

The Hessian biased singular value decomposition method for optimization and analysis of force fields

Siddharth Dasgupta,^{a)} Terumasa Yamasaki,^{b)} and William A. Goddard III^{a)}

Materials and Process Simulation Center, Beckman Institute, (139-74), Division of Chemistry and Chemical Engineering (CN 8768), California Institute of Technology, Pasadena, California 91125

(Received 1 July 1993; accepted 14 November 1995)

We present methodology (HBFF/SVD) for optimizing the form and parameters of force fields (FF) for molecular dynamics simulations through utilizing information about properties such as the geometry, Hessian, polarizability, stress (crystals), and elastic constants (crystals). This method is based on singular value decomposition (SVD) of the Jacobian describing the partial derivatives in various properties with respect to FF parameters. HBFF/SVD is effective for optimizing the parameters for accurate FFs of organic, inorganic, and transition metal compounds. In addition it provides information on the validity of the functional form of the FF for describing the properties of interest. This method is illustrated by application to organic molecules (CH₂O, C₂H₄, C₄H₆, C₆H₈, C₆H₆, and naphthalene) and inorganic molecules (Cl₂CrO₂ and Cl₂MoO₂). © 1996 American Institute of Physics. [S0021-9606(96)01208-8]

I. INTRODUCTION

A great deal of progress has been made in extending the methods of molecular dynamics simulations to the prediction of structures, vibrational frequencies, thermochemical properties, and elastic constants of biological,¹ polymer,^{2,3} ceramic,⁴ and metallic⁵ systems. However, a continuing difficulty is determining the optimum force field (FF) parameters.⁶⁻⁹ Particularly difficult is the application to complex organic, inorganic, and transition metal compounds where the experimental data is generally inadequate to determine the FF parameters. To help solve this problem, the Hessian biased force field (HBFF)⁹ method was developed to utilize *ab initio* quantum chemical (QC) calculations to provide additional information missing in experiments. This HBFF approach has been utilized successfully to combine QC and experimental data to define accurate FF for polymers^{2,3} and ceramic⁴ crystals. There are two aspects of determining the FF for complex systems: (i) the functional form of the FF, and (ii) the refinement of FF parameters. We present herein the HBFF/SVD method, which provides a means for analyzing the efficacy of the functional form (keeping the FF simple and transferable by allowing nearly redundant terms to be identified and eliminated) and for determining the optimum FF parameters. This has been implemented in a computer program FFOPT which is used in conjunction with the commercial POLYGRAF molecular dynamics code.¹⁰

The properties currently considered in optimizing the FF are

- (a) the $(3N-6)$ forces (energy gradients)

$$E'_i = \frac{\partial E}{\partial X_i},$$

where N is number of atoms;

- (b) the $(3N-6)(3N-5)/2$ independent second derivatives, derived from the Hessian

$$E''_{ij} = \frac{\partial^2 E}{\partial X_i \partial X_j};$$

the six independent components of the polarizability tensor, P_α ;

- (c) the six independent lattice stresses (for crystals), Σ_α ;
(d) the 21 independent elastic constants (for crystals), $C_{\alpha\beta}$.

The exact properties may come from experiment or theory (QC).

We define the error in the ability of the FF to fit the properties as

$$\begin{aligned} \text{ERR} = & \sum_{i=1}^{3N-6} (\delta E'_i)^2 + \sum_{i \leq j=1}^{3N-6} (\delta E''_{ij})^2 + \sum_{i \leq j=1}^3 (\delta P_{ij})^2 \\ & + \sum_{i=1}^6 (\delta \Sigma_i)^2 + \sum_{i \leq j=1}^6 (\delta C_{ij})^2, \end{aligned} \quad (1)$$

where δ denotes the difference between the exact quantity and that calculated from the FF. Other properties can be added, but shown are the ones included in our present implementation (FFOPT¹⁰). These have been used to determine force fields for polymers (polyethylene,^{2(a)} polysilane,^{2(b)} polyvinylidene fluoride³) and ceramics (silicon nitride⁴). At the exact geometry the exact forces are all zero. Thus if the FF gives $\mathbf{E}'=0$, it will reproduce the exact geometry. The biased Hessian \mathbf{E}'' is constructed by combining *ab initio* theoretical vibrational modes with experimental frequencies.⁹ In optimizing the polarizability tensor³ we use quantum mechanical values because the experiment often does not determine all components (particularly orientation

^{a)} Authors to whom correspondence should be addressed.

E-mail: sdg@wag.caltech.edu and wag@wag.caltech.edu

^{b)} Present address: Asahi Chemical Industry Co., Ltd., 2-1 Samejima, Fuji, Shizuoka 416, Japan.

of the principal axes). In previous applications the quantities for crystals (Σ_i and C_{ij}) have come from experiment; however the increased accuracy of theory now makes the use of calculated crystal values a viable alternative. In this paper we focus on finite molecules with applications using only the \mathbf{E}' and \mathbf{E}'' terms of (1).

The ERR function depends nonlinearly on the FF parameters and the optimization is carried out iteratively. Evaluation of the Hessian $\delta\mathbf{E}''$ dominates the computational time, and we evaluate these terms using analytical second derivatives.¹⁰ For simple harmonic FF functions one can calculate the Jacobian analytically; however, this restricts the functional forms available for the FF. Instead our current implementation (FFOPT) calculates the Jacobian numerically using finite differences.¹¹

An important issue in developing the FF is the choice of the functional form. Since both *eigenfunctions* and *eigenvalues* of the molecular vibrations are included in the properties, the FF parameters are greatly over-determined. A major virtue of HBFF/SVD is that it leads to the residue projection index (RPI) and the parameter sensitivity index (PSI) which allows deficiencies and redundancies in FF terms to be identified. Current molecular dynamics programs¹⁰ for molecules and crystals allow numerous choices for the functional forms used to describe intramolecular interactions in terms of two-body, three-body, and four-body (and higher) interactions. This allows us to pinpoint where special FF terms are required for such unresolved interactions, providing insight into the nature and size of intramolecular interactions.

In Sec. II we present the HBFF/SVD method, and in Sec. III we define the RPI and PSI indices for analyzing the force field. These methods are applied to formaldehyde in Sec. IV, to linear polyenes and to benzene in Sec. V, and to Cl_2CrO_2 and Cl_2MoO_2 in Sec. VI. Details of the QC calculations are given in Sec. VII.

II. OPTIMIZATION OF THE FORCE FIELD

A. The least-squares problem

We define the residual property vector $\mathbf{b} = (b_1, b_2, \dots, b_{N_{\text{obs}}})$ as

$$\mathbf{b} = \mathbf{b}^{\text{obs}} - \mathbf{b}^{\text{FF}} \quad (2)$$

in terms of the exact properties of a molecule (\mathbf{b}^{obs}) and the values calculated using the FF (\mathbf{b}^{FF}). We want to find parameters $\mathbf{p} = (p_1, p_2, \dots, p_{N_{\text{par}}})$ that lead to $\mathbf{b} \rightarrow 0$. These FF parameters to be optimized include force constants, geometric parameters, nonbonding parameters, charges, polarizabilities, etc.

We define \mathbf{G} as the Jacobian or *sensitivity matrix* consisting of the partial derivatives of residual errors with respect to the FF parameters:

$$(\mathbf{G})_{i\mu} = \frac{\partial b_i}{\partial p_\mu}. \quad (3)$$

To first order, the new properties \mathbf{b}^{new} arising from changing the FF parameters from \mathbf{p}^{old} to \mathbf{p}^{new} is given by

$$\mathbf{b}_i^{\text{new}} = \mathbf{b}_i^{\text{old}} + \sum_{\mu=1}^{N_{\text{par}}} G_{i\mu} (p_\mu^{\text{new}} - p_\mu^{\text{old}}), \quad i = 1, 2, \dots, N_{\text{obs}}. \quad (4)$$

For $\mathbf{b}^{\text{new}} = 0$,

$$\mathbf{b}_i = \sum_{\mu=1}^{N_{\text{par}}} G_{i\mu} \cdot \Delta p_\mu, \quad i = 1, 2, \dots, N_{\text{obs}}, \quad (5a)$$

where $\Delta p_\mu = p_\mu^{\text{old}} - p_\mu^{\text{new}}$. Thus we solve the equation

$$\mathbf{b} = \mathbf{G} \cdot \Delta \mathbf{p}, \quad (5b)$$

for $\Delta \mathbf{p}$. Equation (5) consists of N_{obs} equations and N_{par} unknown variables. The Hessian leads to

$$N_H = \frac{(3N-6)(3N-5)}{2} \quad (6a)$$

constraints, while the geometry (forces) leads to

$$\begin{aligned} N_v &= (3N-5) (\text{linear molecule}), \\ N_v &= (3N-6) (\text{nonlinear molecule}) \end{aligned} \quad (6b)$$

additional constraints (after excluding translation and rotation). (The number of nonzero vibrational frequencies is equal to the number of nonzero forces; hence the notation N_v .) Thus

$$N_{\text{obs}} = N_v + N_H = \frac{(3N-6)(3N-3)}{2},$$

which generally greatly exceeds N_{par} ,

$$N_{\text{obs}} \geq N_{\text{par}}. \quad (7)$$

The applications presented here will consider finite molecules and will not fit polarization; thus, only the force and Hessian parts of (1) will be considered. In this case the vector of error residues becomes

$$\mathbf{b} = (\mathbf{b}_{\text{force}} | \mathbf{b}_{\text{Hessian}}). \quad (8)$$

The force part consists of the N_v residual forces, which should all be zero if the FF reproduces the optimum geometry.

The Hessian part consists of N_H independent components constructed from

$$\mathbf{b}_{\text{Hessian}} = \{H_{ij}^{\text{obs}} - H_{ij}^{\text{FF}}\}, \quad i \leq j = 1, \dots, 3N, \quad (9)$$

where H^{obs} and H^{FF} denote the exact and FF Hessian matrices. H^{obs} is constructed from the mass-weighted *ab initio* Hessian

$$H_{ij}^{\text{QC}} = E''_{ij} / \sqrt{M_i M_j} \quad (10)$$

and the *experimental (or exact) frequencies*. The idea here⁹ is that even high quality *ab initio* wave functions may lead to significant error in the calculated frequencies (e.g., 10% to 20% for HF, 5 to 10% for MP2, even with good basis sets). Thus, we should always use experimental frequencies or else theoretical frequencies scaled to predict experimental values. However, the vibrational modes (wave function described in terms of components on each atom) is available only from

theory. In HBFF we combine the best information from theory and experiment. Diagonalizing the *ab initio* QC Hessian

$$\sum_i^{N_{\text{mode}}} \mathbf{x}_i^t \cdot H^{\text{QC}} \cdot \mathbf{x}_i = \lambda^{\text{QC}} \quad (11)$$

(where λ is diagonal) leads to the *ab initio* QC vibrational frequencies, ν_j^{QC} , where

$$\nu_j^{\text{QC}} = C_{\text{vib}} \sqrt{\lambda_j^{\text{QC}}} \quad (12)$$

(The conversion factor is $C_{\text{vib}} = 108.5913$ if H^{QC} is evaluated using energy in kcal/mol, distance in Å, and mass in a.m.u., and ν is in cm^{-1} .) With HBFF, $\{\nu^{\text{QC}}\}$ is replaced by the experimental or exact frequencies¹² $\{\nu^{\text{EX}}\}$. This leads to a biased Hessian of the form

$$H^{\text{obs}} = \sum_i^{N_{\text{mode}}} \mathbf{x}_i \cdot (\nu_i^{\text{EX}}/C_{\text{vib}})^2 \cdot \mathbf{x}_i^t \quad (13)$$

In the error residue (1), one can use all the elements of the Hessian as in (9) or one can separately consider the frequency information from the diagonalized Hessian

$$\mathbf{b}_{\text{frequency}} = \{(\nu_i^{\text{EX}} - \nu_i^{\text{FF}})^2\}, \quad i = 1, \dots, N_\nu, \quad (14a)$$

and the mode information

$$\mathbf{b}_{\text{mode}} = \sum_i (x_{\mu i}^{\text{obs}} - x_{\mu i}^{\text{FF}}), \quad \mu = 1, \dots, N; \quad i = 1, \dots, N_\nu. \quad (14b)$$

Here (14b) leads to

$$N_m = N_H - N_\nu = \frac{(3N-5)(3N-4)}{2}$$

independent conditions.

For molecules where there is little experimental frequency data and where it is not possible to predict scaling factors, it is better to use the full constraints of the Hessian as in (9). Since extended isotopic shift data implicitly provides experimental information on normal modes, molecules for which there is extensive experimental isotopic shift data do not require a strong weighting of the normal mode constraints (14b). In this case separate consideration of frequencies and modes as in (14) is better since it provides control over weights. As methods for direct determination of accurate Hessians from highly correlated *ab initio* wave functions become more practical, (9) will be more appropriate. FFOPT allows either approach. We refer to the use of (9) as *Hessian optimization* and the use of (14) as *frequency optimization*.

The optimum change in parameters, Δp , is obtained from (5) by left multiplying with \mathbf{G}^- , the generalized inverse of \mathbf{G} to give

$$\Delta \mathbf{p} = \mathbf{G}^- \cdot \mathbf{b}. \quad (15)$$

In general, some FF parameters will be either redundant or ineffective in improving the fit. This leads to a Jacobian matrix \mathbf{G} that is either singular or ill conditioned. Consequently we use SVD to solve (5). This provides two advantages: (i) SVD defines a useful \mathbf{G}^- even when \mathbf{G} is singular, and (ii)

SVD provides diagnostics to help remove redundancies in the FF. \mathbf{G} can always be decomposed into the product of three matrices¹³

$$\mathbf{G} = \mathbf{U} \cdot \mathbf{S} \cdot \mathbf{V}^T, \quad (16)$$

$$G_{j\mu} = \sum_{\nu=1}^{N_{\text{par}}} U_{j\nu} s_\nu V_{\mu\nu},$$

where $\mathbf{V} = (\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{N_{\text{par}}})$ is a square orthogonal matrix (i.e., $\mathbf{V}^{-1} = \mathbf{V}^T$) of dimension N_{par} , \mathbf{S} is a diagonal matrix of dimension N_{par} whose elements s_i are called the *singular values*, and $\mathbf{U} = (\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{N_{\text{par}}})$ is a column orthogonal matrix of dimension $N_{\text{obs}} \times N_{\text{par}}$ (the same shape as \mathbf{G}). The generalized inverse is constructed easily as

$$\mathbf{G}_{\mu j}^- = \sum_{\nu=1}^{N_{\text{par}}} V_{\mu\nu} \frac{1}{s_\nu} U_{j\nu}, \quad (17)$$

$$\mathbf{G}^- = \mathbf{V} \cdot \left[\text{diag} \left(\frac{1}{s_\nu} \right)' \right] \cdot \mathbf{U}^T,$$

where the prime indicates that for any $s_\nu = 0$, we set $1/s_\nu = 0$. This leads to the solution of (15) in the form

$$\Delta p_\mu = \sum_{\nu j} v_{\mu\nu} \frac{1}{s_\nu} u_{j\nu} b_j. \quad (18)$$

However, there are typically near redundancies that lead to very small but nonzero s_i . For such cases the Δp from (18) would be very large and inaccurate (due to round-off errors and nonlinear terms in \mathbf{G}). Such small singular values indicate that certain combinations for the current FF parameters are irrelevant to improving the fit. Thus, we will ignore small nearly singular values in (18).

The singular values are ordered with the largest values first. Thus, s_1 corresponds to the most significant combination of parameters (the one with the biggest impact on ERR). We define the *condition ratio* C_ν for the i th equation in (17) as the ratio

$$C_\nu = \frac{s_1}{s_\nu} \quad (19)$$

so that the largest C_ν corresponds to the least significant combination of parameters. We introduce the *critical condition number* C and construct the approximate inverse matrix, \mathbf{G}_C^- , as

$$(\mathbf{G}_C^-)_{\mu j} = \sum_{\nu}^{C_\nu < C} \mathbf{v}_{\mu\nu} \left(\frac{C_\nu}{s_1} \right) \mathbf{u}_{j\nu}. \quad (20)$$

This ignores modes with $C_\nu > C$ and leads to

$$\Delta p_\mu = \sum_{\nu j}^{C_\nu < C} V_{\mu\nu} \frac{1}{s_\nu} U_{j\nu} b_j, \quad \mu = 1, \dots, N_{\text{par}}. \quad (21)$$

The number of terms in the sum over ν in (21) is denoted as N'_{par} . We find below that $C = 1000$ generally leads to good solutions.

B. Weighting

The convenient units used for evaluating the properties (forces, Hessian elements, polarizabilities, etc.) from which \mathbf{b} is constructed may lead to large numerical values for some elements and small values for others. In addition, the number of terms in various properties differs dramatically. Thus Hessian optimization has $\sim 9N^2/2$ terms, whereas forces and frequencies have $\sim 3N$ terms, and polarizability has six terms.

To account for the differences in the magnitude of the components of \mathbf{G} , we nondimensionalize the Jacobian by writing $\mathbf{G}_{i\mu} = g_{i\mu}^0 G_{i\mu}^0$. Here $g_{i\mu}^0$ is set to the $G_{i\mu}$ for the original FF parameters and fixed. This leads to initial values of $G_{i\mu}^0 = 1$.

To properly weigh various properties for various sized molecules, we modify the error function in (1) to have the form

$$\begin{aligned} \text{ERR}(\text{Hessian}) = & \left(\frac{W_{\text{force}}^2}{N_\nu} \right) \sum_{i=1}^{3N} [w_{\text{force}}^i (\delta E_i')]^2 \\ & + \left(\frac{W_H^2}{N_H} \right) \sum_{i \geq j}^{3N} [w_H^{ij} (\delta E_{ij}'')]^2 \end{aligned} \quad (22a)$$

or

$$\begin{aligned} \text{ERR}(\text{frequency}) = & \left(\frac{W_{\text{force}}^2}{N_\nu} \right) \sum_{i=1}^{3N} [w_{\text{force}}^i (\delta E_i')]^2 \\ & + \left(\frac{W_\nu^2}{N_\nu} \right) \sum_{i=1}^{3N} [w_\nu^i (\delta \nu_i^2)]^2 \\ & + \frac{W_m^2}{(N_H - N_\nu)} \sum_i (w_m^i)^2 (\delta \mathbf{x}_i^t \cdot \delta \mathbf{x}_i), \end{aligned} \quad (22b)$$

TABLE I. Residue weights (for observation errors) and parameter weights used herein.

Weights for observation error residues ^a	
W_{force}	10
w_{force}^i	1
W_H	1
w_H	1
W_ν	1
w_ν^i	1 or $\left(\frac{1}{\nu_j^2}\right)^b$
W_{mode}	0.1 or 1 ^c
w_{mode}^{jk}	1 or $\left(\frac{1}{\nu_j \cdot \nu_k}\right)^b$

^aUpper case W refers to the overall weighting of a category [see Eq. (22)] whereas lower case w refer to the individual weighting of each term.

^bWith inverse frequency weighting, the normalizations are $\|w_H\|^2 = N_H$, $\|w_\nu\|^2 = N_\nu$, and $\|w_{H\text{-off-diag}}\|^2 = N_{H\text{-off-diag}}$. These weights were set to unity in this work.

^cIn this work, $W_{H\text{-off-diag}} = 0.1$ is used with the *frequency optimization* scheme.

where ERR(Hessian) and ERR(frequency) denote total error functions for *Hessian optimization* and *frequency optimization*, respectively. Here the normalization constants N_H , N_ν , and $(N_H - N_\nu)$ correspond to the number of property elements being fitted, leading to an rms error. This makes the error expression consistent for different-sized molecules. The property weighted W_x allow independent weighting of the various types of terms, for example, to make forces (geometry) more important than frequency and frequency more important than modes. Table I shows the weights used. There is more premium on force components ($W_{\text{force}} = 10$) than on the frequencies ($W_\nu = 1$), and even less on the modes ($W_{\text{mode}} = 0.1$). This ensures accurate reproduction of the correct structure.⁹ The preconditioning factors, w_{force} , w_H , w_ν etc. allow each observable to be weighted separately by their experimental uncertainties. However, for the results presented herein, these values are set to unity.

For the numerical calculation of the Jacobian we use a 1% displacement of all the force field parameters to be optimized with only forward differences (to save time).

III. ANALYSIS OF FORCE FIELDS

A. Residue projection index (RPI)

Even a FF with numerous cross terms (for example, the MSXX FF)^{2-4,9} may not satisfy all of the constraints imposed in the HBFF formalism. The RPI is defined to provide an index facilitating the identification of constraints difficult for the FF to satisfy. For each error residue b_i , the RPI is constructed from the columns of the \mathbf{U} matrix as

$$\text{RPI}(b_i) = \sum_{\nu}^{c_\nu < C} (U_{\nu i})^2. \quad (23)$$

If $\text{RPI}(b_i) = 1$, the corresponding residue component is exactly describable with the current set of terms in the FF and

TABLE II. Geometry and vibrational frequencies from Hartree Fock (HF) calculations of formaldehyde (HCHO). The 6-31G** basis was used.

	Expt.	HF	MP2
Geometry			
$R_{\text{C=O}}$ (Å)	1.203	1.189	1.219
$R_{\text{C-H}}$ (Å)	1.099	1.094	1.099
$\theta_{\text{H-C-H}}$ (deg.)	116.5	116.2	115.52
rmsD (Å) ^a	...	0.007	
Vibrational frequencies (cm ⁻¹)			
$A_1 \nu_s(\text{C-H})$	2783	3149.2	3030.3
$A_1 \nu_s(\text{C=O})$	1746	2007.9	1796.3
$A_1 \delta(\text{CH}_2)$	1500	1656.0	1589.1
$B_1 \nu_{as}(\text{C-H})$	2843	3226.7	3108.5
$B_1 \rho(\text{CH}_2)$	1249	1367.6	1297.7
$B_2 \omega(\text{CH}_2)$	1167	1335.3	1219.5
$\delta \nu_{\text{rms}}^b$...	263.5	156.6
		(12.82%)	(5.91%)

^arms error in x, y, z coordinates.

^brms error in frequencies.

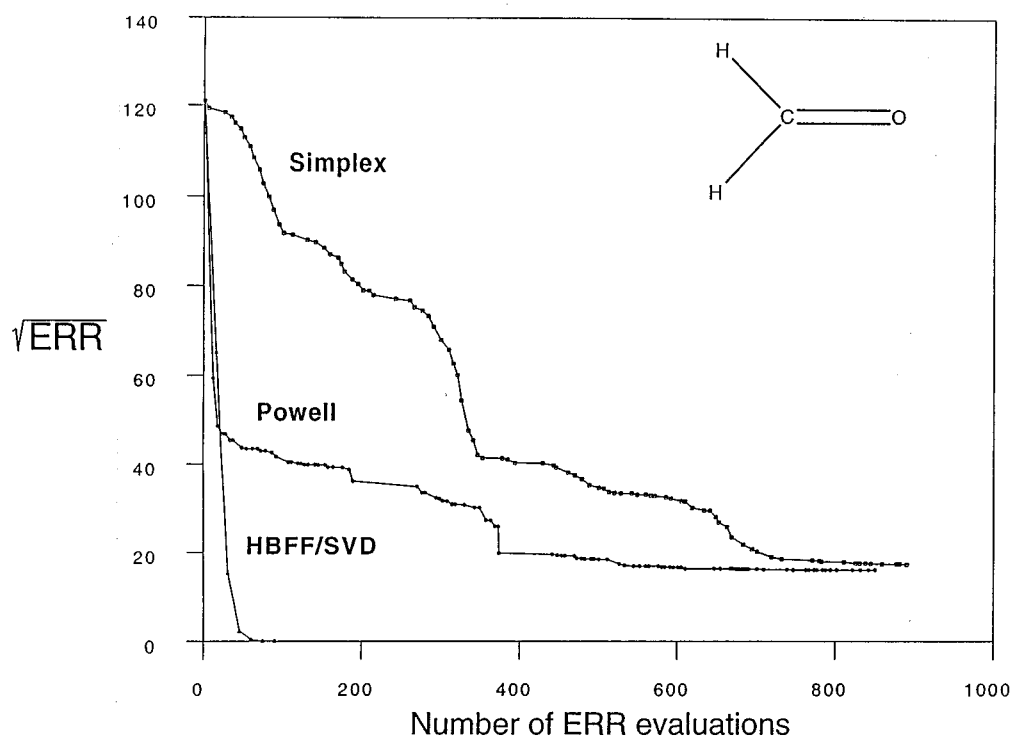


FIG. 1. Hessian optimizations of the MSX FF of HCHO. The square root of ERR is plotted.

TABLE III. HBFF optimization^a of the MSX FF for formaldehyde (HCHO) by Simplex, Powell, and HBFF/SVD methods. Results after one round of optimization are shown.

		Initial ^b	Simplex	Powell	HBFF/SVD
C-H	R_e	0.990	1.099	1.079	1.094
	K_R	700.00	621.85	716.06	649.45
	D_e	87.00	87.00	87.00	87.00
C-O	R_e	1.220	1.203	1.196	1.199
	K_R	1400.00	1632.43	1962.23	1826.33
	D_e	175.00	175.00	175.00	175.00
H-C-H	θ_e	120.00	116.35	125.14	124.23
	K_θ	100.00	66.29	46.91	46.72
	$K_{\theta,R}$	-10.00	-22.16	-68.00	-23.45
	$K_{R,R}$	10.00	-6.59	-4.11	10.92
O-C-H	θ_e	120.00	121.31	125.71	125.59
	K_θ	100.00	118.59	104.58	100.18
	$K_{\theta,R_{O-C}}$	-10.00	2.31	-122.02	-66.97
	$K_{\theta,R_{H-C}}$	-10.00	15.32	-79.73	-19.00
	$K_{R,R}$	10.00	-10.04	19.22	67.07
C-X-X-X	K_ψ	40.00	58.86	45.41	44.05
iter ^c (ERR eval) ^d			669(900)	10(854)	6(91)
ERR _{initial} ^e			14 634.65	14 634.65	14 634.65
ERR _{final} ^e			300.66	256.98	0.00
δv_{rms}^f			63.81	18.28	0.00
$\delta \text{force}_{\text{rms}}^g$			0.5004	0.0121	0.0001

^aUsing *Hessian optimization*.

^bExcept for cross terms and bond energies, all initial values are from DREIDING, Ref. 14.

^cNumber of iterations required for convergence to $|\text{ERR}^{(k+1)} - \text{ERR}^{(k)}| < 0.01$.

^dTotal number of error residue evaluations (corresponds to computational cost).

^eSum of squares error value at the beginning (ERR_{init}) and the end (ERR_{final}) of optimizations.

^frms frequency error between experimental and FF frequencies in cm^{-1} .

^grms residual force in $\text{kcal mol}^{-1} \text{\AA}^{-1}$.

the current condition number C . Thus proper choice of parameters should lead to an accurate description of b_i . If RPI (b_i) $\ll 1$, then no adjustment of parameters for the current FF will lead to a small error in b_i . Consequently RPI helps find deficiencies in the *form* of the FF for describing the selected set of molecular properties.

B. Parameter sensitivity index (PSI)

We use the columns of \mathbf{V} to define a PSI for each FF parameter

$$\text{PSI}(p_\mu) = \sum_{\nu}^{C_\nu < C} (\mathbf{V}_{\nu\mu})^2. \quad (24)$$

When $\text{PSI}(p_\mu)$ is close to unity, the parameter p_μ and its corresponding FF term is unique and active in the optimization. A small value of PSI indicates that the corresponding FF terms is redundant and/or that (p_μ) it is not sensitive for the current constraints. Parameters with very small PSI are not well determined and care must be taken in using FFs with such parameters.

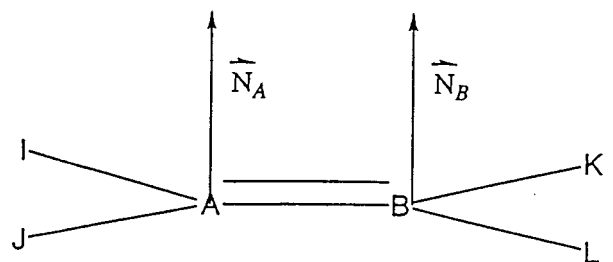
IV. FORMALDEHYDE

To illustrate the efficacy of HBFF/SVD, an MSX level FF⁹ was optimized for formaldehyde molecule using other standard methods (Simplex and Powell) in the HBFF formalism. The structure and vibrational frequencies used in optimizations are shown in Table II. We started with the DREIDING-II parameter values¹⁴ except for bond energies [which were set to experimental (adiabatic) values] and cross terms (for which initial values of 10.0 were assumed). The results after one round of *Hessian optimization* by these methods are shown in Fig. 1 and Table III. In this example, HBFF/SVD converges smoothly in six iterations. This shows the linear relationship between observed error components and FF parameters implied in (4) and strongly validates the derivative based approach of HBFF/SVD (over nonderivative approaches) for both efficiency and quality of optimization. HBFF/SVD requires evaluations of energy derivatives N_{par} times to construct the numerical \mathbf{G} matrix; however, it converges more quickly to better solutions. With Simplex or Powell, starting with initial parameters far from the optimum leads to significant drifting before eventually finding a satisfactory parameter set. This example also shows the efficiency of the Powell method over the Simplex method, especially at the early stage of optimizations. The Simplex method finally converged to a solution as good as for Powell, but it tended to become almost trapped in local minima during the course of optimization.

V. LINEAR POLYENES AND BENZENE

A. Ethylene (C_2H_4)

1. *Pi-twist*



(25)

For ethylene, there are three out-of-plane (oop) vibrational modes: the two wagging modes (B_{2g} and B_{1u}) and the twisting mode (A_u). It is common to describe oop modes in terms of torsions involving the four dihedral angles: $I-A-B-K$, $I-A-B-L$, $J-A-B-K$, and $J-A-B-L$ (where A and B are carbon atoms and the others are H atoms).

However, Table IV shows that such dihedral terms *cannot* describe the oop modes for C_2H_4 . This dihedral form for the torsional potential leads to a coupling between the twist and wag modes, that cannot be removed by adjusting parameters. Adjusting the dihedral torsional barrier to fit the twisting mode necessarily causes an error in the frequencies for the wagging (inversion) modes. Thus the splitting should be 8.8 cm^{-1} but dihedral torsions lead to a splitting of -129.6 cm^{-1} . This contrasts with the case of single bond systems where dihedral torsions have proved successful.

The problem here is that the physics of the torsional barrier for a double bond is different than that for a single bond. For single bonds the torsional barrier is dominated by Pauli orthogonality (nonbonded) interactions between the bonds connected to the bond of interest [involving next-nearest neighbor (1-4) atoms], referred to as nonbond repulsions. In principle this might be described in terms of the electrostatic and van der Waals interactions; however, the parameters for such interactions are determined from *intermolecular* packing of molecules and do not adequately describe the barrier for single bond dihedrals. Consequently, for single bonds this deficiency is removed by adding explicit torsional terms to the FF.

In contrast a double bond as in ethylene involves a σ bond that is relatively constant as one CH_2 group is rotated about the CC axis plus a π bond that is a maximum for the planar geometry ($\phi=0^\circ$) and zero for $\phi=90^\circ$. Thus the optimum geometry has eclipsed neighboring bonds, a behavior opposite that of nonbond repulsions which prefers staggered. In order to describe the FF of double bonded molecules such as ethylene, we reformulated the torsional term to reflect π overlap rather than the $X-A-B-X$ dihedrals. This new *pi-twist* form allows the twisting motion to be decoupled from the inversion modes.

TABLE IV. Comparison of RPI and PSI for ethylene FF optimization. Errors are in cm^{-1} , while RPI and PSI are dimensionless.

Experimental frequency	Torsion		Set 1		Set 2		Set 3	
	Errors	RPI	Errors	RPI	Errors	RPI	Errors	RPI
825.9	1.2	0.99	0.0	1.00	-0.1	1.00	0.0	1.00
948.8	-73.1	0.40	0.0	1.00	-0.1	1.00	0.0	1.00
940.0	59.8	0.60	0.0	1.00	-0.1	1.00	0.0	1.00
1026.5	1.9	1.00	0.0	1.00	-0.1	1.00	0.0	1.00
1220.0	-0.4	1.00	-0.4	1.00	-0.4	1.00	-0.4	1.00
1343.5	-0.5	1.00	0.0	1.00	0.0	1.00	0.0	1.00
1443.5	0.1	1.00	0.0	1.00	-0.1	1.00	0.0	1.00
1630.0	0.1	1.00	0.0	1.01	-0.2	1.00	0.0	1.00
3012.3	0.4	1.00	0.0	1.00	-11.2	0.74	0.0	1.00
3013.6	0.4	1.00	0.0	1.00	-11.1	0.74	0.0	1.00
3083.5	-0.6	1.00	13.5	0.50	25.1	0.26	13.5	0.50
3104.9	-0.1	1.00	-13.4	0.50	-1.7	0.26	-13.4	0.50
Parameter		PSI		PSI		PSI		PSI
C-H R_e		0.80		1.00		1.00		1.00
K_R		0.85		1.00		1.00		1.00
C-C R_e		1.00		1.00		1.00		1.00
K_R		1.00		1.00		1.00		1.00
H-C-H θ_e		0.94		0.77		0.77		0.75
K_θ		0.74		0.89		0.89		0.78
$K_{\theta,R_{C-H}}$		0.95						
$K_{R,R}$		0.14		1.00				1.00
C-C-H θ_e		0.99		0.79		0.79		0.79
K_θ		0.97		0.85		0.85		0.80
$K_{\theta,R_{C-H}}$		1.00						
$K_{\theta,R_{C-C}}$		1.00		1.00		1.00		1.00
K_ψ		0.00		0.70		0.70		0.72
K_{tor}		1.00						
$K_{\text{pi-twist}}$...		1.00		1.00		1.00
$K_{\text{wag/wag}}$...		1.00		1.00		1.00
K_{2c-aa}		1.00		1.00		1.00		1.00
$K_{C-H/C-H}$		1.00		1.00		1.00		1.00
K_{1c-aa}		0.05/0.58						0.07/0.09

The pi-twist coordinate is defined in terms of the normals to the planes defined by each central atom and its two substituents. Thus in (25) the cross product of the IA and JA vectors defines the normal to one plane (denoted \mathbf{N}_A) while the cross product of KB and LB defines the other (denoted \mathbf{N}_B). The cross product of \mathbf{N}_A with the AB vector and of \mathbf{N}_B with the AB vector defines two new vectors ($\mathbf{V}_{\pi A}$ and $\mathbf{V}_{\pi B}$) perpendicular to the bond. The angle between $\mathbf{V}_{\pi A}$ and $\mathbf{V}_{\pi B}$ is defined as the pi-twist coordinate ϕ_π . The potential is then taken as

$$E_\pi = \frac{1}{2}V_2(1 - \cos 2\phi_\pi) \quad (26)$$

(additional $\cos 4\phi_\pi$, $\cos 6\phi_\pi$,... terms can be included). Thus E_π describes the overlap between the π orbitals on atoms A and B . If \mathbf{N}_A or \mathbf{N}_B is not perpendicular to the AB axis, then the π orbitals on A and B are modified, leading to decreased bonding. This is included in the force field by defining an inversion coordinate ψ_A as the deviation of \mathbf{N}_A from $\mathbf{V}_{\pi A}$ and similarly for ψ_B . The new inversion term is defined as

$$E_{\text{inv}} = \frac{1}{2}V_{\text{wag-wag}}(\psi_A - \psi_B)^2 + \frac{1}{2}V_{\text{inv}}(\psi_A + \psi_B)^2. \quad (27)$$

2. Discussion of procedure

Using a large amount of experimental isotopic data, Duncan and Hamilton (DH)¹⁵ determined the most complete spectroscopic FF for ethylene. The spectroscopic FF uses simple harmonic terms expressed in terms of symmetry adapted valence coordinates. In contrast the MSXX FF, Table V, uses local modes, allows nonharmonic functional forms for angle bending, inversion, and pi-twist and includes charges and van der Waals terms.

Table VI reports the calculated and experimental isotopic shifts for various species. The calculated frequencies arise from expansions through second order in displacements from equilibrium. Hence the calculated frequencies do not include anharmonic effects in the modes even though the FF contains anharmonic terms. For ethylene there is sufficient experimental data to determine the anharmonic components and hence to obtain harmonic experimental values for comparison to a harmonic theoretical analysis. However for most molecules one must compare to direct experimental (anharmonic) frequencies, and to be consistent we do so here. For

TABLE V. Comparison of optimized parameters for ethylene.

		Dihedral	π Twist		
			Set 1	Set 2	Set 3
Bonds					
C-H	R_e	1.0633	1.0823	1.0823	1.0823
	K_R	838.20	728.66	728.42	728.66
C-C	R_e	1.3239	1.3385	1.3385	1.3391
	K_R	1528.07	1289.01	1291.58	1289.01
Angles					
H-C-H	θ_e	129.07	116.93	116.92	116.49
	K_θ	38.45	60.30	60.28	59.11
	$K_{\theta,R}$	-74.88	0.00	0.00	0.00
	$K_{R,R}$	1.61	5.52	0.00^a	5.52
C-C-H	θ_e	126.37	118.30	118.30	117.88
	K_θ	61.84	74.69	74.67	72.70
	$K_{\theta,R_{C-C}}$	-56.68	-32.38	-32.43	-32.38
	$K_{\theta,R_{H-C}}$	-41.51	0.00	0.00	0.00
	$K_{R,R}$	0.00	0.00	0.00	0.00
Inversion					
C-X-X-X	K_ψ	0.00	31.22	31.22	31.41
	$K_{\text{wag/wag}}$	0.00	1.20	1.20	1.20
Torsion					
H-C-C-H	K_{dihedral}	39.39	0.00	0.00	0.00
	$K_{\pi\text{-twist}}$	0.00	37.13	37.13	37.13
1-c ang-ang					
C-C-H-H	$K_{C-C-H/C-C-H}$	-8.02	0.00	0.00	-3.44^a
	$K_{C-C-H/H-C-H}$	-3.40	0.00	0.00	-2.80^a
2-c ang-ang					
H-C-C-H	$K_{2c\text{-angang}}$	-9.37	-13.70	-13.68	-13.70
2-c bnd-bnd					
H-C-C-H	$K_{C-H/C-H}$	-2.80	-4.43	-4.47	-4.43

^aDifferences with set 1 are in boldface.

cases where the anharmonicity is important this leads to additional errors in predicting isotopic shifts, but the results indicate the accuracy expected for typical cases.

3. Use of RPI and PSI

To illustrate the utility of HBFF SVD, Table IV shows the effect of individual terms on the optimization process. Parameter set 1 contains the complete FF reported in Table V. With set 1, all modes are well described except for the last two antisymmetric C-H stretching modes, where the errors are 13.5 cm^{-1} . The RPI indicate that the first 10 modes are adequately described (RPI=1) whereas the last two modes are not properly described (RPI=0.5). This is because the symmetry of these two modes is such that the two-center bond-bond coupling term cancels for each mode, leading to zero splitting. Allowing the sign of this bond-bond coupling term to change depending on whether the H's are *cis* or *trans*, leads to the proper splitting of these two modes. However, in MD various motions might convert *cis* atoms into *trans* and hence we eschew such distinctions. Alternatively we could have included a $\cos \phi$ term in the force constant; however, this is not yet allowed in POLYGRAF. Consequently, the present FF cannot eliminate the error in these

two modes, as indicated by RPI=0.5. On the other hand, the geminal C-H/C-H coupling constant is adequate to describe the splitting between the symmetric and antisymmetric CH_2 modes. This is evident from set 2, where this term is removed from the FF. Here there are four modes with RPI<1 leading to four modes with errors that cannot be removed by parameter optimization. Set 3 which starts with set 1 and adds one-center angle-angle coupling terms leads to no change in RPI. This indicates that the FF completeness is not improved. However, this additional coupling term causes curvature in several bending modes, which is reflected in the reduction of the PSI for the H-C-H K_θ and C-C-H K_θ . The small PSIs for the coupling terms (0.38 for C-C-H/C-C-H and 0.22 for C-C-H/H-C-H) indicate that these terms are not very active in the parameter set and can be eliminated, as in set 1. This illustrates the utility of both RPI and PSI in identifying those modes that can be adequately fit as well as those parameters that are redundant.

While FFOPT allows other types of cross terms, we have chosen only the small subset of Table V based on physical reasoning. We consider that the local coordinates should be in close proximity to have significant coupling. Thus for two C-H bonds separated by more than one C-C bond, we do

TABLE VI. Comparison of predicted and experimental frequencies (Ref. 15) (in cm^{-1}) for ethylene.

Mode	Source	$^{12}\text{C}_2\text{H}_4$	$^{12}\text{C}_2\text{D}_4$	$^{12}\text{C}_2\text{H}_2\text{D}_2$	$\text{HD}^{12}\text{C}_2\text{H}_2$	$^{13}\text{C}_2\text{H}_4$
Wag(B_{2g})	HBFF/SVD	940.0	775.6	943.8	943.7	927.0
	Expt.	940.0	780.0	942.4	943.4	...
Wag(B_{1u})	HBFF/SVD	948.8	717.5	747.8	803.9	943.6
	Expt.	948.8	720.0	750.5	807.9	...
Twist(A_u)	HBFF/SVD	1026.5	725.8	888.9	1000.9	1026.5
	Expt.	1026.5	726.0	887.0	1000.0	1026.5
Rock(B_{2u})	HBFF/SVD	825.9	592.9	682.7	730.8	824.7
	Expt.	825.9	595.0	687.2	730.0	824.9
Rock(B_{1g})	HBFF/SVD	1219.6	982.0	1128.2	1117.3	1203.3
	Expt.	1220.0	1000.0	1142.4	1128.5	...
Scissor(A_g)	HBFF/SVD	1343.5	980.9	1023.3	1284.3	1328.9
	Expt.	1343.5	984.6	1031.0	1289.5	1329.6
Scissor(B_{3u})	HBFF/SVD	1443.5	1067.0	1379.4	1398.3	1438.0
	Expt.	1443.5	1077.9	1384.0	1400.7	1437.8
C=C(A_g)	HBFF/SVD	1630.0	1496.5	1576.5	1598.2	1588.6
	Expt.	1630.0	1518.4	1585.0	1607.0	1587.6
C-H <i>s</i> -str. (B_{3u})	HBFF/SVD	3012.3	2179.0	2215.4	2265.1	3007.0
	Expt.	3012.3	2201.0	2229.8	2273.8	3007.7
C-H <i>s</i> -str. (A_g)	HBFF/SVD	3013.6	2247.9	3012.9	3012.9	3004.6
	Expt.	3013.6	2261.6	3009.0	3016.2	3006.5
C-H <i>a</i> -str. (B_{2u})	HBFF/SVD	3091.5	2302.5	3093.9	3094.1	3078.5
	Expt.	3104.9	2341.8	3094.1	3096.1	3092.8
C-H <i>a</i> -str. (B_{1g})	HBFF/SVD	3097.0	2332.3	2318.5	3056.0	3081.9
	Expt.	3083.5	2315.4	2335.0	3061.6	3072.1
rms error		5.49	16.99	8.57	5.59	5.83

not allow an interaction constant. Second, two local coordinates with widely spaced primary or group frequencies are not coupled even if the local coordinates are in close proximity. This latter case is illustrated by the absence of the C-H/C-C-H bond-angle coupling term. The C-H modes are at about 3000 cm^{-1} , whereas the C-C-H deformation modes are in the range of $800\text{--}1500\text{ cm}^{-1}$. Thus splittings between these two classes of modes can be handled adequately with the diagonal force constants. Where cross terms are clearly necessary is for local modes of the same type, for example, two rocking modes, two scissoring modes, or two symmetric CH_2 stretching modes.

B. Butadiene

Butadiene (C_4H_6) the second member of the polyene family, provides an interesting contrast. Clearly, the two double bonds must be treated differently from the central single bond. Since all atoms are C2 (i.e., sp^2 centers), we use the polyene FF form based on bond orders. For each C2-C2 bond, there are separate parameters r_e and K_b corresponding to bond orders of 2 and 1. (In a more sophisticated FF such as UFF^{16(a)} and MM3,^{16(b)} quantum mechanical or semi-empirical calculations are used to compute the bond order, allowing scaling of the bond parameters uniformly with the bond order.) For the angle term there are two bond-angle coupling constants (C=C/CCC and C-C/CCC). Similarly, the two-center bond-bond coupling is different for C=C/C=C vs C-C/C-C. The pi-twist barriers are also different around C=C and C-C bonds as are the couplings of inversions at two adjacent centers.

The optimized HBFF is reported in Table VII. The addi-

TABLE VII. The polyene FF (HBFF/SVD) for butadiene and hexatriene.

Term	Type	C_4H_6	C_6H_8
C-H	R_e	1.0733	1.0736
	K_R	725.7811	717.7557
C-C	R_e	1.4585	1.4627
	K_R	726.9573	760.1899
C=C	R_e	1.3169	1.3214
	K_R	1233.6013	1196.3834
H-C-H	θ_e	114.80	113.56
	K_θ	63.8094	63.6054
	$K_{R,R}$	-1.3776	-1.4567
C-C-H	θ_e	116.83	115.17
	K_θ	64.2198	63.5594
	$K_{\theta,R-C}$	-24.6066	-28.9749
C-C-C	θ_e	120.03	118.97
	K_θ	53.6713	55.5773
	$K_{\theta,R-C=C}$	60.1463	-37.1391
C-X-X-X	$K_{\theta,R-C-C}$	-6.7886	-4.5030
	$K_{R,R}$	136.4053	50.0876
	K_ψ	29.1760	30.4579
C-C-C-H	K_ψ	60.8119	59.6923
C=C	$K_{\text{pi-twist}}$	37.0845	28.5709
C-C	$K_{\text{pi-twist}}$	11.2376	11.2693
C=C	$K_{\text{wag/wag}}$	0.5553	1.1085
C-C	$K_{\text{wag/wag}}$	-1.0491	0.2121
H-C-C-H	$K_{2c\text{-angang}}$	-5.6231	8.4885
H-C=C-H	$K_{C\text{-H/C-H}}$	-7.3592	-6.7484
H-C-C-H	$K_{C\text{-H/C-H}}$	-11.9518	-8.6818
C-C-C-H	$K_{2c\text{-angang}}$	-3.3265	2.6262
C=C-C=C	$K_{2c\text{-angang}}$	-61.0503	69.0851
C=C-C=C	$K_{C\text{-C/C=C}}$	0.8171	-7.1634
C-C=C-C	$K_{C\text{-C/C-C}}$		-53.0395

TABLE VIII. Comparison of (a) RPI and (b) PSI for the butadiene and hexatriene FF optimizations.

C ₄ H ₆			C ₆ H ₈		
Frequency (cm ⁻¹)	Errors (cm ⁻¹)	RPI	Frequency (cm ⁻¹)	Errors (cm ⁻¹)	RPI
(a)					
163.0	6.0	1.00	94.0	4.3	0.84
301.0	-1.0	1.00	152.0	-2.3	0.33
513.0	-1.1	1.00	215.0	-7.1	0.76
524.0	-20.9	1.00	248.0	-24.5	0.09
753.0	-67.2	0.84	349.0	-2.4	0.41
890.0	-8.6	0.94	441.0	0.8	0.59
911.0	7.9	0.61	541.0	10.4	0.68
908.0	11.6	0.61	615.0	-52.9	0.21
991.0	-7.7	0.43	683.0	-88.2	0.22
967.0	50.1	0.94	868.0	-10.4	0.85
1013.0	10.8	1.00	900.0	-15.7	0.75
1205.0	3.0	0.98	901.0	-9.1	0.58
1291.0	-1.9	1.00	966.0	-60.5	0.26
1296.0	11.0	0.65	930.0	-19.7	0.17
1385.0	1.7	0.42	938.0	34.9	0.66
1442.0	-1.0	0.81	985.0	33.9	0.43
1599.0	2.5	0.87	1009.0	52.9	0.45
1643.0	-3.0	1.00	1188.0	-10.8	0.90
3014.0	-21.9	0.16	1132.0	24.5	0.90
2985.0	6.7	0.16	1255.0	-0.5	0.60
3014.0	14.3	0.68	1283.0	-9.4	0.97
3056.0	18.4	0.64	1296.0	36.5	0.24
3101.0	-7.4	0.17	1288.0	54.6	0.28
3102.0	-9.1	0.17	1397.0	5.6	0.45
			1433.0	-16.2	0.58
			1574.0	7.5	0.87
			1629.0	-22.8	0.34
			1628.0	13.4	0.84
			2997.0	-19.9	0.20
			2992.0	-14.7	0.19
			3018.0	4.6	0.76
			3000.0	27.1	0.51
			3017.0	26.6	0.43
			3046.0	8.0	0.64
			3100.0	-18.4	0.13
			3088.0	-6.4	0.13
Term	Parameter	C ₄ H ₆ PSI	C ₆ H ₈ PSI		
(b)					
C-H	R _e	1.00	1.00		
	K _R	1.00	1.00		
C-C	R _e	1.00	1.00		
	K _R	1.00	1.00		
	R _e	1.00	1.00		
	K _R	1.00	1.00		
H-C-H	θ _e	0.88	1.00		
	K _θ	0.97	1.00		
	K _{R,R}	0.00	0.01		
C-C-H	θ _e	0.74	1.00		
	K _θ	0.82	1.00		
	K _{θ,R-C}	1.00	1.00		
C-C-C	θ _e	0.98	1.00		
	K _θ	0.99	1.00		
	K _{θ,R-C=C}	0.94	0.94		
	K _{θ,R-C-C}	1.00	1.00		
	K _{R,R}	0.99	0.99		
C-X-X-X	K _ψ	0.76	1.00		
C-C-C-H	K _ψ	0.99	1.00		
C=C	K _{pi-twist}	1.00	1.00		

TABLE VIII. (Continued.)

Term	Parameter	C ₄ H ₆ PSI	C ₆ H ₈ PSI
C-C	K _{pi-twist}	1.00	0.99
C=C	K _{wag/wag}	1.00	1.00
C-C	K _{wag/wag}	1.00	0.02
H-C-C-H	K _{2c-angang}	0.87	1.00
H-C=C-H	K _{C-H/C-H}	0.00	0.99
H-C-C-H	K _{C-H/C-H}	1.00	1.00
C-C-C-H	K _{2c-angang}	0.98	0.99
C-C-C-C	K _{2c-angang}	1.00	1.00
C-C=C-C	K _{C-C/C-C}	0.00	0.08
C=C-C=C	K _{C=C/C=C}	0.11	0.99

tional complexity of butadiene compared to ethylene could justify including additional interactions in the FF. However, we did not allow all possible such couplings. For example, we do not include coupling between the twist about the two adjacent partial double bonds (which would split the two CH₂ twisting modes). Table VIII clearly indicates this deficiency, where a number of modes have RPI smaller than 1.00. The normal mode compositions are reasonably accurate, and the isotopic shifts (reported in Table IX) are in reasonable agreement with the experimental values.

C. Hexatriene

The parameters of hexatriene (C₆H₈) can be used to define the FF for all longer polyene analogs. The structure and Hessian for hexatriene were predicted using the optimized butadiene FF. It would have been straightforward to use the *ab initio* values; however, we wanted to illustrate how to handle cases where it is not practical to perform the *ab initio* calculations. Rather than the *ab initio* Hessian,¹⁸ we construct the Hessian of hexatriene using C4H6FF. The frequencies computed from the C4H6FF Hessian are about 1% off from the values that would be obtained from the *ab initio* Hessian of hexatriene. However, with HBFF they are replaced by the experimental values (reported by Langkilde *et al.*¹⁹ including extensive isotopic substitution data). There are two reasons for these slight errors in the calculated frequencies of C₆H₈ using C4H6FF:

- Errors in structure: The geometry of C₆H₈ predicted with C4H6FF will differ slightly from the exact geometry or the calculated *ab initio* geometry.²⁰ *Ab initio* calculations of the geometry of C₆H₈ are in good agreement amongst themselves but differ substantially from the only gas-phase experimental determination (a quite old electron diffraction study²¹). The theoretical values are more accurate but they should be corrected for vibrational averaging.
- π electron delocalization: As the length of the polyene chain increases^{20(c)} resonance or delocalization systematically reduces the difference between the double and single bonds from Δ=0.13 Å in C₄H₆ to Δ=0.11 Å for C₄₄H₄₆. This electronic effect is a function of the bond order for each CC bond and could be calculated from

TABLE IX. Comparison of predicted and experimental frequencies (in cm^{-1}) for butadiene.

Mode	Symmetry	C_4H_6		C_4D_6	
		HBFF/SVD	Expt. (Ref. 17)	HBFF/SVD	Expt. (Ref. 17)
(a) Out-of-plane					
C–C twist	A_u	154.7	163.0	133.2	140.0
C=C twist	A_u	493.5	524.0	363.6	381.0
C=C twist	B_g	661.7	753.0	527.3	603.0
CH_2 wag	B_g	903.7	911.0	703.1	702.0
CH_2 wag	A_u	908.1	908.0	719.5	718.0
CH wag	B_g	1014.0	967.0	815.1	795.0
CH wag	A_u	1033.5	1013.0	766.3	770.0
(b) In-plane					
CCC δ	B_u	296.3	301.0	261.9	240.0
CCC δ	A_g	506.5	513.0	471.4	440.0
C–C ν + CH_2 ro.	A_g	897.1	890.0	684.8	741.0
CH_2 ro.	B_u	952.5	991.0	722.0	739.0
CH_2 ro.+C–C ν	A_g	1202.8	1205.0	1208.7	1186.0
CH^ν ro.+C=C ν	A_g	1300.8	1291.0	915.0	919.0
CH^ν ro.	B_u	1310.7	1296.0	963.7	1009.0
CH_2 sc.+C=C ν	B_u	1390.1	1385.0	1021.5	1042.0
CH_2 sc.	A_g	1433.2	1442.0	1015.1	1048.0
C=C ν + CH_2 sc.	B_u	1597.0	1599.0	1509.5	1523.0
C=C ν + CH_2 sc.	A_g	1643.1	1643.0	1568.9	1583.0
CH_2 s-st.	B_u	2992.2	2985.0	2194.8	2215.0
CH_2 s-st.	A_g	2991.7	3014.0	2189.1	2205.0
CH^ν st.	A_g	3033.6	3014.0	2276.4	2262.0
CH^ν st.	B_u	3071.4	3056.0	2278.5	2255.0
CH_2 a-st.	B_u	3095.1	3102.0	2310.5	2355.0
CH_2 a-st.	A_g	3095.9	3101.0	2318.5	2341.0

the wave function or estimated from various approximate methods.²² However, we have not incorporated such effects into the FF.

Ab initio calculations show that as the chain length increases, the inner double bond constants become smaller than the outer and terminal double bonds, with the opposite for the single bonds. The CCC deformation constants vary little. The torsional force constant decreases for the inner double bonds and increases for the inner single bonds. Couplings between adjacent double bonds get more negative with increasing chain length, as do couplings between adjacent single bonds (sign negative). Coupling between adjacent single and double bonds are positive and increase with chain length but decrease with proximity. All such couplings go toward zero for more distant pairs.

These FF trends are found for the $\text{C}_4\text{H}_6\text{FF}$ and $\text{C}_4\text{H}_8\text{FF}$, Table VII. The exceptions are: the pi-twist barrier around C–C does not increase noticeably, the C=C/C=C coupling is positive for butadiene but negative for hexatriene, the C=C/C–C coupling decreases in hexatriene. The polyene FF does *not* distinguish between inner and outer parts; hence, the force constants must represent an average over these values. Within these limitations, we optimized the polyene FF using the experimental vibrational frequencies for hexatriene. Table VII shows the FF and Table X shows the vibrational frequencies.

D. Higher polyenes

We used the $\text{C}_6\text{H}_8\text{FF}$ to predict the frequencies for octatetraene^{23(a)} which are compared with experiment in Table XI. The structure for octatetraene from the x-ray analysis²⁴ disagrees with the *ab initio* calculations.²³ In particular, the x-ray structure has the inner double bonds shorter by 0.009 Å than the outer double bonds, whereas *ab initio* calculations lead to the inner double bonds being longer by 0.007 Å.

E. Aromatics

For benzene and the higher polycyclic aromatic hydrocarbons, the dependence of the force constants on the π -bond order becomes critical in distinguishing various C–C bonds. Since our polyene FF does not allow the use of continuous bond orders, we optimized the FF only for benzene and test it by calculating the vibrational frequencies for naphthalene.

For benzene, the long history of FF development has been reviewed by Goodman, *et al.*²⁵ Essentially all previous work used valence-only FF in terms of symmetry adapted coordinates. Hence, it is difficult to compare with our local FF which also includes nonbond terms (charge and vdW). There are quite a few complicated interaction constants in such FF for benzene (e.g., between the two b_{2u} modes ν_{14} and ν_{15}) which are not included in our local FF description.

TABLE X. Comparison of predicted and experimental frequencies (in cm^{-1}) for all-*trans* hexatriene. Values in parentheses are calculated, not observed.

Mode	Symmetry	Natural abundance		3,4-deutero-hexatriene	
		HBFF/SVD	Expt. (Ref. 19)	HBFF/SVD	Expt. (Ref. 19)
Out-of-plane					
C–C twist	A_u	98.3	94.0	97.6	87.0
C=C twist	A_u	222.0	248.0	216.9	247.0
C=C twist	A_u	593.4	683.0	587.3	658.0
CH ₂ wag	A_u	883.6	900.0	886.2	902.0
CH wag inner	A_u	973.9	938.0	751.8	736.0
CH wag outer	A_u	1062.5	1008.0	1023.3	992.0
C–C twist	B_g	207.7	215.0	192.1	201.0
C=C twist	B_g	560.5	615.0	535.4	558.0
CH wag inner	B_g	857.2	868.0	759.0	787.0
CH ₂ wag	B_g	890.3	901.0	885.0	901.0
CH wag outer	B_g	1019.5	985.0	1018.6	984.0
In-plane					
CCC δ outer	A_g	344.9	353.0	341.9	345.0
CCC δ inner	A_g	441.6	444.0	437.2	438.0
CH ₂ rock	A_g	909.3	930.0	878.2	871.0
C–C ν +CH rock	A_g	1176.9	1188.0	1176.9	1201.0
CH rock outer+inner	A_g	1274.2	1283.0	970.5	1001.0
CH rock inner+outer	A_g	1342.9	1288.0	1288.1	1285.0
CH ₂ sciss	A_g	1401.4	1397.0	1388.8	1396.0
C=C ν	A_g	1581.5	1574.0	1568.1	1565.0
C=C ν	A_g	1641.2	1574.0	1613.9	1607.0
CH ₂ <i>s</i> -str.	A_g	2977.3	2992.0	2977.3	2999.0
CH str. inner+outer	A_g	3027.2	3000.0	2280.6	2241.0
CH str. outer+inner	A_g	3043.6	3017.0	3037.2	3015.0
CH ₂ <i>a</i> -str.	A_g	3081.7	3088.0	3081.4	3089.0
CCC δ outer	B_u	149.1	152.0	146.9	151.0
CCC δ inner	B_u	549.9	541.0	537.1	512.0
CH ₂ rock	B_u	904.9	966.0	880.0	(875)
C–C ν	B_u	1155.7	1132.0	1149.8	1133.0
CH rock inner+outer	B_u	1254.2	1255.0	973.7	1044.0
CH rock outer	B_u	1333.2	1296.0	1297.7	1294.0
CH ₂ sciss	B_u	1415.8	1433.0	1401.9	1425.0
C=C ν	B_u	1606.4	1629.0	1604.8	1623.0
CH str.	B_u	3022.5	3018	2238.5	2230.0
CH ₂ <i>s</i> -str.	B_u	2977.1	(2997.0)	2977.1	3016.0
CH str.	B_u	3054.0	3046.0	3036.5	3051.0
CH ₂ <i>a</i> -str.	B_u	3081.7	3100.0	3081.5	3099.0
rms err. ^a		32.35(26.38)		25.54(22.94)	

^aValues in parentheses for in-plane modes only, which are better described with our cross-terms than the out-of-plane modes.

Even so the simple HBFF for benzene does well (Table XII), especially for the in-plane modes (all are well described with the set of force and interaction constants reported in Table XII). The fit for out-of-plane modes is much poorer (they require additional interaction constants). This deficiency in the FF is also evident from the RPI reported in Table XII.

Using the C6H6FF we calculated the structure and frequencies for naphthalene.²⁶ The results are shown in Table XIII. The average frequency error is 50.8 cm^{-1} while the average error in the bond distances is 0.008 \AA .

VI. Cl_2CrO_2 AND Cl_2MoO_2

A. Molecular properties

In order to illustrate the use of HBFF/SVD on complex inorganic molecules, we obtained a FF of Cl_2CrO_2 and

Cl_2MoO_2 , Fig. 2. Experimental structures²⁷ and vibrational frequencies²⁸ are available for these compounds.

As the first step of HBFF, we carried out *ab initio* HF calculations, leading to the results in Table XIV. Clearly HF calculations are inadequate for predicting either molecular structure or vibrational frequencies for these molecules. The rms differences in experimental and HF structures are 0.092 \AA for Cl_2CrO_2 and 0.142 \AA for Cl_2MoO_2 , with the $\text{M}=\text{O}$ bond distance underestimated by 0.23 \AA for Cl_2MoO_2 ! This difficulty for HF to describe an oxo bond to a transition metal is well known.^{28(j)–28(k)} In addition, experiments and HF calculations lead to a *different order* in the vibrational modes of A_1 symmetry. For both Cl_2CrO_2 and Cl_2MoO_2 , experiment puts $A_1 \nu_s(\text{M}-\text{Cl})$ above $A_1 \delta(\text{MO}_2)$ (by 114 and 92 cm^{-1} , respectively) while HF puts it below (by 231 and

TABLE XI. Comparison of predicted and experimental frequencies (in cm^{-1}) for all-*trans* octatetraene. Values in parentheses are calculated, not observed.

	Mode	Symmetry	HBFF/SVD	Expt. [Ref. 23(a)]	SQM [Ref. 23(b)]	
Out-of-plane	C–C twist	A_u	52.6	(60.0)	58.0	
	C=C twist	A_u	151.6	181.0	167.0	
	C–C twist	A_u	221.9	245.0	239.0	
	C=C twist	A_u	570.0	629.0	617.0	
	CH wag inner	A_u	833.5	840.0	843.0	
	CH ₂ wag	A_u	882.0	900.0	926.0	
	CH wag outer	A_u	997.5	960.0	976.0	
	CH wag inner+outer	A_u	1064.6	1011.0	1027.0	
	C–C twist	B_g	141.8	(151.0)	146.0	
	C=C twist	B_g	294.2	343.0	336.0	
	C=C twist	B_g	575.9	(647.0)	654.0	
	CH wag inner	B_g	872.0	877.0	892.0	
	CH ₂ wag	B_g	886.5	896.0	923.0	
	CH wag inner+outer	B_g	965.9	(926.0)	943.0	
	H wag outer	B_g	1037.0	(1002.0)	1011.0	
	In-plane	CCC δ	A_g	219.0	(219.0)	215.0
		CCC δ	A_g	357.1	343.0	334.0
		CCC δ	A_g	544.3	538.0	528.0
		CH ₂ rock	A_g	914.9	956.0	954.0
C–C ν		A_g	1177.8	1136.0	1124.0	
C–C ν		A_g	1189.1	1179.0	1187.0	
CH rock		A_g	1305.7	1281.0	1288.0	
CH rock		A_g	1329.1	1291.0	1304.0	
CH rock		A_g	1382.4	1299.0	1316.0	
CH ₂ sciss		A_g	1436.0	1423.0	1441.0	
C=C ν		A_g	1602.1	1613.0	1614.0	
C=C ν		A_g	1650.0	1613.0	1617.0	
CH ₂ <i>s</i> -str.		A_g	2987.4	(3005.0)	3018.0	
CH str.		A_g	3022.1	(3009.0)	3021.0	
CH str.		A_g	3029.3	(3015.0)	3027.0	
CH str.		A_g	3049.1	(3021.0)	3035.0	
CH ₂ <i>a</i> -str.		A_g	3093.0	(3096.0)	3103.0	
CCC δ		B_u	109.0	96.0	84.0	
CCC δ		B_u	391.6	390.0	377.0	
CCC δ		B_u	582.8	565.0	559.0	
CH ₂ rock		B_u	914.9	(928.0)	929.0	
C–C ν		B_u	1167.1	1139.0	1138.0	
CH rock		B_u	1258.1	1229.0	1245.0	
CH rock		B_u	1336.1	1280.0	1293.0	
CH rock		B_u	1389.7	1303.0	1317.0	
CH ₂ sciss		B_u	1412.2	1405.0	1423.0	
C=C ν		B_u	1582.3	1569.0	1584.0	
C=C ν		B_u	1629.8	1632.0	1634.0	
CH ₂ <i>s</i> -str.		B_u	2986.2	2967.0	3019.0	
CH str.		B_u	3025.3	3009.0	3024.0	
CH str.		B_u	3042.3	(3018.0)	3030.0	
CH str.		B_u	3057.0	3030.0	3037.0	
CH ₂ <i>a</i> -str.		B_u	3093.8	3091.0	3102.0	
rms err. ^a			32.63(30.69)		14.25(14.28)	

^aValues in parentheses for in-plane modes only, which are better described with our cross-terms than the out-of-plane modes.

168 cm^{-1} , respectively). These HF vibrational modes are illustrated in Fig. 3. The FF developed and discussed here is based on the experimental assignment (as usual with the HBFF formalism). However, we also developed a FF based on the HF description of modes (hereafter referred to as the *HF assignment*) for comparison purposes. For Cl_2MoO_2 , experimental results^{28(g)–28(i)} also show uncertainties in the frequencies of some vibrational modes [e.g., for $\rho(\text{MoCl}_2)$ Ref. 28(g) reports 267 cm^{-1} whereas Ref. 28(l) reports 180

cm^{-1}]. In the following optimizations, we use the frequency values in Ref. 28(g) (the values not in parentheses).

B. Optimization

The MSX FF has standard diagonal terms plus $K_{\theta,R}$ and $K_{R,R}$ cross terms for all three-body interactions.^{4,9} With HBFF/SVD, we first carried out *Hessian optimization* (fitting to the experimental structure and Hessian-vibrational fre-

TABLE XII. Comparison of predicted and experimental frequencies (in cm^{-1}) for benzene. RPIs are also included.

Symmetry	C_6H_6			C_6D_6		
	HBFF/SVD	Expt. (Ref. 25)	RPI	HBFF/SVD	Expt. (Ref. 25)	
Out-of-plane						
E_{2u}	417.5	398.0	0.35	375.9	345.0	
B_{2g}	631.0	707.0	0.73	576.0	599.0	
A_{2u}	679.7	674.0	0.83	498.9	496.2	
E_{1g}	835.8	847.0	0.36	649.8	660.0	
E_{2u}	984.8	967.0	0.25	773.5	787.0	
B_{2g}	1003.5	990.0	0.50	777.3	829.0	
In-plane						
E_{2g}	609.9	608.0	0.38	578.3	580.2	
A_{1g}	992.7	993.0	0.84	944.6	945.6	
B_{1u}	1010.3	1010.0	0.94	959.4	970.0	
E_{1u}	1038.4	1038.0	0.30	794.8	814.3	
B_{2u}	1149.0	1149.7	0.89	829.3	823.7	
E_{2g}	1177.6	1177.8	0.41	857.5	867.0	
B_{2u}	1309.9	1309.4	0.96	1283.4	1286.3	
A_{2g}	1349.2	1350.0	0.78	1049.1	1059.0	
E_{1u}	1484.7	1484.0	0.27	1353.7	1335.2	
E_{2g}	1600.5	1601.0	0.44	1550.3	1558.3	
B_{1u}	3053.1	3057.0	0.17	2273.2	2285.0	
E_{2g}	3059.1	3056.7	0.17	2286.6	2272.5	
E_{1u}	3065.1	3064.4	0.17	2279.5	2289.3	
A_{1g}	3071.7	3073.9	0.17	2282.6	2303.4	
rms err.	18.5			18.0		
Force constants						
Bonds						
C–H	R_e	1.0797				
	K_R	729.23				
C–C	R_e	1.3725				
	K_R	885.46				
Angles						
C–C–H	θ_e	118.31				
	K_θ	71.04				
	$K_{\theta,R}$	-51.98				
C–C–C	θ_e	104.61				
	K_θ	79.69				
	$K_{\theta,R_{C-C}}$	3.06				
	$K_{R,R}$	68.01				
Inversion						
C–C–C–H	K_ψ	53.08				
C–C–C–H	K_ψ	26.54				
	$K_{\text{wag/wag}}$	1.00				
Torsion						
	$K_{\pi\text{-twist}}$	24.48				
2-c ang–ang						
H–C–C–H	$K_{2c\text{-angang}}$	-5.27				
C–C–C–H	$K_{2c\text{-angang}}$	-14.00				
C–C–C–C	$K_{2c\text{-angang}}$	-13.43				
2-c bnd–bnd						
H–C–C–H	$K_{\text{C-H/C-H}}$	0.13				
C–C–C–C	$K_{\text{C-C/C-C}}$	-32.77				

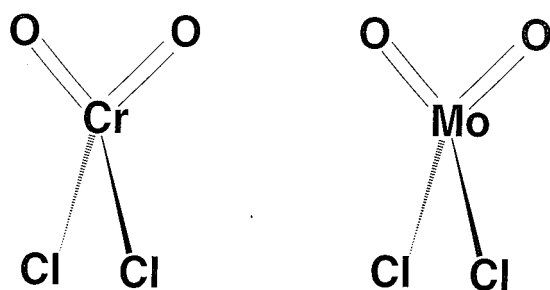
quencies and the theoretical vibrational modes). Then we applied *frequency optimization* to fit experimental frequencies more tightly, leading to the final FF parameters in Table XV. FF parameters were also independently optimized based on the *HF assignment*. In both cases we did a second FF MSX/R, in which several terms were dropped from the MSX FF.

TABLE XIII. Comparison of predicted and experimental frequencies (in cm^{-1}) for naphthalene. Values in parentheses are calculated, not observed.

	Symmetry	HBFF/SVD	Expt. (Ref. 26)
Out-of-plane	B_{1u}	173.8	166.0
	A_u	224.5	(188.0)
	B_{3g}	399.8	385.0
	B_{1u}	404.4	474.0
	B_{2g}	438.6	395.0
	A_u	578.1	(622.0)
	B_{2g}	610.4	772.0
	B_{3g}	738.1	717.0
	B_{1u}	771.8	780.0
	A_u	858.5	(825.0)
	B_{2g}	883.3	875.0
	B_{3g}	952.3	951.0
	B_{1u}	972.0	955.0
	A_u	1008.2	981.0
	In-plane	B_{2g}	1011.9
B_{3u}		377.4	359.0
A_g		486.7	514.0
B_{1g}		521.7	508.0
B_{2u}		649.1	619.0
A_g		802.4	761.0
B_{3u}		803.8	810.0
B_{1g}		951.4	939.0
B_{2u}		1030.1	1008.0
A_g		1069.9	1020.0
B_{1g}		1158.4	1158.0
B_{3u}		1159.5	1125.0
A_g		1190.3	1163.0
B_{2u}		1229.8	1163.0
B_{1g}		1278.6	1240.0
B_{2u}	1285.0	1209.0	
A_g	1330.4	1380.0	
B_{2u}	1341.0	1361.0	
B_{3u}	1388.1	1265.0	
B_{3u}	1479.1	1389.0	
B_{1g}	1485.2	1458.0	
A_g	1504.9	1460.0	
B_{3u}	1574.0	1595.0	
B_{2u}	1587.8	1509.0	
B_{1g}	1643.9	1624.0	
A_g	1734.6	1578.0	
B_{1g}	3054.8	3060.0	
B_{3u}	3057.0	3058.0	
B_{2u}	3059.8	3027.0	
A_g	3063.4	3031.0	
B_{1g}	3064.5	3055.0	
B_{3u}	3068.5	3065.0	
B_{2u}	3070.4	3090.0	
A_g	3072.0	3060.0	
rms err. ^a		50.81(50.44)	

^aValues in parentheses for in-plane modes only, which are better described with our cross-terms than the out-of-plane modes.

Major components in the potential energy distributions (PED)²⁹ are compared in Table XVI for these different sets of FF. The PED shows that the two different assignments in the $\nu_s(\text{MCl})$ and $\delta(\text{MO}_2)$ modes lead to quite different FFs, although both (MSX and MSX/R FF) are optimized to equal quality. Table XV shows that the two assignments affect most the $K_R(\text{M-Cl})$ and $K_{R,R}(\text{Cl-M-Cl})$ force constants. Most diagnostic are $K_{R,R}(\text{Cl-M-Cl})$ and $K_\theta(\text{O=M=O})$ which tend to be more positive for the *experimental assign-*

FIG. 2. The structures of Cl_2CrO_2 and Cl_2MoO_2 .

ment and more negative for the *HF assignment*.

We optimized these FF using the SVD, Simplex, and Powell methods, leading to the results in Table XVII. We see that HBFF/SVD leads to dramatically better performance than either Simplex or Powell.

In order to compare the nature of optimization with different methods, Table XVIII shows the residue components and corresponding *condition ratio* evaluated from the least-squares equation right after each *Hessian optimization* step. Residues corresponding to small *condition ratio* are deleted relatively well with Simplex and Powell methods, but residues with larger *condition ratio* corresponding to the more *difficult* errors are almost intact in these optimizations. This point is verified by computing the residue weighted average *condition ratio* $\langle C \rangle$, also shown in Table XVIII, which starts at 86. Simplex shifts the domain of errors very slightly to

TABLE XIV. Comparison of experimental structures and vibrational frequencies with Hartree-Fock (HF) calculations for Cl_2CrO_2 and Cl_2MoO_2 .

Geometry	Cl_2CrO_2		Cl_2MoO_2		
	Expt. ^a	HF	Expt. ^a	HF	
$R_{\text{M}=\text{O}}$ (Å)	1.57	1.403	1.75	1.522	
$R_{\text{M}-\text{Cl}}$ (Å)	2.12	2.160	2.28	2.307	
$\theta_{\text{O}=\text{M}=\text{O}}$ (deg)	105.0	106.9	109.5	105.3	
$\theta_{\text{Cl}-\text{M}-\text{Cl}}$ (deg)	113.0	126.6	113.0	127.4	
rmsD(Å) ^b	...	0.092	...	0.142	
Vibrational frequencies (cm^{-1})					
sym	mode	Expt. ^c	HF	Expt. ^d (Expt. ^e)	HF
A_1	$\nu_s(\text{M}=\text{O})$	991.0	1206.1	996(996)	1160.6
A_1	$\nu_s(\text{M}-\text{Cl})^f$	470.0	383.7	432(429)	364.2
A_1	$\delta(\text{MO}_2)^f$	355.8	614.9	339(338)	532.0
A_1	$\delta(\text{MCl}_2)$	139.2	134.2	113(167)	120.6
A_2	τ	224	260.6	194(116)	231.7
B_1	$\nu_{as}(\text{M}=\text{O})$	1002.0	971.1	966(970)	1026.6
B_1	$\rho(\text{MO}_2)$	212.0	231.5	211(202)	194.4
B_2	$\nu_{as}(\text{M}-\text{Cl})$	502.7	482.4	470(470)	427.9
B_2	$\rho(\text{MCl}_2)$	257.0	301.2	267(180)	270.3
	$\delta\nu_{\text{rms}}^g(\text{cm}^{-1})$		118.3		92.0

^aReference 27.

^brms error in x, y, z coordinates.

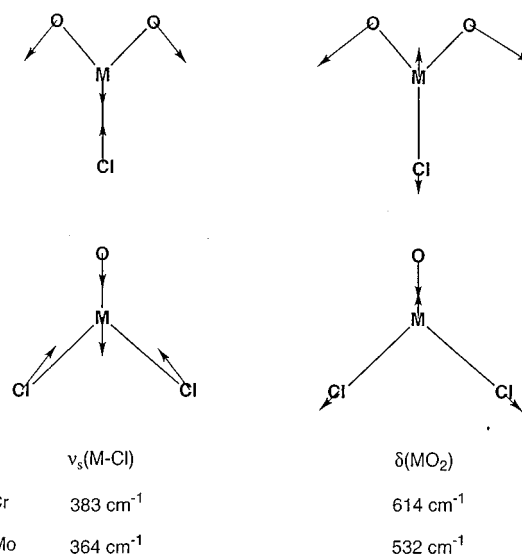
^cReference 28(b).

^dReference 28(g).

^eReferences 28(h) and 28(i).

^fHF and Expt lead to different assignments.

^grms error in frequencies.

FIG. 3. The A_1 vibrational modes for which the HF and experimental assignments disagree.

$\langle C \rangle = 149$, while Powell moves it to intermediate regions of $\langle C \rangle = 449$. In contrast HBFF/SVD with threshold $C = 1000$ deletes most of errors for $C_i < C$. Thus for Simplex and Powell, to obtain an adequate starting set of parameters requires trial and error changes to delete the higher condition number errors. HBFF/SVD does not require such trial and error, providing a much more robust and automatic procedure when parameters are far from the optimum values. Figure 4 shows the convergence behavior for HBFF/SVD optimization of the MSXFF. There is an initial decrease in the total error function values for *Hessian optimization*. It then increases substantially in the early steps of *frequency optimization*. At these steps, HBFF/SVD deletes errors for modes having higher *condition ratios*, leading to drastic changes in some parameter values. This induces additional errors into the smaller condition number modes. Once the higher condition number errors are eliminated, the smaller condition number errors are automatically deleted and optimization converges quickly. Such jumps in the total error value is not allowed in approaches such as Powell or Simplex that strive to minimize the function value, resulting in convergence to local minima near the starting values of parameters.

The *critical condition number* threshold C used with HBFF/SVD is an adjustable parameter in the optimization. In this work, we used the value of 1000, which in our experience is usually best. In order to understand how this parameter affects optimizations, we optimized the MSX FF for Cl_2CrO_2 with various C , obtaining the results in Fig. 5. Here *Hessian optimization* was applied first, converging within five cycles for all cases; then *frequency optimization* was applied. The results show that convergence depends on C . The numbers of least squares equations solved with *frequency optimization* are 5 to 6 for $C = 10$, 13 for $C = 100$, 13 for $C = 1000$, 15 to 17 for $C = 10\,000$, and 17 for $C = 100\,000$. Satisfactory fittings were obtained for the

TABLE XV. Optimized MSX and MSX/R^a FFs for (a) Cl₂CrO₂ and (b) Cl₂MoO₂ molecules. One is based on the *experimental assignment*, the other is based on the *HF assignment* (both use the experimental geometry). Distances in Å, angles in degree, force constants in kcal/mol, Å, radian units.

		Initial	MSX Expt assignment	MSX/R	MSX HF assignment	MSX/R
(a)						
Cr=O	R_e	1.570	1.571	1.570	1.574	1.570
	K_R	1000.00	1023.55	1013.94	1009.11	1014.10
Cr=Cl	R_e	2.120	2.101	2.120	2.115	2.120
	K_R	500.00	427.70	397.67	343.80	338.93
O=Cr=O	θ_e	105.00	107.30	107.72	106.96	106.96
	K_θ	100.00	169.67	172.10	243.35	238.31
	$K_{\theta,R}$	-100.00	-28.07	...	-32.57	...
	$K_{R,R}$	100.00	40.54	40.93	47.22	41.08
Cl-Cr=O	θ_e	109.60	114.36	115.14	115.17	115.14
	K_θ	100.00	90.05	91.07	90.54	91.08
	$K_{\theta,R_{Cl-Cr}}$	-100.00	7.29	...	4.67	...
	$K_{\theta,R_{O-Cr}}$	-100.00	11.57	...	-12.94	...
	$K_{R,R}$	100.00	34.95	...	34.60	...
	θ_e	113.00	121.40	121.79	121.63	121.74
Cl-Cr-Cl	K_θ	100.00	67.94	61.54	62.29	61.88
	$K_{\theta,R_{Cl-Cr}}$	-100.00	55.20	...	3.77	...
	$K_{R,R}$	100.00	66.90	54.20	-10.19	-4.53
	rmsD (Å) ^b	...	0.000	0.000	0.000	0.000
	$\delta\nu_{rms}$ (cm ⁻¹) ^c	...	0.00	1.29	0.00	1.29
(b)						
Mo=O	R_e	1.75	1.754	1.750	1.749	1.750
	K_R	1000.00	1106.97	1116.69	1110.28	1116.69
Mo-Cl	R_e	2.280	2.245	2.280	2.247	2.280
	K_R	500.00	443.88	425.75	370.49	363.02
O=Mo=O	θ_e	109.50	120.69	122.72	117.00	118.43
	K_θ	100.00	200.25	181.04	293.03	268.11
	$K_{\theta,R}$	-100.00	-84.90	...	-75.02	...
	$K_{R,R}$	100.00	67.18	85.34	70.99	85.35
Cl-Mo=O	θ_e	108.60	133.79	134.89	133.81	134.89
	K_θ	100.00	94.16	95.58	94.15	95.59
	$K_{\theta,R_{Cl-Cr}}$	-100.00	30.37	...	31.95	...
	$K_{\theta,R_{O-Cr}}$	-100.00	16.91	...	15.87	...
	$K_{R,R}$	100.00	38.95	...	52.55	...
	θ_e	113.00	159.25	173.97	146.06	174.81
Cl-Mo-Cl	K_θ	100.00	49.59	41.89	67.45	41.32
	$K_{\theta,R_{Cl-Cr}}$	-100.00	-14.51	...	-32.42	...
	$K_{R,R}$	100.00	-8.36	13.47	-83.43	-49.26
	rmsD (Å) ^b	...	0.000	0.000	0.000	0.000
	$\delta\nu_{rms}$ (cm ⁻¹) ^c	...	0.34	2.03	0.32	2.03

^aSome FF terms are excluded from MSX FF.

^brms error in x , y , z coordinates.

^crms error in frequencies.

$C = 100$, 1000 , and $10\,000$. For $C = 10$, additional error residues must be deleted to obtain a good fit to observed properties. For $C = 100\,000$ the least-squares equations attempted to solve for errors to which the FF is insensitive, resulting in oscillations. As the condition number increases, the chance of introducing artifacts into the FF also increases. We have used these procedures for several cases and find that $C = 1000$ generally leads to good optimizations.³⁰

C. Deficiency in force field terms

Table XIX shows the residue components and RPI values after one round of *Hessian optimization*. For Cl₂CrO₂ (or Cl₂MoO₂) this point has $N'_{\text{par}} = 15$ (or 12) equations solved to

the given threshold ($C = 1000$) out of the $N_{\text{par}} = 17$ equations corresponding to the number of properties in the MSXFF problem. This implies that this FF is more redundant and therefore less complete for Cl₂MoO₂ than for Cl₂CrO₂. Reflecting such difficulties, the RPI values are relatively smaller for Cl₂MoO₂ than for Cl₂CrO₂. In Table XIX, the RPI for force components are all unity, indicating that the geometry can be exactly optimized (with appropriately large weights for forces). We observe some deficiencies in the Hessian components for both cases. HF calculations are not accurate enough to judge the magnitude of such observed error components, but the RPI must be close to unity to resolve disagreements between the FF and observed proper-

TABLE XVI. Potential energy distributions (PED)^a of vibrational modes for (a) Cl₂CrO₂ and (b) Cl₂MoO₂.

Vibrations									
Frequency (cm ⁻¹)	Mode	MSX				MSX			
		Experimental assignment		MSX/R		HF assignment		MSX/R	
(a)									
991.0	A ₁ ν _s (Cr=O)	97%	ν(Cr=O)	97%	ν(Cr=O)	96%	ν(Cr=O)	96%	ν(Cr=O)
470.0	...	82%	ν(Cr-Cl) ^b	71%	ν(Cr-Cl) ^b	69%	δ(O=Cr=O) ^c	69%	δ(O=Cr=O) ^c
355.8	...	76%	δ(O=Cr=O) ^c	71%	δ(O=Cr=O) ^c	75%	ν(Cr-Cl) ^b	76%	ν(Cr-Cl) ^b
139.2	A ₁ δ(CrCl ₂)	61%	δ(Cl-Cr-Cl)	68%	δ(Cl-Cr-Cl)	67%	δ(Cl-Cr-Cl)	68%	δ(Cl-Cr-Cl)
224	A ₂ τ	97%	δ(Cl-Cr=O)	97%	δ(Cl-Cr=O)	94%	δ(Cl-Cr=O)	97%	δ(Cl-Cr=O)
1002.0	B ₁ ν _{as} (Cr=O)	99%	ν(Cr=O)	99%	ν(Cr=O)	99%	ν(Cr=O)	99%	ν(Cr=O)
212.0	B ₁ ρ(CrO ₂)	93%	δ(Cl-Cr=O)	96%	δ(Cl-Cr=O)	96%	δ(Cl-Cr=O)	96%	δ(Cl-Cr=O)
502.7	B ₂ ν _{as} (Cr-Cl)	88%	ν(Cr-Cl)	86%	ν(Cr-Cl)	85%	ν(Cr-Cl)	84%	ν(Cr-Cl)
257.0	B ₂ ρ(CrCl ₂)	83%	δ(Cl-Cr=O)	77%	δ(Cl-Cr=O)	80%	δ(Cl-Cr=O)	79%	δ(Cl-Cr=O)
(b)									
996	A ₁ ν _s (Mo=O)	99%	ν(Mo=O)	99%	ν(Mo=O)	98%	ν(Mo=O)	99%	ν(Mo=O)
432	...	86%	ν(Mo-Cl) ^d	82%	ν(Mo-Cl) ^d	72%	δ(O=Mo=O) ^e	72%	δ(O=Mo=O) ^e
339	...	72%	δ(O=Mo=O) ^e	71%	δ(O=Mo=O) ^e	82%	ν(Mo-Cl) ^d	86%	ν(Mo-Cl) ^d
113	A ₁ δ(MoCl ₂)	48%	δ(Cl-Mo-Cl)	55%	δ(Cl-Mo-Cl)	49%	δ(Cl-Mo-Cl)	55%	δ(Cl-Mo-Cl)
194	A ₂ τ	85%	δ(Cl-Mo=O)	88%	δ(Cl-Mo=O)	88%	δ(Cl-Mo=O)	88%	δ(Cl-Mo=O)
966	B ₁ ν _{as} (Mo=O)	100%	ν(Mo=O)	100%	ν(Mo=O)	100%	ν(Mo=O)	100%	ν(Mo=O)
211	B ₁ ρ(MoO ₂)	85%	δ(Cl-Mo=O)	90%	δ(Cl-Mo=O)	88%	δ(Cl-Mo=O)	90%	δ(Cl-Mo=O)
470	B ₂ ν _{as} (Mo-Cl)	96%	ν(Mo-Cl)	91%	ν(Mo-Cl)	95%	ν(Mo-Cl)	90%	ν(Mo-Cl)
267	B ₂ ρ(MoCl ₂)	76%	δ(Cl-Mo=O)	70%	δ(Cl-Mo=O)	78%	δ(Cl-Mo=O)	71%	δ(Cl-Mo=O)

^aThe PED for each vibrational mode is calculated with diagonal FF terms only and values are scaled so that absolute values of contributions sum up to 100%.

^bMode assigned as A₁ ν_s(M-Cl).

^cMode assigned as A₁ δ(MO₂).

^dMode assigned as A₁ ν_s(Mo-Cl).

^eMode assigned as A₁ δ(MoO₂).

ties. For example, the δ(MO₂) and ν_s(M-Cl) components show the largest error values of 4.82 and -3.71 and the RPI are relatively small, 0.29 for Cl₂CrO₂ and 0.27 for Cl₂MoO₂. This implies that a δ(MO₂)-ν_s(M-Cl) cross term is needed in the FF to fit this off-diagonal Hessian component.

In larger systems, the interpretation of the RPI may not be as straightforward as in above examples, because the index is represented by normal coordinates. Determining which FF terms are responsible for small RPI values may require transforming from normal coordinates to local va-

TABLE XVII. Convergence for optimization of the MSX FF for Cl₂CrO₂ and Cl₂MoO₂ molecules.

	Cl ₂ CrO ₂			Cl ₂ MoO ₂		
	Simplex	Powell	HBFF/SVD	Simplex	Powell	HBFF/SVD
Hessian optimization						
Iterations ^a	272	13	5	229	13	5
ERR evaluations ^b	405	1202	91	355	1278	91
ERR _{initial}	18.6014	18.6014	18.6014	13.0369	13.0369	13.0369
ERR _{final}	6.6459	2.2403	0.9849	4.5669	1.3518	0.6780
δν _{rms} ^c	119.81	80.17	16.69	62.85	24.50	19.69
δforce _{rms} ^d	0.0029	0.0001	0.0000	0.0018	0.0002	0.0000
Frequency optimization						
Iterations	323	14	5	480	6	6
ERR evaluations	419	1396	91	609	518	109
ERR _{initial} ^e	40.2816	26.7586	5.5764	21.1693	8.1983	6.5720
ERR _{final}	11.0916	3.5098	0.0161	13.6936	5.8305	0.2741
δν _{rms}	31.37	7.80	0.00	39.16	17.34	0.34
δforce _{rms}	0.0059	0.0082	0.0000	0.0060	0.0003	0.0016

^aNumber of iterations required to converge to |ERR^(k+1) - ERR^(k)| < 0.01.

^bTotal number of error residue evaluations. This corresponds to computational cost.

^crms frequency error (cm⁻¹) between experimental and FF frequencies.

^drms residual force in kcal mol⁻¹ Å⁻¹.

^eInitial ERR is evaluation in frequency optimization scheme with parameters from previous Hessian optimization step.

TABLE XVIII. Error residues and *condition ratios* for the Simplex, Powell, and HBFF/SVD methods. Results are for MSX FF of Cl₂CrO₂ after one round of optimization.^a

Initial FF		Simplex		Powell		SVD	
C_i	b_i	C_i	b_i	C_i	b_i	C_i	b_i
1	0.0521	1	0.0137	1	-0.0019	1	0.0020
1	0.7779	1	0.0021	1	0.0008	1	0.0006
4	0.1738	4	0.0044	3	0.0001	2	0.0003
10	-0.0221	16	0.0273	13	-0.0010	3	0.0000
34	-2.2569	34	-0.2554	35	-0.0170	35	0.0000
35	0.6202	49	-0.6783	39	0.0950	37	0.0000
70	0.2416	62	0.1258	92	-0.1327	104	0.0000
91	1.2061	90	-0.9965	113	-0.0808	116	0.0000
115	0.3344	108	-0.8396	132	-0.0515	134	0.0000
130	-0.1716	116	0.1356	171	0.0370	198	0.0000
138	-0.1223	129	0.1915	184	0.3445	211	0.0000
167	1.0668	148	1.1273	210	-0.1983	267	0.0000
174	0.0083	158	0.4683	258	0.1164	435	0.0001
249	1.0082	224	1.2029	304	-0.8697	568	-0.0001
315	0.5381	272	0.3744	332	0.5105	677	-0.0003
600	0.0343	556	0.0554	6094	0.0578	1040	0.0003
14207	-0.0207	35205	0.0202	10342	0.1343	2625	-0.0216
$\langle C \rangle^b$	86	149		449		2599	

^aHessian optimization using Eqs. (4) and (5).^bResidue weighted mean *condition ratio*, $\langle C \rangle = \sum_i C_i \cdot b_i^2 / \sum_j b_j^2$.

lence coordinates (e.g., bond-stretch, angle-bend, torsion, and so on).³¹ Instead we used a PED analysis to associate particular FF parameters with each vibrational mode. In the above example, the normal vibrational modes $\nu_s(\text{M}-\text{Cl})$ and $\delta(\text{MO}_2)$ are determined mostly by $K_{R(\text{R}-\text{Cl})}$ and $K_{\theta(\text{O}=\text{M}=\text{O})}$, respectively.

D. Balance in experimental and theoretical constraints

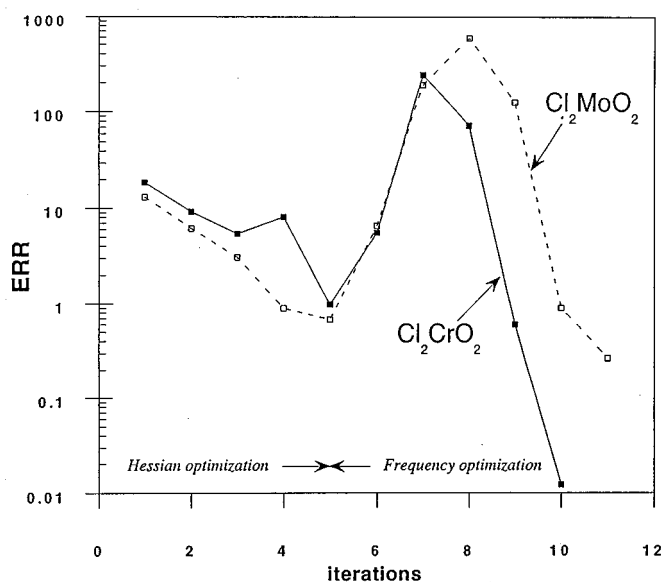
The MSXFF is not adequate to satisfy all of the constraints imposed in the *Hessian optimization*. The optimiza-

tions resulted in rms frequency errors of 16.69 and 19.69 cm^{-1} for Cl₂CrO₂ and Cl₂MoO₂ respectively while force residues are reduced satisfactorily for both cases. For this reason, we switched the optimization to *frequency optimization* with reduced weights ($\bar{W}_{H_{\text{off-diag}}} = 0.1$) for *ab initio* vibrational mode constraints. This led to excellent fits in both rms frequency errors and forces, Table XVII.

Although *frequency optimization* emphasizes experimental data (forces and frequencies), the *ab initio* vibrational mode information still affects the course of optimization. For example, when very small weights are used for the vibrational mode constraints, we found that the optimization sometimes oscillates. Thus we recommend using the vibrational modes of HBFF even for cases when only structure and frequency information is available. This can be done by using a trial FF to generate vibrational modes.

E. Redundancy and inactive force field terms

Use of PSI information is illustrated in Table XX(a) for optimizing the MSXX FF of formaldehyde.⁹ In this example, only the geometric constant parameters are optimized (without geometric constraints). We expected redundancies related to the two angle terms (H-C-H and O-C-H) and the two inversion terms. Indeed the PSI for MSXX shows such redundancies, where the PSI of two K_θ , two K_ψ , and two $K_{\theta,\theta}$ each sum up to nearly unity. We then reoptimized various modified MSXX/R FF with one of the force constants excluded. In each case this leads to a significant increase in the PSI for the remaining parameter. For MSXX/R (K_θ) and MSXX/R ($K_{\text{Cl,Cl}}$) we also observed some coupling in force constants. Thus the PSI values of three-body bond-bond cross terms and the four-body angle-angle cross terms also changed.

FIG. 4. Convergence of HBFF/SVD for Cl₂CrO₂ and Cl₂MoO₂. ($C = 1000$).

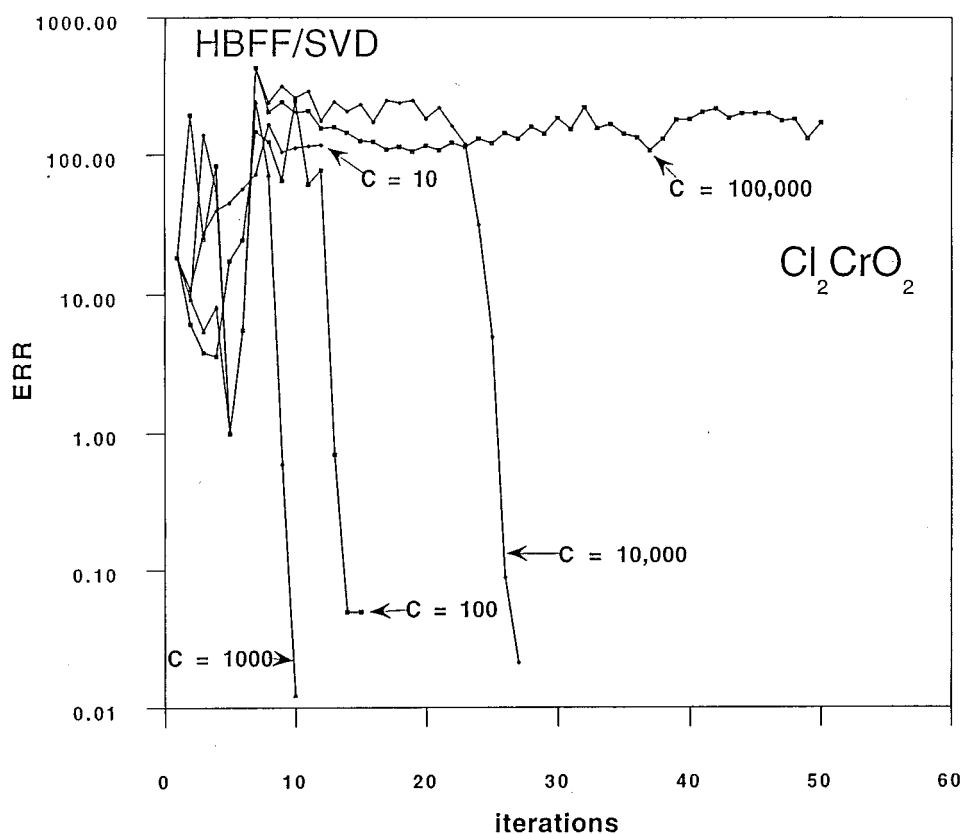


FIG. 5. Convergence of HBFF/SVD for Cl_2CrO_2 as a function of *critical condition number* C .

Similar situations were observed in the optimizations of the MSX FF for Cl_2CrO_2 and Cl_2MoO_2 . At the end of *frequency optimization* with MSX FF, there are 13 (or 12) equations solved out of 17 for Cl_2CrO_2 (or Cl_2MoO_2) with the

given weights and threshold ($C=1000$). This indicates how many parameters are active in the optimization, suggesting redundancy or inactiveness in the MSX FF set. Such redundant or inactive parameters in a FF will sometimes be nec-

TABLE XIX. Residue components and RPI of MSX FF for Cl_2CrO_2 and Cl_2MoO_2 at the end of the *Hessian optimization* step. The RPI are in parentheses.

	CrO_2Cl_2				MoO_2Cl_2			
Force residue components (A_1 modes only)								
A_1 symmetry								
$\nu_s(\text{M}=\text{O})$	-0.0001(1.00)				-0.0001(1.00)			
$\nu_s(\text{M}-\text{Cl})$	-0.0004(1.00)					0.0001(1.00)		
$\delta(\text{MO}_2)$	0.0004(1.00)				0.0000(1.00)			
$\delta(\text{MCl}_2)$	-0.0004(1.00)				0.0007(1.00)			
Hessian residue components								
A_1 symmetry								
$\nu_s(\text{M}=\text{O})$	-0.07(0.89)				-0.25(0.00)			
$\nu_s(\text{M}-\text{Cl})$	1.20(0.16)	-1.32(0.95)			0.77(0.01)	-1.27(0.92)		
$\delta(\text{MO}_2)$	-1.28(0.15)	4.82(0.29)	3.01(0.77)		0.56(0.00)	-3.71(0.27)	2.13(0.81)	
$\delta(\text{MCl}_2)$	0.51(0.16)	0.00(0.99)	-0.70(0.99)	0.00(1.00)	-0.39(0.00)	0.01(1.00)	0.65(0.98)	0.02(1.00)
A_2 symmetry								
τ	-1.07(0.75)				-1.91(0.37)			
B_1 symmetry								
$\nu_{as}(\text{M}=\text{O})$	-0.16(0.40)		$\rho(\text{MO}_2)$		0.30(0.17)		$\rho(\text{MO}_2)$	
$\rho(\text{MO}_2)$	0.18(0.87)	0.01(1.00)			-0.12(0.01)	0.00(1.00)		
B_2 symmetry								
$\nu_{as}(\text{M}-\text{Cl})$	-1.23(0.32)		$\rho(\text{MCl}_2)$		0.01(0.26)		$\rho(\text{MCl}_2)$	
$\rho(\text{MCl}_2)$	-1.77(0.35)	0.56(0.97)			-2.18(0.22)	0.64(0.97)		

TABLE XX. (a) The PSI for the MSXX^a and MSXX/R FF of formaldehyde. Force constants were fitted using *Hessian optimization* scheme. In each case of MSXX/R, a single term is deleted. (b) PSI for the MSX and MSX/R FF of Cl₂CrO₂ and Cl₂MoO₂. PSIs are shown for both *experimental assignment* and *HF assignment* (in parentheses).

	MSXX	MSXX/R	(K_θ)	MSXX/R	(K_ψ)	MSXX/R	($K_{\theta,\theta}$)
(a)							
Stretch							
K_{CH}	1.0000	1.0000	0.9998	1.0000	1.0000	0.9993	0.9991
K_{CO}	1.0000	1.0000	0.9978	1.0000	1.0000	0.9950	0.7297
H-C-H							
K_θ	0.3798	...	0.9969	0.3806	0.3806	0.9978	0.7844
$K_{\theta,R}$	1.0000	1.0000	0.8385	1.0000	1.0000	0.8427	0.9885
$K_{R,R}$	1.0000	1.0000	0.2797	1.0000	1.0000	0.2998	0.9974
O=C-H							
K_θ	0.6180	0.6946	...	0.6185	0.6185	0.9910	0.9599
$K_{\theta,R}$	1.0000	1.0000	0.9945	1.0000	1.0000	0.9982	0.6830
$K_{\theta,R}$	1.0000	1.0000	0.9382	1.0000	1.0000	0.9378	0.9932
$K_{R,R}$	1.0000	1.0000	0.9619	1.0000	1.0000	0.9407	0.9835
Inversion							
$K_{\psi_{C-O-H-H_C}}$ ^b	0.4991	0.4998	0.5000	...	0.9965	0.5000	0.4955
$K_{\psi_{C-H-H-O}}$	0.4991	0.4998	0.5000	0.9965	...	0.5000	0.4955
Angle-angle							
$K_{\theta_{O-C-H},\theta_{H-C-H}}$	0.6239	0.6938	0.9956	0.6244	0.6244	0.9977	...
$K_{\theta_{O-C-H},\theta_{O-C-H}}$	0.3800	0.6120	0.9970	0.3800	0.3800	...	0.8901
(b)							
		CrO ₂ Cl ₂		MoO ₂ Cl ₂			
		MSX	MSX/R	MSX	MSX/R		
M=O							
R_e		0.97(0.97)	1.00(1.00)	0.50(0.55)	1.00(1.00)		
K_R		0.93(0.92)	1.00(1.00)	0.94(0.93)	1.00(1.00)		
M-Cl							
R_e		0.90(1.00)	1.00(1.00)	0.53(0.39)	1.00(1.00)		
K_R		0.91(0.99)	1.00(1.00)	0.97(0.95)	1.00(1.00)		
O=M=O							
θ_e		1.00(1.00)	1.00(1.00)	0.96(0.98)	1.00(1.00)		
K_θ		0.96(0.96)	1.00(1.00)	0.80(0.88)	1.00(1.00)		
$K_{\theta,R}$		0.12(0.16)	1.00(1.00)	0.52(0.31)	1.00(1.00)		
$K_{R,R}$		0.84(0.88)	1.00(1.00)	0.90(0.91)	1.00(1.00)		
Cl-M=O							
θ_e		1.00(1.00)	1.00(1.00)	1.00(0.99)	1.00(1.00)		
K_θ		1.00(1.00)	1.00(1.00)	1.00(1.00)	1.00(1.00)		
$K_{\theta,R_{Cl-M}}$		1.00(0.87)	1.00(1.00)	0.68(0.72)	1.00(1.00)		
$K_{\theta,R_{O=M}}$		0.04(0.25)	1.00(1.00)	0.35(0.30)	1.00(1.00)		
$K_{R,R}$		0.14(0.38)	1.00(1.00)	0.13(0.15)	1.00(1.00)		
Cl-M-Cl							
θ_e		1.00(1.00)	1.00(1.00)	0.98(0.72)	1.00(1.00)		
K_θ		0.92(1.00)	1.00(1.00)	0.99(0.79)	1.00(1.00)		
$K_{\theta,R}$		0.37(0.00)	1.00(1.00)	0.36(0.50)	1.00(1.00)		
$K_{R,R}$		0.91(0.64)	1.00(1.00)	0.40(0.92)	1.00(1.00)		
Errors							
rmsD(Å) ^d		0.000(0.000)	0.000(0.000)	0.000(0.000)	0.000(0.000)		
$\delta\nu_{rms}$ (cm ⁻¹) ^e		0.00(0.00)	1.29(1.29)	0.34(0.32)	2.03(2.03)		

^aReference 9.

^bInversion is defined with the angle between C-H bond and O-C-H plane.

^cInversion is defined with the angle between C-O bond and H-C-H plane.

^drms error in *x*, *y*, *z* coordinates.

^erms error in frequencies.

essary in using a general form for the FF.^{32,33} However, FF terms that are completely inactive throughout the optimization should probably be excluded from the FF to enhance the efficiency of energy calculations and to help with transferability.

With a view to reducing such redundant parameters in the MSX FF for Cl₂CrO₂ and Cl₂MoO₂ molecules, we exam-

ined the PSI of MSX FF at the end of the *frequency optimization* step. These PSI values are shown in Table XX(b) for both Cl₂CrO₂ and Cl₂MoO₂. For Cl₂CrO₂, the FF parameters with small PSI are $K_{\theta,R}$ (PSI) ≤ 0.37, except $K_{\theta,R_{Cl-Cr}}$ and $K_{R,R}$ for Cl-Cr=O interaction (PSI=0.14, 0.38). The Cl₂MoO₂ case seems rather tricky as PSI shows small values

of 0.50 (0.55) and 0.53 (0.39) for the fundamental geometric parameters of $R_e(\text{Mo}=\text{O})$ and $R_e(\text{Mo}-\text{Cl})$, respectively, implying these parameters are not very active in this optimization. This is due to the difficulties in satisfying all observational constraints for Cl_2MoO_2 —the number of least-squares equations solved is only 12 for optimizing 17 FF parameters. The $K_{\theta,R}$ terms and $K_{R,R}(\text{Cl}-\text{Mo}=\text{O})$ also show small PSI values. In developing MSX/R, we tried to drop all $K_{\theta,R}$ and $K_{R,R}(\text{Cl}-\text{M}=\text{O})$ ($\text{M}=\text{Cr}, \text{Mo}$), keeping the number of parameters at 12. With this MSX/R, all PSI are exactly unity and reoptimization led to the FF in Table XV, which fits well both structure and vibrational frequencies. The PED in Table XVI also shows that MSX/R is quite similar to MSX. However the MSX/R FF of Cl_2MoO_2 , has a very large (almost linear) equilibrium angle $\theta_e(\text{Cl}-\text{Mo}-\text{Cl})$. This may be due to our choice for the specific terms in the FF—in this example, there were several choices of terms for MSX/R. The MSX FF also has similar tendency for the Cl_2MoO_2 case. We also tried to drop all three-body cross terms except $K_{R,R}(\text{O}=\text{M}=\text{O})$ and $K_{\theta,R_{\text{Cl-M}}}(\text{Cl}-\text{M}=\text{O})$ and optimized the FF. In this case we could not obtain a good fit for Cl_2MoO_2 using *HF assignment*. This provides evidence in favor of the experimental assignment and shows how PSI are useful in sorting out the important FF terms.

VII. CALCULATIONAL DETAILS

Experimental structure data was taken from the equilibrium geometry for formaldehyde³⁴ and from Sutton's table²⁷ for chromium and molybdenum–dioxo-dichlorides. For the other organic molecules, they were obtained from the references for the experimental frequencies as listed below.

Ab initio calculations were carried out at the restricted Hartree–Fock level using the GAUSSIAN-90 program.³⁵ For the organic molecules, either the D95V** or 6-31G** basis sets³⁶ were used. For Cl_2CrO_2 and Cl_2MoO_2 we used the D95V* basis set for O and the Hay–Wadt effective core potentials³⁷ for Cr, Mo, and Cl. Geometries were optimized and all the vibrational frequencies and modes were calculated at the energy minimum.

Experimental frequencies used were taken from the following: formaldehyde,³⁸ ethylene,¹⁵ butadiene,¹² hexatriene,^{18,19} octatetraene,²³ benzene,²⁵ naphthalene,²⁶ Cl_2CrO_2 ,^{28(b)} and Cl_2MoO_2 .^{28(g),28(i)}

The HBFF/SVD optimization was implemented in the FFOPT program^{10(a)} which was used with POLYGRAF.^{10(b)} The force field terms used are

$$E = E_b + E_a + E_t + E_i + E_x + E_{\text{vdw}} + E_Q, \quad (28)$$

Bond stretching is described by the harmonic potential,

$$E_b = \frac{1}{2}K_b[r - r_e]^2, \quad (29)$$

where r is the bond length, r_e is the equilibrium bond length, and K_b is the force constant. Angle bending is described by the harmonic cosine expression,

$$E_a = \frac{1}{2}C_a[\cos \theta - \cos \theta_e]^2, \quad (30)$$

where θ is the angle between bonds IJ and JK , θ_e is the equilibrium angle, and $K_\theta = C_a \sin^2 \theta_e$ is the diagonal force constant.

In addition to the pi-twist form of the torsional energy as described in Sec. V A 1, a Fourier series is often used

$$E_t = \sum_{m=0}^{12} V_m \cos m\phi. \quad (31)$$

For a given $J-K$ dihedral about a single bond, there are several possible $I-J-K-L$ terms. Each torsional term is scaled by the number of torsions about the bond so that the torsional barriers V_m represent the full torsional barrier about each bond.

Given an atom I bonded to exactly three other atoms J , K , and L , the angle term (30) for JIK , JIL , and KIL will generally lead to a double well for umbrella motion (as in NH_3). However we can include an explicit inversion term to better describe the energy associated with the umbrella motion. We use the harmonic cosine expression

$$E_i = \frac{1}{2}C_i(\cos \psi - \cos \psi_e)^2, \quad (32)$$

where ψ is the angle between the IL bond and the IJK plane and ψ_e is the equilibrium value ($\psi_e = 0$ corresponds to the planar configuration with the longest distance from L to J and K). Here the force constant is $K_\psi = C_i \sin^2 \psi_i$ and the barrier to planarization is

$$V_{\text{bar}}^{\text{inv}} = \frac{1}{2}C_i(1 - \cos \psi_e)^2. \quad (33)$$

There are three possible choices for L . To remove any biases we sum over all three and multiply by 1/3. Additionally, there is a coupling term between inversions at adjacent pi-twist centers as described in Sec. V A 1.

To obtain an accurate description of the vibrational levels generally requires couplings (cross terms) between various bonds and angles. For each angle term (30) we use the bond–angle and bond–bond terms

$$E_{ax} = D_1(\cos \theta - \cos \theta_e)(r_1 - r_{e1}) + D_2(\cos \theta - \cos \theta_e) \times (r_2 - r_{e2}) + K_{rr}(r_1 - r_{e1})(r_2 - r_{e2}), \quad (34)$$

where r_1 and r_2 are the lengths of the IJ and JK bonds, $K_{r\theta} = -D \sin \theta_e$ is the angle–stretch force constant, and K_{rr} is the stretch–stretch force constant. In addition, we considered one-center angle–angle cross terms of the form

$$E_{1aa} = G(\cos \theta_{IJK} - \cos \theta_{eIJK})(\cos \theta_{IJL} - \cos \theta_{eIJL}), \quad (35)$$

where $K_{1\theta\theta} = G \sin \theta_{eIJK} \sin \theta_{eIJL}$ is the force constant for two angle terms (IJK and IJL) sharing a common central bond (IJ) and a common central atom (J). However, we found that these terms do not significantly improve the results, and we eschew them. Also we included two-center angle–angle terms described by

$$E_{2aa} = F \cos \phi (\cos \theta_{IJK} - \cos \theta_{eIJK}) \times (\cos \theta_{JKL} - \cos \theta_{eJKL}), \quad (36)$$

where $K_{2\theta\theta} = F \sin \theta_{eIJK} \sin \theta_{eJKL}$ is the force constant for angle terms (IJK and JKL) in which the central atoms (J and K) are bonded to each other. These cross terms are considered collectively as

$$E_x = E_{ax} + E_{1aa} + E_{2aa}. \quad (37)$$

The van der Waals part of the nonbond interaction for atoms I and J are described using the exponential-6 potential

$$E_{\text{vdw}} = D_v \left[\left(\frac{6}{\zeta - 6} \right) e^{\zeta(1-\rho)} - \left(\frac{\zeta}{\zeta - 6} \right) \rho^{-6} \right], \quad (38)$$

where $\rho = R_{IJ}/R_v$. Here D_v is the well depth, R_v is the distance at the minimum and ζ is a dimensionless constant related to the stiffness of the inner wall. The electrostatic interactions (E_Q) are described using the Coulomb expression

$$E_Q = \frac{Q_I Q_J}{\epsilon_0 \epsilon R_{IJ}}, \quad (39)$$

where Q_I is the charge (electron units) on center I , $\epsilon = 1$, and $1/\epsilon_0 = 332.0637$ gives E in kcal/mol when distances are in Å.

VIII. CONCLUSION

The HBFF/SVD optimization technique utilizes information both from experiment and *ab initio* QC calculations. The application of HBFF/SVD to typical problems in organic and transition metal compounds leads to reasonably fast convergence, allowing interactive development of the FF. The RPI and PSI indices help to identify deficiencies and redundancies in the FF, allowing systematic and efficient development of the FF.

ACKNOWLEDGMENTS

We thank Asahi Chemical for allowing one of the coauthors (T.E.Y.) to have an extended stay at the MSC. This work was supported by DOE-BCTR. The facilities of the MSC are supported by grants from NSF-GCAG (ACS-92-17368), NSF-Chemistry (94-13930), Allied-Signal, Asahi Chemical, Asahi Glass, BP Chemical, Chevron Petroleum Technology, BF Goodrich, Xerox, and Beckman Institute.

¹(a) J. A. McCammon and S. C. Harvey, *Dynamics of Proteins and Nucleic Acids* (Cambridge U.P., Cambridge, 1987); (b) *Computer Simulation of Biomolecular Systems*, edited by W. F. Gunsteren and P. K. Weiner (ESCOM, Leiden, 1989).

²(a) N. Karasawa, S. Dasgupta, and W. A. Goddard III, *J. Phys. Chem.* **95**, 2260 (1991); (b) C. B. Musgrave, S. Dasgupta, and W. A. Goddard III, *ibid.* **99**, 13321 (1995).

³N. Karasawa and W. A. Goddard III, *Macromolecules* **25**, 7268 (1992).

⁴J. A. Wendel and W. A. Goddard III, *J. Chem. Phys.* **97**, 5048 (1992).

⁵(a) M. Li and W. A. Goddard III, *Phys. Rev. B* **40**, 12155 (1989); (b) M. Li and W. A. Goddard III, *J. Chem. Phys.* **98**, 7995 (1993).

⁶S. Lifson and A. Warshel, *J. Chem. Phys.* **49**, 5116 (1968).

⁷K. Rasmussen, *Lecture Notes in Chemistry* (Springer-Verlag, Heidelberg, 1985), Vol. 37.

⁸N. L. Allinger, K. Chen, M. Rahman, and A. Pathiaseril, *J. Am. Chem. Soc.* **113**, 4505 (1991), and references therein.

⁹S. Dasgupta and W. A. Goddard III, *J. Chem. Phys.* **90**, 7207 (1989).

¹⁰(a) FFOPT (unpublished) was written by W. A. Goddard III, S. Dasgupta, and T. E. Yamasaki; (b) It is used in conjunction with the Vibrate module of Pro-POLYGRAF from Molecular Simulations, Inc., Burlington, MA 02154.

¹¹Residue components in the Hessian optimization might be approximated by assuming that the vibrational wave function is unchanged for small changes in the force constants (see Ref. 6). However we computed these residues by finite difference of each parameter without any such approximation.

¹²P. Pulay, G. Fogarasi, G. Pongor, J. E. Boggs, and A. Vargha, *J. Am. Chem. Soc.* **105**, 7037 (1983).

¹³(a) G. Strang, *Linear Algebra and its Applications* (Academic, New York, 1980); (b) W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes* (Cambridge U.P., Cambridge, 1989).

¹⁴S. L. Mayo, B. D. Olafson, and W. A. Goddard III, *J. Phys. Chem.* **94**, 8897 (1990).

¹⁵J. L. Duncan and E. Hamilton, *Theochem.* **76**, 65 (1981).

¹⁶(a) A. K. Rappé, C. J. Casewit, K. S. Colwell, W. A. Goddard III, and W. M. Skiff, *J. Am. Chem. Soc.* **114**, 10024 (1992); (b) J. H. Lii and N. L. Allinger, *ibid.* **111**, 8576 (1989).

¹⁷P. Pulay, G. Fogarasi, G. Pongor, J. E. Boggs, and A. Vargha, *J. Am. Chem. Soc.* **105**, 7037 (1983).

¹⁸G. Fogarasi, P. G. Szalay, P. P. Liescheski, J. E. Boggs, and P. Pulay, *J. Mol. Struct. (Theochem)* **151**, 341 (1987).

¹⁹F. W. Langkilde, R. Wilbrandt, and A. M. Brouwer, *J. Phys. Chem.* **94**, 4809 (1990).

²⁰(a) R. J. Hemley, B. R. Brooks, and M. Karplus, *J. Chem. Phys.* **85**, 6550 (1986); (b) C. W. Bock, Y. N. Panchenko, S. V. Krasnoschiokov, and V. I. Pupyshv, *J. Mol. Struct.* **148**, 131 (1986); (c) P. G. Szalay, A. Karpfen, and H. Lischka, *J. Chem. Phys.* **87**, 3530 (1987).

²¹M. Traetteberg, *Acta. Chem. Scand.* **22**, 628 (1968).

²²R. G. Parr, *Quantum Theory of Molecular Electronic Structure* (Benjamin, New York, 1964).

²³(a) H. Yoshida and M. Tasumi, *J. Chem. Phys.* **89**, 2803 (1988); (b) T. P. Hamilton and P. Pulay, *J. Phys. Chem.* **93**, 2341 (1989).

²⁴R. H. Baughman, B. E. Kohler, I. J. Levy, and C. Spangler, *Synth. Metals* **11**, 37 (1985).

²⁵L. Goodman, A. G. Ozkabak, and S. N. Thakur, *J. Phys. Chem.* **95**, 9044 (1991).

²⁶H. Sellers, P. Pulay, and J. E. Boggs, *J. Am. Chem. Soc.* **107**, 6487 (1985).

²⁷*Tables of Interatomic Distances and Configuration in Molecules and Ions*, edited by L. E. Sutton (The Chemical Society, London, 1958).

²⁸(a) B. G. Ward and F. E. Stafford, *Inorg. Chem.* **7**, 2569 (1968); (b) E. L. Varetto and A. Müller, *Spectrochimica Acta.* **34A**, 895 (1978); (c) C. J. Marsden, L. Hedberg, and K. Hedberg, *Inorg. Chem.* **21**, 1115 (1982); (d) M. Cieślak-Golonka, *Coordinat. Chem. Rev.* **109**, 223 (1991); (e) T. V. Iorns and F. E. Stafford, *J. Am. Chem. Soc.* **88**, 4819 (1966); (f) D. M. Adams and R. G. Churchill, *J. Chem. Soc. A*, 2310 (1968); (g) I. R. Beattie, K. M. S. Livingston, D. J. Reynolds, and G. A. Ozin, *ibid.* 1210 (1970); (h) W. Levason, R. Narayanaswamy, J. S. Ogden, A. J. Rest, and J. W. Turff, *J. Chem. Soc. Dalton Trans.* 2009 (1982); (i) D. L. Neikirk, J. C. Fagerli, M. L. Smith, D. Mosman, and T. C. Devore, *J. Mol. Struct.* **244**, 165 (1991); (j) A. K. Rappé and W. A. Goddard III, in *Potential Energy Surfaces and Dynamics Calculations*, edited by D. G. Truhlar (Plenum, New York, 1981), pp. 661–684; (k) A. K. Rappé and W. A. Goddard III, *J. Am. Chem. Soc.* **102**, 5114 (1980); (l) V. M. Kouba and A. A. Mal'tsev, *Russ. J. Inorg. Chem.* **21**, 2569 (1968).

²⁹For example, see: H. Fuhrer, V. B. Kartha, K. G. Kidd, P. J. Krueger, and H. H. Mantsch, *Normal Coordinate Analysis* (N.R.C.C. Bulletin No. 15, 1976).

³⁰We may write the least-squares equation (2a) as, $\mathbf{G} + \mathbf{b} = \mathbf{G} + \mathbf{G}\Delta\mathbf{p} = \Delta\mathbf{p}$ and solve by diagonalization of $\mathbf{G}^+ \mathbf{G}$, instead of using the SVD of \mathbf{G} . However, in this approach the condition ratio is also squared making the control of the numerical solution quite difficult.

- ³¹ K. Palmö, L.-O. Pietilä, and S. Krimm, *J. Comput. Chem.* **12**, 385 (1991).
- ³² P. Pulay, G. Fogarasi, F. Pang, and J. E. Boggs, *J. Am. Chem. Soc.* **101**, 2550 (1979).
- ³³ L.-O. Pietilä, *J. Mol. Struct.* **195**, 111 (1989).
- ³⁴ K. Yamada, T. Nakagawa, K. Kuchitsu, and Y. Morino, *J. Mol. Spectrosc.* **38**, 70 (1971).
- ³⁵ GAUSSIAN 90, Revision F, M. J. Frisch, M. Head-Gordon, G. W. Trucks, J. B. Foresman, H. B. Schlegel, K. Raghavachari, M. Robb, J. S. Binkley, C. Gonzalez, D. J. Defrees, D. J. Fox, R. A. Whiteside, R. Seeger, C. F. Melius, J. Baker, R. L. Martin, L. R. Kahn, J. J. P. Stewart, S. Topiol, and J. A. Pople (Gaussian, Inc., Pittsburgh, PA, 1990).
- ³⁶ T. H. Dunning and P. J. Hay, in *Modern Theoretical Chemistry*, edited by H. F. Schaefer III (Plenum, New York, 1976), Vol. 3.
- ³⁷ P. J. Hay and W. R. Wadt, *J. Chem. Phys.* **82**, 270, 284 (1985).
- ³⁸ T. Shimanouchi, *Tables of Molecular Vibrational Frequencies*, Nat. Stand. Ref. Data. Ser., Nat. Bur. Stand. **39**, 1972.