampling boundaries. The oscillatory behavior and the large swings in the value of the high-order polynomials are detrimental for the resulting input-output maps. But simply adding more input variables in general does not necessarily improve the accuracy of the corresponding RS-HDMR input-output map, because the map function is usually too complicated to be fitted accurately by low degree polynomials. Moreover, we found that some PES parameters, which weigh heavily in the RS-HDMR input-output maps, can obscure the roles played by other parameters. To overcome these technical difficulties, we plan to adopt a non-parametric RS-HDMR scheme such that the underlying component functions can be represented accurately by numerical data tables at chosen grid points. From the tabulated data, the component function values can then be calculated using the Interpolating Moving Least-Squares (IMLS) method. Other new HDMR tools also need to be formulated to address the difficulties with the existing procedures. The new tools should enable future studies using larger sampling ranges of PES parameters for MD. As a final step, we would then be able to fully implement the closed-loop adaptive procedure for the refinement of various chemical reaction dynamical models [27].

References

[1] Alis, O. F., Rabitz, H. (1999). General foundations of high dimensional model representations, *J. Math. Chem.,* 25, 197-233.

[2] Rabitz, H., Alis, O. F., Shorter, J. and Shim K. (1999). "Efficient input-output model representations, *Comput. Phys. Comm.*, 117, 11-20.

[3] Alis, O. F., Rabitz, H. (2001). Efficient implementation of high dimensional model representations, *J. Math. Chem.,* 29, 127-142.

[4] Li, G., Rosenthal, C. and Rabitz, H. (2001). High dimensional model representations, *J. Phys. Chem. A,* 105, 7765-7777.

[5] Li, G., Wang, S. W. and Rabitz, H. (2002), Practical approaches to construct RS-HDMR component functions, *J. Phys. Chem. A,* 106, 8721-8733.

[6] Li, G., Hu, J. S., Wang, S. W., Georgopoulos, P.G., Schoendorf, J., Rabitz, H. (2006). Random sampling-high dimensional model representation (RS-HDMR) and orthogonality of its different order component functions, *J. Phys. Chem. A,* 110, 2474-2485.

[7] * Li, G., Rabitz, H., Yelvington, P. E., Oluwole, O. O., Bacon, F., Kolb, C. E., Chen, Z., Ju, Y., Wang, S. W., Georgopoulos, P. G., Schoendorf, J. (2008). Global sensitivity analysis by random sampling- high dimensional model representation, *Technometrics*, submitted.

[8] Geremia, J. M., Rabitz, H. and Rosenthal, C. (2001). Constructing global functional maps between molecular potentials and quantum observables, *J. Chem. Phys.*, 114, 9325-9336..

[9] Geremia, J. M. and Rabitz, H. (2001). The Ar-HCL energy surface from a global map-facilitated inversion of state-to-state rotationally resolved differential scattering cross section and rovibrational spectral data, *J. Chem. Phys.*, 115, 8899-8912.

[10] Ho, T.-S. and H. Rabitz, (2003). Reproducing Kernel Hilbert Space Interpolation Methods as a Paradigm of High Dimensional Model Representations: Application to Multidimensional Potential Energy Surface Construction, *J. Chem. Phys.*, 119(13), 6433-6442.

[11] Wang, H. B., Zhu, L., Hase, W. L. (1994). A model multidimensional analytic potential-energy function for the $Cl^- + CH_3Br \rightarrow ClCH_3 + Br^-$ reaction, *J. Phys. Chem.*, 98(6), 1608-1619.

[12] Wang, H. B., Peslherbe, G. H., Hase, W. L. (1994). Trajectory studies of S_N2 nucleophilic-substitution 4. intramolecular and unimolecular dynamics of the $Cl^-...CH_3Br$ and $ClCH_3...Br^-$ complexes, *J. Amer. Chem. Soc.*, 116(21), 9644-9651.

[13] Hase, W. L. (1994). Simulations of gas-phase chemical-reaction - applications to S_N2 nucleophilic-substitution, *Science*, 266(5187) 998-1002.

[14] Wang, H. B., Hase, W. L. (1995). Statistical rate theory calculations of the $Cl^- + CH_3Br \rightarrow ClCH_3 + Br^-$ rate-constant versus temperature, *J. Amer. Chem. Soc.*, 117(36), 9347-9356.

[15] Peslherbe G. H., Wang H. B., Hase W. L. (1996). Trajectory studies of S_N2 nucleophilic substitution 5. Semiempirical direct dynamics of Cl^-CH_3Br unimolecular decomposition, *J. Amer. Chem. Soc.*, 118(9), 2257-2266.

[16] Wang H. B., Hase W. L. (1996). Reaction path Hamiltonian analysis of the dynamics for $Cl^- + CH_3Br \rightarrow ClCH_3 + Br^-$ S_N2 nucleophilic substitution, *Chem. Phys.*, 212(2-3), 247-258.

[17] Raugei, S., Cardini, G., Schettino, V. (1999). An ab initio molecular dynamics study of the S_N2 reaction $Cl^- + CH_3Br \rightarrow ClCH_3 + Br^-$, *J. Chem. Phys.*, 111(24), 10887-10894.

[18] Sun, L., Hase, W. L., Song, K. (2001). Trajectory studies of S_N2 nucleophilic substitution 8. Central barrier dynamics for gas phase $Cl^- + CH_3Cl$, *J. Amer. Chem. Soc.*, 123(24), 5753-5756.

[19] Wang, Y. F., Hase, W. L., Wang, H. B. (2003). Trajectory studies of S_N2 nucleophilic substitution. IX. Microscopic reaction pathways and kinetics for $Cl^- + CH_3Br$, *J. Chem. Phys.*, 118(6), 2688-2695.

[20] Hu, X. C., Hase, W. L. (1988). Effect of anharmonicity on intermolecular energy-transfer from highly vibrationally excited molecules, *J. Chem. Phys.*, 92(14), 4040-4046.

[21] Oref, I., Tardy, D. C. (1990). Energy-transfer in highly excited large polyatomic-molecules, *Chem. Rev.*, 90(8), 1407-1445.

[22] Gilbert, R. G. (1995). Mechanism and models for collisional energy-transfer in highly excited large polyatomic-molecules - invited review, *Australian J. of Chem.*, 48(11),1787-1817.

[23] Marques, J. M. C., Martinez-Nunez, E., Fernadez-Ramos, A., et al. (2005). Trajectory dynamics study of the Ar+CH_4 dissociation reaction at high temperatures: the importance of zero-point-energy effects, *J. Phys. Chem. A*, 109(24), 5415-5423.

[24] Troya, D. (2005). Quasiclassical trajectory study of energy transfer and collision-induced dissociation in hyperthermal Ar+CH_4 and Ar+CF_4 collisions, *J. Phys. Chem. A*, 109(26), 5814-5824.

[25] Marques, J. M. C., Martinez-Nunez, E., Vazquez, S. A. (2006). Trajectory dynamics study of collision-induced dissociation of the Ar + CH_4 reaction at hyperthermal conditions: Vibrational excitation and isotope substitution, *J. Phys. Chem. A*, 110(22), 7113-7121.

[26] * Bieniasz, L. K., Rabitz, H. (2006). Extraction of parameters and their error distributions from cyclic voltammograms using bootstrap resampling enhanced by solution maps: computational study, *Anal. Chem.*, 78, 8430-8437.

[27] * Li, G., Ho, T. -S., Rabitz, H. (2008). Random sampling-high dimensional model representation (RS-HDMR) with application on molecular dynamics simulations. manuscript in preparation.

* Support partly or totally under the present grant.