1-2006

Optimization of Multilayer Wear-Resistant Thin Films Using Finite Element Analysis on Stiff and Compliant Substrates

R. K. Lakkaraju
University of Nebraska - Lincoln

Florin Bobaru
University of Nebraska - Lincoln, fbobaru2@unl.edu

S. L. Rohde
University of Nebraska-Lincoln, srohde1@unl.edu

Published in *Journal of Vacuum Science & Technology A:Vacuum, Surfaces, and Films* **24**:1 (January 2006), pp. 146–155; doi 10.1116/1.2121750

Submitted August 11, 2005; accepted September 19, 2005; published online December 23, 2005.

Optimization of multilayer wear-resistant thin films using finite element analysis on stiff and compliant substrates

R. K. Lakkaraju, Department of Mechanical Engineering, University of Nebraska-Lincoln, Lincoln, Nebraska 68588

F. Bobaru, Department of Engineering Mechanics, University of Nebraska-Lincoln, Lincoln, Nebraska 68588

S. L. Rohde, Department of Mechanical Engineering, University of Nebraska-Lincoln, Lincoln, Nebraska 68588; srohde@unlnotes.unl.edu

Abstract

Extensive research has been carried out by researchers on the growth and characterization of multilayer protective coatings, but the design of these coatings still remains largely empirical. In this regard, recent progress has been made in developing a design approach for optimizing a multilayer coating structure before deposition, which would help save time and material. In pursuit of an optimal design, finite element analysis using a plane strain Hertzian contact model was developed to investigate the stress/strain behavior within the layers of the system. The present study looks to find the optimal thicknesses of individual layers in a multilayer coating/substrate system that can reduce stresses and/or strains in the system. Multilayer Cr/CrN thin films were modeled and optimized to have effective "load support" by the films on stiff A2 steel and compliant 2024-Al substrates. Optimization was carried out using both multiobjective and single-objective procedures for the models of eight-layer film on substrates. For the multilayer on A2 steel substrate, the first test case is a multiobjective optimization performed by minimizing the strain discontinuities at the coating/substrate interface and the stresses developed in the uppermost layer under combined normal and tangential load conditions. Another option is a single-objective optimization (minimizing the strain discontinuity) and constraining the stress to values below the yield stress. The same two test cases were employed on the 2024-Al model, but the stresses considered were those in the substrate in order to keep the model within the elastic regime. Efficiency of several optimization algorithms, such as genetic algorithms and gradient based routines are discussed and the preliminary results are compared to experimental pin-on-disk wear results of empirically designed coatings. Architectures were found with improvements in the elastic measures employed here.

Introduction

Over the last two decades, extensive research has been carried out to find the best possible protective coating to decrease the wear of machine tools and engineering components. The past decade has seen a tremendous growth in the utilization of CrN films because of their reduced residual stress, improved oxidation resistance at high temperatures, and lower coefficients of friction when compared to TiN films.[1–4] Recently, our focus has shifted to Cr/CrN multilayer films,[5] which can be very effectively used to further improve the fracture toughness, hardness, and adhesion of these films.[6]

Design of an optimal multilayer coating structure requires a detailed knowledge of the stress/strain profiles produced within the multilayer coating structure under mechanical loading. Plastic deformation of the substrate, coating delamination from the substrate, and through-thickness cracking of the coating has experimentally been observed in multilayered coated systems. If one could minimize the stresses and strains developed in the coating/substrate system by changing the architecture of multilayered coating, the propensity for delamination, plastic deformation, and fracture would be reduced. Hence, wear properties, as a consequence, would be improved.

The two most important stresses to be considered for the design are the residual and mechanical stresses. The former are produced from the mismatch of the coefficients of thermal expansion during deposition and the latter are due to external loadings. In the present article, as a first step in optimizing the coating architecture, only the stresses produced by the external loading are considered. To the authors' knowledge, there are limited design and/or optimization methodologies for obtaining optimal architectures of multilayered coatings under complex loadings. This study is aimed at beginning to meet this need.

Finite element analysis (FEA) has been used in the past decade to investigate the deformations of single, bilayer, and multilayer coating structures under normal, or normal and tangential loadings, which model the interaction between the indenter and the coated structure.[7–14] The influence of friction, film thickness, modulus ratios, and local maximum stresses has been studied for single-layer and multilayer films. Recently, it has also been shown that stresses are relaxed by keeping a metallic layer at the substrate interface and by placing metallic layers between ceramic layers.[15] To date, however, only very limited work

has been done to design a coating architecture with the potential of enhancing performance and avoiding premature failure of long-wearing components.

The eight-layer system with total thickness fixed to 2 μm was used since a number of empirically designed coatings of this form have been previously deposited and experimentally tested.[16] In the previous work examined experimentally, various architectures were obtained by varying the thickness of individual layers, thus providing the motivation for searching for the *optimized* thicknesses in an eight-layer coating/substrate system. It was observed that these different architectures had significantly different wear properties.

The objective of the present work is to couple a FEA for a load-specific case with optimization algorithms in order to produce an *optimal* coating architecture for a multilayer system that reduces certain measures of elastic stresses and/or strains. The hope is that by doing so, the optimized architecture will have improved wear properties, conserve deposition material, reduce mechanical testing, and minimize downtime. For a model with a fixed number (8) of alternating Cr/CrN layers on A2 steel substrate, a multiobjective optimization algorithm is employed to find the optimal thickness of each coating layer that minimizes the maximum elastic strain discontinuity across the coating/surface interface and the von Mises stress in the uppermost coating layer. Single-objective constrained optimization is also used, in which the objective function is based on the strain discontinuities at the coating/substrate interface and the constraint is to keep von Mises stresses within the elastic limit. A similar eight-layer Cr/CrN on 2024-Al substrate was also evaluated. The von Mises stress in the substrate was observed to be reaching the yield stress and hence, for the optimization of this architecture, maximum strain discontinuity at the coating/substrate interface was minimized and the von Mises stress in the substrate was constrained in order to keep the model from reaching plasticity. Genetic algorithms (GA) and gradient based methods (GBM) were used in both cases and efficiency of these techniques was evaluated.

Finite Element Analysis Of Multilayered Coatings

Finite element analysis was performed to study the stress, strain, and displacements produced in a multilayer coating system. Thermal, intrinsic, and external loading stresses are the typical film stresses of interest. Under many loading conditions, thermal stresses, such as low-to-moderate speed sliding wear, can be assumed to be negligible. Thus, these stresses are not incorporated in the present analysis. Intrinsic or residual stress induced in the growth process can be quite significant (even >4 GPa); however, in efficient wear coating these stresses are normally compressive and may be treated in future studies by superimposing an experimentally determined residual stress on the mechanical model. The present study focuses on the details of the stresses/strains produced only by external mechanical loading under normal and tangential loading conditions. These conditions are used to provide a simple simulation for a pin-on-disk experiment involving a spherical indenter under normal and tangential loading. The tangential loading (due to friction) destroys the axisymmetry of the problem; therefore, a three-dimensional numerical model is necessary for an accurate representation of the deformation in the coating/substrate system. Since a coupling between the analysis engine (FEA in present case) and optimization algorithms involve iterative solutions of the mechanical problem, in this study, a two-dimensional plane strain approximation is used. The profile of the normal loading is obtained from the Hertzian theory of elastic bodies in contact.[17] The frictional force is modeled as constant across the contact region.[15] Other nonconstant values of forces can be used with our implementation, but this is not pursued here.

Table I. Mechanical properties of the materials used (see Reference 15).

Material	Young's modulus (GPa)	Poisson's Ratio
A2 steel	200	0.33
2024-Al	72.4	0.33
Cr	180	0.3
CrN	280	0.3
WC	620	0.186

A. Model development

In the FEA, it is assumed that the film is perfectly bonded to the bulk material; additionally, the model is defined to be homogeneous and isotropic.[17] It has been investigated by Gupta and Walowit[18] that in multilayer thin films the elastic modulus ratio between film and substrate should be less than 4 to avoid the deviation from the Hertzian theory. The model used in present work has a ratio less than 4 and, hence, the normal distribution falls under Hertzian theory. Normal and tangential loads (to simulate friction-induced loading) were applied on to the film by the tungsten carbide indenter. The indenter used has a radius of 10 mm, and the properties of the materials used are shown in Table I. The normal load used for the present study is 10 N when applied on the steel substrates[19] and 2 N on 2024-Al substrates. It was observed that when a 4 N load was applied on the Cr/CrN/Al architecture, as was used in the experimental study, the model stresses exceeded the yield stress in the Al substrate. The load was reduced to 2 N in order to maintain the model in elastic regime. The coating architectures obtained using an elastic model may also give a good indication for the models which reach plasticity, but this is subject of future research.

To obtain the normal loading profile to be implemented in FEA, the Hertzian theory is used. The Hertzian theory considers a spherical indenter to impart the normal pressure distribution over the contact region with the elastic half space. The equations of the Hertzian contact are detailed

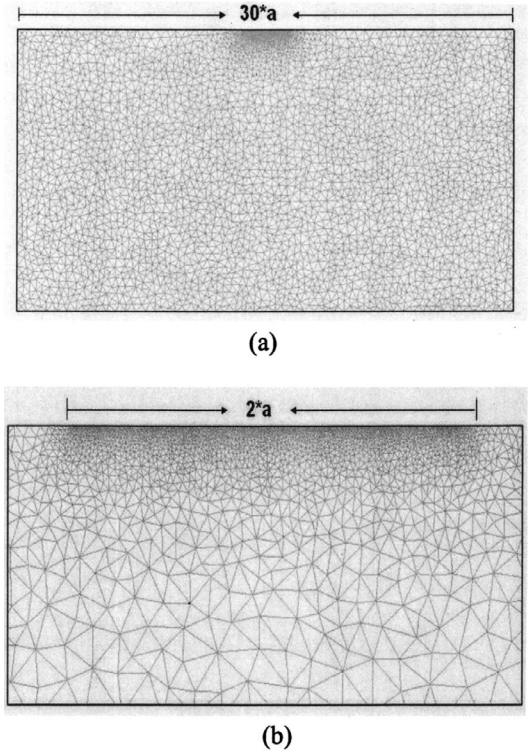

Figure 1. (a) FEM mesh used for an uncoated A2 steel substrate, **(b)** Magnification of the mesh at the contact region showing a larger number of elements.

below.[20] The half width a of the spherical indenter is given by

$$a = \left[\frac{3PR}{4E^*}\right]^{1/3}, \tag{1}$$

where R is the radius of the spherical indenter and P is the load applied on the surface. Equivalent modulus E^* between the two contacting bodies of differing modulus is given by

$$E^* = \left[\frac{1-\upsilon_1^2}{E_1} + \frac{1-\upsilon_2^2}{E_2}\right]^{-1}, \tag{2}$$

where υ_1 and υ_2 are the Poisson's ratio and E_1 and E_2 are the Young's modulus of the coating and indenter, respectively.

The normal pressure distribution is given by

$$p(x) = P_0\left[1 - \left(\frac{x}{a}\right)^2\right]^{1/2}, \tag{3}$$

where, the maximum contact pressure is given by

$$P_0 = \left[\frac{3P}{2\pi a^2}\right]. \tag{4}$$

The tangential load with coefficient of friction of μ=0.5 is given by

$$t(x) = \mu P_0\left[1 - \left(\frac{x}{a}\right)^2\right]^{1/2}. \tag{5}$$

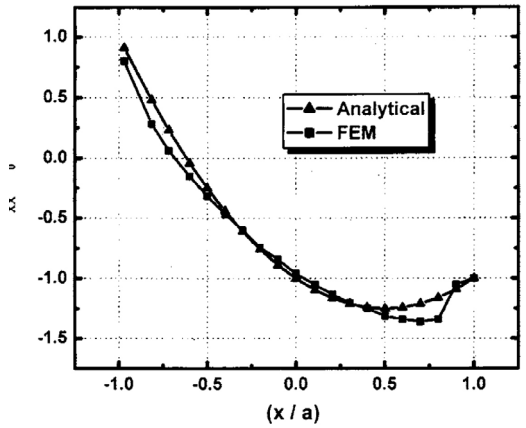

Figure 2. Comparison of analytical and FEM results for normalized σ_{xx} stress vs position of x normalized by half contact width a, along the contact region for uncoated A2 steel substrate under combined normal and tangential loading.

The finite element (FE) model was developed using the FE software, FEMLAB®. The geometry was built according to the contact half width, a. After considerable research, it was determined that a block with width equaling 30 times the half contact area and depth equaling fifteen times the half contact area (i.e., 30a by 15a) was necessary to ensure minimal influence of boundaries on the results of the model.[19] The base node displacements were restrained in all directions in order to avoid the rigid body motions which could be produced due to the tangential loads.[11]

B. Meshing and validation

In order to select an appropriate mesh for the coated system, numerical tests were performed on the uncoated system for which analytical solutions are available. The typical model geometry with the mesh for an uncoated A2 steel substrate is shown in Figure 1(a). The mesh size was finer in the contact region to improve resolution in the region of greatest interest. An enlarged picture of the meshing under the contact region is shown in Figure 1(b). For the uncoated substrate, the number of elements in the contact region was increased by having an element size of $a/100$, with the mesh becoming coarser as it moves towards the boundaries.

Due to the computer's memory limitations, the FE model was solved within FEMLAB® using a conjugate gradient iterative method with a geometric multigrid preconditioner. The validation of the model was performed by comparing the analytical results to the field emission microscopy (FEM) results obtained within the contact region for an uncoated substrate. In Figure 2, the normalized surface stresses (σ_{xx}) produced under normal and tangential loading on an uncoated substrate with a coefficient of friction of 0.5 are plotted against position x normalized by the half contact length a. A satisfactory approximation of the analytical solution was observed. For the optimization problem of the multilayer coated system, the size of the elements in the contact region was taken as earlier, while in the lower part of the substrate, an increas-

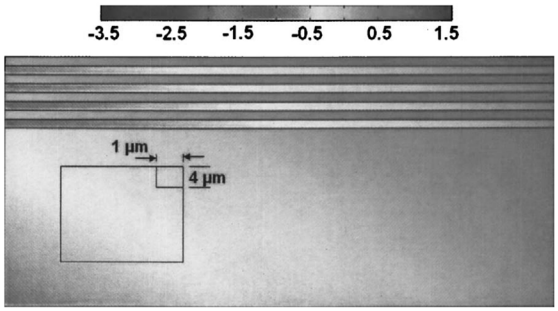

Figure 3. Strain (ϵ_{yy}) discontinuities observed across the layers under combined normal and tangential loading on A2 steel substrate. **Inset:** The original figure is the magnified version of the top right corner within the contact region.

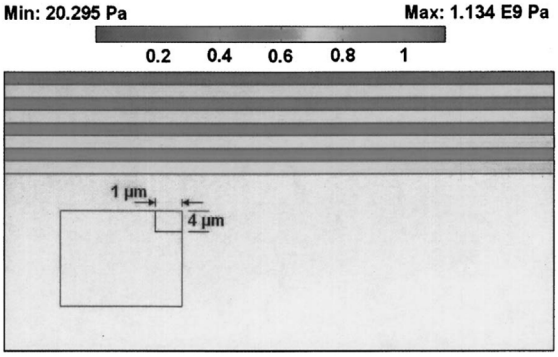

Figure 4. von Mises stress distribution in the contact region in the eight-layer Cr/CrN film on A2 steel substrate under normal and tangential loading (results plotted on the undeformed configuration).

ingly coarser mesh was used in order to reduce the computational burden. A convergence study was performed in order to ensure the accuracy of the stresses and strains developed in the system.

It has been observed experimentally that having a metallic Cr layer at the interface improves the adhesion between the coatings and substrate at the interface, and that by having metal layers sandwiched between ceramic CrN layers results in better pin-on-disk wear performance, presumably due to relaxed mechanical stresses induced between the layers and in the coating.[15] Hence, Cr/CrN layers were built with Cr at the interface and CrN as the uppermost layer.

The normal strains and von Mises stress developed in the eight-layer Cr/CrN/A2 steel model are shown in Figures 3 and 4, respectively. The maximum von Mises stresses appeared in the uppermost layer, while the normal strains were discontinuous across the layers. Hence, for reducing the probability of mechanical failure we proposed to minimize the maximum strain discontinuity ($\Delta\epsilon_{yy}$) across the interfacial coating layer (Cr) and the substrate (A2 steel), as well as minimize the largest von Mises stress in the uppermost layer (CrN).

The von Mises stresses and the strain discontinuities in the x direction ($\Delta\epsilon_{xx}$) were also monitored, since the location of highest value of these can move to different layers. The ϵ_{xx} was observed to be continuous and ϵ_{yy} to be discontinuous through the layers and across the coating/substrate interface. For the optimization setup, the thicknesses of the first seven layers are defined as control parameters: t_1, t_2, …, t_7. The thickness of the eighth layer is then given by t_8: $\left[2 - \sum_{i=1}^{7} t_i\right]$ μm so that the total thickness of the coating stays constant during optimization iterations. These are the design variables, where t_1 is the Cr layer at the interface and t_8 is the uppermost CrN layer. These design variables, as they vary, change the values of the objective function and constraints in the optimization problem described in the next section.

Optimization

Optimization is defined as the process of finding the values of the design variables that maximize or minimize a function. It is customary to cast all optimization problems as minimizations and therefore, if the actual goal is to find a maximum of function f, one can minimize the function $-f$ instead. Optimization algorithms converge to either a global optimum (truly the lowest value) or a local optimum (the lowest value in a certain neighborhood).

An optimization problem is defined by objectives and constraints. A multiobjective unconstrained problem is defined as

$$\begin{aligned} &\min f_1(x_1, x_2, \ldots, x_p), \\ &\min f_2(x_1, x_2, \ldots, x_p), \\ &\text{and } l_i \le x_i \le u_i, \ i = 1, 2, \ldots, p \end{aligned} \tag{6}$$

where x_1, …, x_p are the design variables on which the objective function depends and the design variables may be bounded by lower (l_i) and upper (u_i) bounds. The solution to Equation (6) can be expressed by the Pareto front which is a nondominated solution with respect to the objectives. In order to select a solution from the Pareto set, a multiobjective optimization problem is transformed into a single-objective optimization problem by considering the minimization of a weighted sum

$$F(x) = \alpha_1 f_1 + \alpha_2 f_2 + \ldots + \alpha_s f_s \tag{7}$$

where α_1, …, α_s are user-selected weights such that, $\sum_{i=1}^{s} \alpha_i = 1$.

A single-objective constrained optimization problem is defined by

$$\begin{aligned} &\min f(x_1, x_2, \ldots, x_p), \\ &\text{subject to } g(x_1, x_2, \ldots, x_p) \ge 0, \\ &\text{and } l_i \le x_i \le u_i, \ i = 1, 2, \ldots, p, \end{aligned} \tag{8}$$

where the $x_1, \ldots, x_p$ are the design variables on which the objective function f and the constraint g depend. The design variables may be bounded by lower (l_i) and upper (u_i) bounds. An example of a constraint is limiting stresses to remain below the value of a material's yield stress. In our case, in addition to a multiobjective optimization problem defined by two objectives, constrained single-objective optimization is also employed.

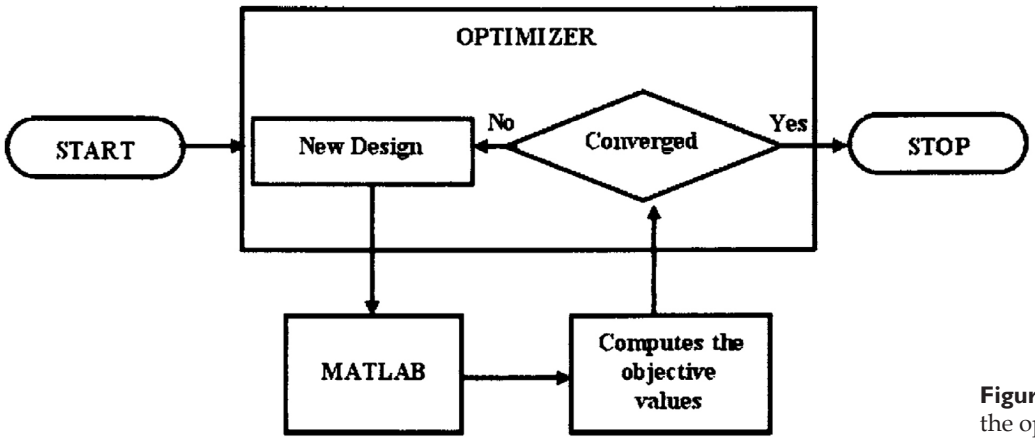

Figure 5. Functional flowchart of the optimization setup.

One of the most successful methods for solving non-linear constrained optimization algorithms is sequential quadratic programming.[21,22] It is a robust and efficient method to solve a smooth non-linear constrained optimization problem. In the present work, NLPQLP (Reference 23) and NBI-NLPQLP (Reference 24) algorithms are implemented in modeFRONTIER™ (Reference 25) for the single-objective constrained optimization and multiobjective unconstrained optimization, respectively additionally, a MOGA-II (Reference 26) algorithm was employed for both the cases.

Evolutionary algorithms, such as GAs form a different class of optimization methods which generally tend to converge to the global optimum. The GA reaches an optimum by starting with an initial population and evolves through natural selection theory using reproduction, crossover, and mutation. Reproduction uses the different techniques (roulette wheel, tournament selection, etc.) in finding the best possible offspring for the next generation with higher fitness values. Crossover is used to mate the strings to obtain values for next generation. Finally, mutation is used to avoid a premature loss of data.[27] One of the major reasons for selecting a GA for optimization problems is due to its robustness and its tendency to reach a global optimum. However, GAs are very slow in convergence and there are many other problem-dependent parameters in a GA which influence the convergence rate and for which there are no general methods of selection.

On the other hand, in a GBM a solution is derived by moving the design variables in the direction computed using the gradient of the function to reach a lower value of the objective function than the previous iteration. GBMs converge fast if the starting guess is close enough to the minimizer and they use information pertaining to the specific problem since gradient and Hessians are employed. These algorithms generally converge to a local minimum. By changing the values of the starting design variable, one can find a number of local minima, and the chance that one of them will be global minima is increased.

A. Problem definition

The multiobjective optimization problem examined has eight design variables (individual thicknesses t_i, i=1, …, 7 and $t_8 = :\left\lfloor 2 - \sum_{i=1}^{7} t_i \right\rfloor$) and two objectives functions f_1 and f_2 to be minimized

$$f_1 = \Delta\epsilon_{yy}, \tag{9}$$

$$f_2 = \max(\sigma_{\mathrm{vm}}). \tag{10}$$

The first objective function f_1 corresponds to the maximum normal strain discontinuity across the substrate/coating interface ($\Delta\epsilon_{yy}$). The value $\Delta\epsilon_{yy}$ was computed by

$$\Delta\epsilon_{yy} = \max_{k=1,\,1500} \left| \Delta\epsilon^{s}_{yy}(X_k) - \Delta\epsilon^{c}_{yy}(X_k) \right|, \tag{11}$$

where ϵ^{s}_{yy} and ϵ^{c}_{yy} are normal strains in the substrate and interfacial layer of the coating, respectively. The X_k values are the x coordinates of approximately 1500 equally spaced points which have the y coordinate 0.05 μm above the coating/substrate interface and 0.2 μm below it. The second objective function f_2 corresponds to the maximum von Mises stress produced in the uppermost layer of the coating for the model with A2 steel and maximum stress in substrate for the model with 2024-Al. Equal weights are given to the two objectives, and by using the multicriterion decision maker,[25] a solution is extracted from the Pareto set.

The second problem tested uses a constrained single-objective optimization carried out with the same eight design variables defined earlier. The objective function f_1 is given as

$$f_1 = \Delta\epsilon_{yy} \quad \text{subject to } \sigma_{\mathrm{vm}} < 2\ \mathrm{GPa}, \tag{12}$$

where σ_{vm} is the maximum von Mises stress in the uppermost layer for A2 steel and in the substrate for 2024-Al.

B. Methods employed

A flowchart explaining the coupling of the FE software FEMLAB® and optimization software modeFRONTIER™

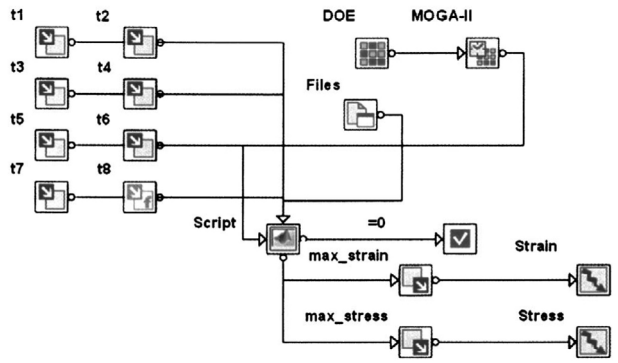

Figure 6. Project setup of a multiobjective optimization using MOGA-II algorithm in modeFRONTIER™.

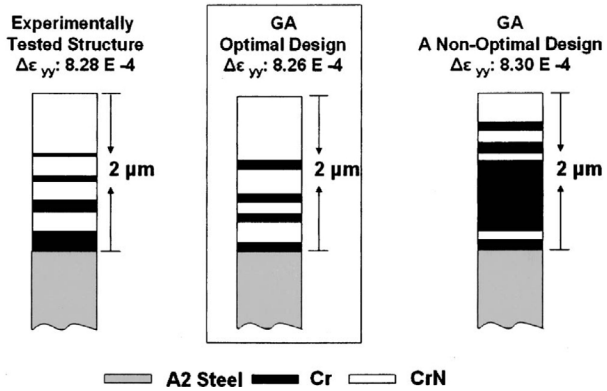

Figure 7. Schematic configurations obtained using a multiobjective optimization with GA, compared to an empirically developed and previously tested structure for Cr/CrN/A2 steel.

is shown in Figure 5. The finite element code in MATLAB® acts as the interface between the optimization and finite element software. The output values are read by the optimization algorithm and in turn the design values are changed to produce a different architecture. This iterative process continues until an end condition is reached. The first test case of multiobjective optimization was carried out using the genetic algorithm MOGA-II and the gradient based NBI-NLPQLP. MOGA-II is developed based on a standard genetic algorithm with advanced features such as elitism and directional crossover. Elitism helps to keep the best solutions from generation to generation; directional crossover helps to increase the efficiency of the algorithm.

For the present study, when using MOGA-II the probability of classical crossover and probability of mutation were taken to be 0.6 and 0.0333, respectively, and 32 designs were used for the initial population. Since the problem deals with eight design variables and two objectives, the size of initial population used was 32, according to the general rule: initial population = 2 × (number of variables) × (number ofobjectives). Tournament selection method is used to find the best individuals in a group instead of the traditional roulette wheel method in order to reduce the computational expense. A generational evolution-type algorithm was used with 50 generations to be evaluated. Mutation rates are kept low so that the search does not become random. The total setup of the project in modeFRONTIER™ for multiobjective optimization using GA is shown in Figure 6.

In order to compare the efficiency of the GA and GBM, a multiobjective gradient based algorithm, NBI-NLPQLP, was used. The normal boundary intersection method (NBI) is coupled with the NLPQLP algorithm in modeFRONTIER™ software to solve the multiobjective optimization. The NBI method divides the objective into subproblems and assigns the individual objectives to the NLPQLP scheduler. Each subproblem is solved using the initial designs, and the NBI sets the internal parameters accordingly to reach the optimal Pareto set. Information from the previous subproblem is used as a starting point for next subproblem; all the subproblems are solved successively to reach the Pareto optimal solution set. Hence, an increase in the number of subproblems leads to a Pareto set with good resolution. The central difference method is used for the approximation of the derivatives because of their second order approximation properties. The magnitude of error in this method is low when compared to the forward difference method which converges to the derivative with the order of one. Random sequence technique algorithm present within modeFRONTIER™ is used to generate an initial set of 16 designs and 50 subproblems were evaluated for the present problem.

The second test was a constrained single-objective optimization performed using MOGA-II and NLPQLP methods. Sixteen designs were generated for the initial population and again evaluated for 50 generations. A NLPQLP gradient based method was used and 500 iterations were performed. The present work was carried out on an Intel P-4, 2.8 GHz processor machine with 1.5 GB of random access memory, running a Windows operating system.

Numerical Results and Discussions

The goal was to design and optimize the coating architecture of wear resistant multilayer thin films with respect to reducing elastic stresses and strains, which in turn may reduce the mechanical failure. Gradient based methods

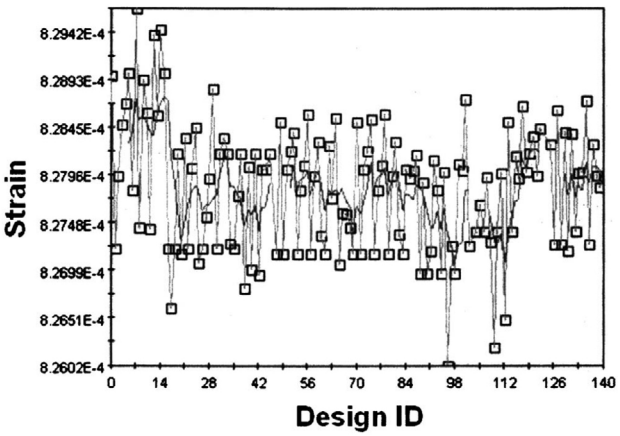

Figure 8. Strain vs design ID generated using a gradient-based multiobjective optimization (NBI-NLPQLP) with Pareto solution.

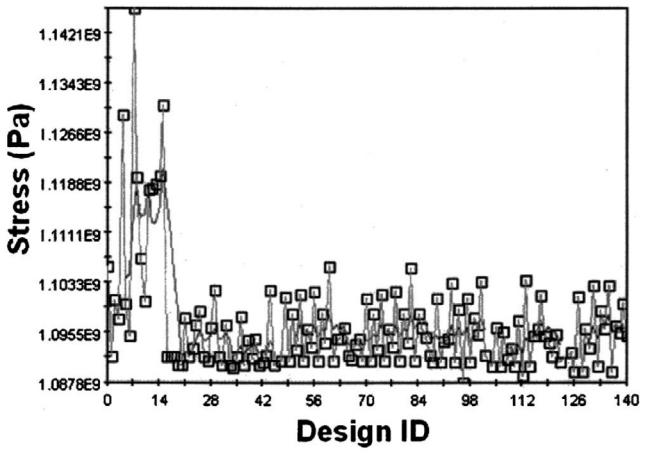

Figure 9. Stress vs design ID generated using a gradient-based multiobjective optimization (NBI-NLPQLP) with Pareto solution.

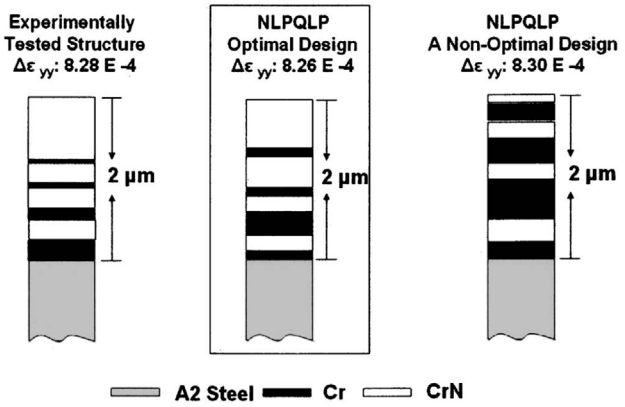

Figure 11. Schematic configurations obtained using a single-objective optimization with sequential quadratic programming (NLPQLP) compared to an empirically developed and previously tested structure for Cr/CrN/A2 steel.

and GA were used to reach this goal, and these techniques were compared to each other to evaluate their efficiency.

A. Cr/CrN multilayer thin films on A2 steel substrate

Displayed in Figure 7 is the optimized structure compared to an empirically designed structure that has been successfully tested,[5] as well as a nonoptimal solution obtained by multiobjective optimization using GA. It was observed that the optimal solution obtained had a strain value approximately 0.5% less than one of the nonoptimal solution tested in the optimization iterations and appears to be a "better" design, in terms of the measures used here as objectives and constraints, than the most successful empirically developed coating.

Shown in Figures 8 and 9 are the strain and stress values versus number of iterations, respectively. It was observed that number of iterations used for this method was less than in the multiobjective GA technique, thereby reducing the overall computational time to reach an optimal solution. Though the solution might have reached the local minimum, the minimized objective function values were found to be very close to the GA values.

Using an alternative single-objective optimization problem, two additional test cases were generated. First, a constrained single-objective optimization was performed using a MOGA-II algorithm, and it was observed that as the number of designs increased, the objective function value approached its minimum, as shown in Figure 10. Second, a single-objective gradient based NLPQLP algorithm was used, giving the solutions shown schematically in Figure 11. It can be observed that the Cr layer at the interface is thinner and that the uppermost CrN layer is thicker than intermediate ceramic layers for the optimal structure produced using a single-objective GBM. A thinner interface layer and a thicker uppermost ceramic layer are observed in single-objective optimization using both algorithms.

The four optimal structures produced using the GA and gradient based techniques for single and multiobjective optimization are compared in Table II. It was observed that the computational time taken by the GA tech-

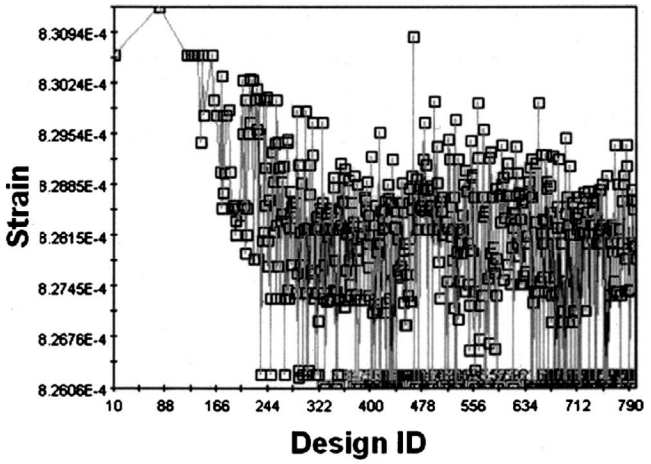

Figure 10. Strain vs design ID generated using a single-objective optimization by GA (using MOGA-II algorithm) with Pareto solution.

Table II. List of the optimal solutions for Cr/CrN/A2 steel obtained using GA and gradient based techniques with computational time and minimized strain and stress values.

Configuration	Multiobjective GA	Multiobjective, NBI-NLPQLP	Single-objective-GA	Single-objective NLPQLP
T1	0.125	0.1	0.1	0.1
T2	0.275	0.3	0.15	0.2
T3	0.1	0.1	0.125	0.3
T4	0.15	0.18	0.375	0.2
T5	0.1	0.1	0.15	0.1
T6	0.325	0.28	0.125	0.4
T7	0.1	0.1	0.15	0.1
T8	0.825	0.84	0.825	0.6
Strain	8.26 *E*-4	8.26 *E*-4	8.26 *E*-4	8.26 *E*-4
Stress (Pa)	1.0873 *E*9	1.0878 *E*9	1.0949 *E*9	1.1005 *E*9
Time	~17 h	~3 h	~9 h	~6 h

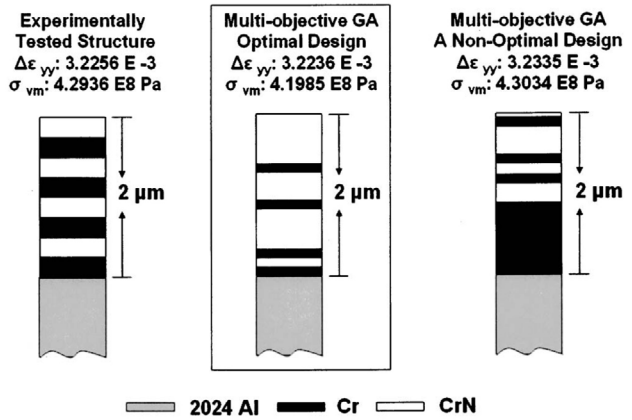

Figure 12. Schematic representation of the optimal solution, experimentally tested architecture and a nonoptimal solution using multiobjective GA for Cr/CrN/Al.

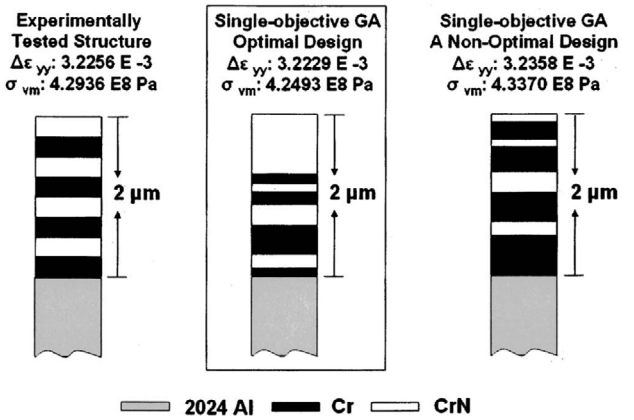

Figure 14. Schematic representation of optimal structure, experimentally tested architecture and a nonoptimal solution using single-objective GA for Cr/CrN/Al.

niques is significantly larger, as expected, when compared to the gradient based techniques. It was also noted that the final minimized objective values of stresses for both techniques were within 2% of each other. Since the difference is small, it is likely that gradient based techniques can be efficiently used instead of GA to minimize the computational time.

B. Optimized Cr/CrN multilayer thin films on compliant Al substrate

The first test case considered was multiobjective optimization, and GA and GBM were used and compared. Optimal coating architecture produced by GA compared to previously tested architecture and a nonoptimal architecture is shown in Figure 12. Though there was a negligible decrease in strain discontinuity and a 2.4% decrease in the stress values compared to a sample of nonoptimal case computed during optimization iterations, the optimal design has the thinnest metallic Cr layers and thickest uppermost CrN layer following the same trend as optimal architectures on stiff A2 steel substrates.

The NBI-NLPQLP method used for the gradient based multiobjective optimization used much less time when compared to the GA, 5 h compared to 28 h. The optimal coating architecture is shown in Figure 13, and it can be observed that the optimized structure has similar architecture and minimized strain values compared to the GA optimization. Again, the reduction in strain was negligible, and the optimized stresses showed a reduction of 2.2% when compared to one of the nonoptimal designs from the optimization iterations. Given the similar final results and greatly reduced computational time, the NBI-NLPQLP is likely to be a better method for these particular problem criteria.

Results for second test case with single-objective optimization were analyzed, and the optimal coating architecture obtained using GA is shown in Figure 14. It can be seen that the obtained optimal architecture has thinnest metallic Cr layer at interface and thickest CrN uppermost layer, showing the consistency of all the algorithms. Though the stress was constrained, the final architecture resulted in the stresses being reduced to the range of the values in multiobjective case.

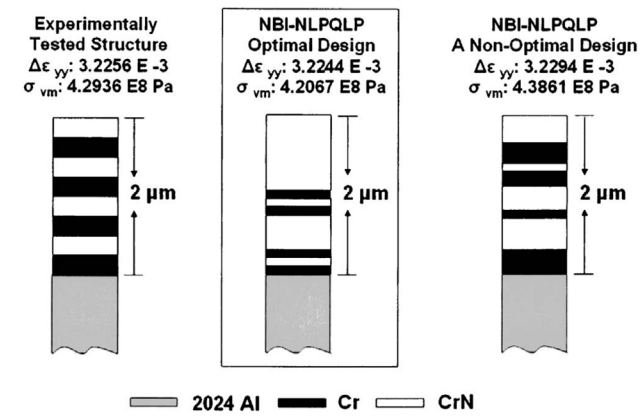

Figure 13. Schematic representation of optimal solution, experimentally tested architecture and a nonoptimal solution using multiobjective NBI-NLPQLP method for Cr/CrN/Al.

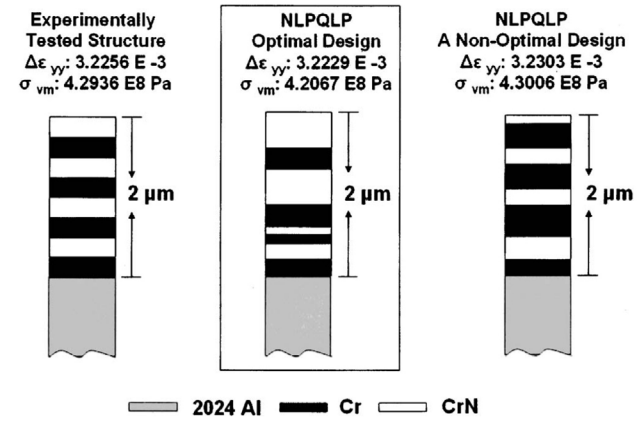

Figure 15. Schematic representation of optimal solution, experimentally tested architecture and a nonoptimal solution using single-objective NLPQLP method for Cr/CrN/Al.

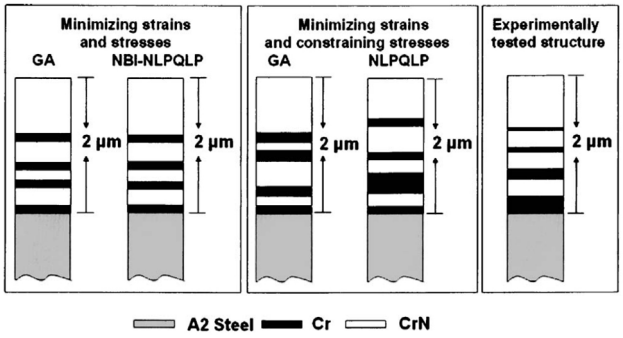

Figure 16. Schematic of the optimal coating architectures obtained using GA and GBM, compared to successful, but empirically designed architecture for Cr/CrN/A2 steel.

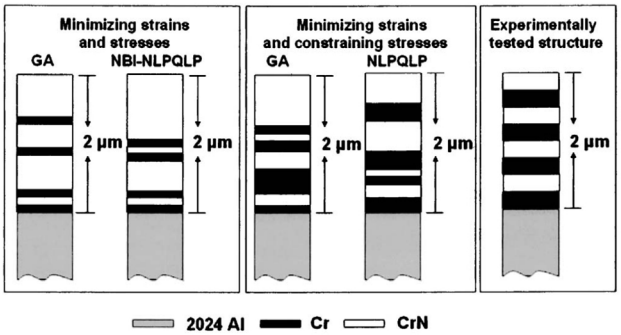

Figure 17. Schematic of the optimal coating architectures obtained using GA and GBM, compared to successful, but empirically designed architecture for Cr/CrN/Al.

Finally, the optimal coating structure compared to the empirical structure and a nonoptimal solution for NLPQLP method is shown in Figure 15. This architecture shows a deviation from the other algorithms with slightly thicker metallic layers. This might be due to the fact that the solution likely reached a local minimum. Nevertheless, again the objective values were reduced to the range of GA, but far less time was required for this to be achieved.

The final architectures for the optimized multilayer thin films on compliant Al substrate, along with the optimization times, are shown in Table III. It can be clearly observed that GA took a very long time, as expected, and reached nearly the same objective values as that of GBM. A multiobjective GBM is preferred for this particular problem in order to get an optimal architecture with considerably less time.

The optimal coating architectures for A2 steel and 2024-Al obtained from four methods are shown in Figures 16 and 17, respectively. It is observed that the designs obtained by GA and the gradient based algorithms show similar trends for both the steel and aluminum. The possibility of other local minima exists, and in future we plan to analyze this. It remains to be seen experimentally which of the optimal designs obtained here would perform better.

Table III. List of the optimal solutions for Cr/CrN/Al models obtained using GA and gradient based techniques with computational time and minimized strain and stress values.

Configuration	Multiobjective GA	Multiobjective, NBI-NLPQLP	Single-objective GA	Single-objective NLPQLP
T1	0.1	0.1	0.1	0.2
T2	0.125	0.1	0.175	0.2
T3	0.1	0.1	0.35	0.1
T4	0.5	0.44	0.275	0.1
T5	0.1	0.1	0.15	0.25
T6	0.35	0.1	0.1	0.45
T7	0.1	0.1	0.1	0.25
T8	0.625	0.96	0.75	0.45
Strain	3.2236 *E*-3	3.2244 *E*-3	3.2229 *E*-3	3.2229 *E*-3
Stress (Pa)	4.1985 *E*8	4.2067 *E*8	4.2493 *E*8	4.2607 *E*8
Time	~28 h	~5 h	~12 h	~5 h

Conclusions

A coating design methodology was developed that integrates a finite element solver with an optimization tool. Defining two different optimization problems and solving each of them with a corresponding genetic algorithm and gradient based method, four similar coating designs with optimized layer thicknesses were found. These optimal architectures minimized a strain discontinuity across the coating/substrate interface and the von Mises stress in the uppermost layers. These optimizations were conducted for both stiff (A2-steel) and (2024-Al) compliant substrate materials. Even though the improvements in the values of the objectives used here are small, it is important to note that the measures used are point-wise maximum values and by reducing them even by a small amount, it is possible that the wear behavior and fracture resistance are significantly improved. It remains to be seen if the improvements, however small, do lead to improved wear properties and reduced failure of the actual system. Presently, this is subject of ongoing experimental research in our lab. Gradient based algorithms produced local minima very close in value to the global minima obtained by GA, while taking less time to arrive at a solution. In terms of modeling, resorting to micromechanics model, fracture and damage criteria may lead to further improvements in the mechanical performance of the multilayered coated systems. The present work provides the first step towards the development of FEA based wear resistant coating design methodology.

Acknowledgments

This work was supported by NSF (CMS-0409728) and Nebraska Research Initiative (NRI). Special thanks to ESTECO for providing the modeFRONTIER™ software used in this work. Particular thanks to research colleagues S. Sevvana, S. Chennadi, J. Li, D. M. Mihut, and B. Hirt for their valuable inputs during the course of the work.

References

1. C. Gautier, H. Moussaoui, F. Elstner, and J. Machet, *Surf. Coat. Technol.* **86–87**, 254 (1996).
2. F. Lévy, P. Hones, P. E. Schmid, R. Sanjinés, M. Diserens, and C. Wiemer, *Surf. Coat. Technol.* **120–121**, 284 (1999).
3. G. Betrand, C. Savall, and C. Meunier, *Surf. Coat. Technol.* **82**, 42 (1996).
4. J. Albert Sue, A. J. Perry, and J. Vetter, *Surf. Coat. Technol.* **68–69**, 126 (1994).
5. M. L. Kuruppu, G. Negrea, I. P. Ivanov, and S. L. Rohde, *J. Vac. Sci. Technol.* A **16**, 1949 (1998).
6. J. Romero, J. Esteve, and A. Lousa, *Surf. Coat. Technol.* **188–189**, 338 (2004).
7. D. J. Ward and R. D. Arnell, *Thin Solid Films* **420–421**, 269 (2002).
8. H. Djabella and R. D. Arnell, *Thin Solid Films* **223**, 98 (1993).
9. H. Djabella and R. D. Arnell, *Thin Solid Films* **245**, 27 (1994).
10. H. Djabella and R. D. Arnell, *Thin Solid Films* **235**, 156 (1993).
11. H. Djabella and R. D. Arnell, *Thin Solid Films* **223**, 87 (1993).
12. R. M. Souza, G. G. W. Mustoe, and J. J. Moore, *Thin Solid Films* **392**, 65 (2001).
13. S. Ramalingam and L. Zheng, *Tribol. Int.* **28**, 145 (1995).
14. N. Ye and K. Komvopoulos, *Trans. ASME* **125**, 692 (2003).
15. T. Gorishnyy, L. G. Olson, M. Oden, S. M. Aouadi, and S. L. Rohde, *J. Vac. Sci. Technol. A* **21**, 332 (2003).
16. M. L. Kuruppu, M.S. thesis, University of Nebraska, Lincoln, 1997.
17. K. Holmberg and A. Matthews, *Coatings Tribology*, Tribology series 28 (Elsevier, New York, 1994), p. 65.
18. P. K. Gupta and J. A. Walowit, *J. Lubr. Technol.* **96**, 250 (1974).
19. D. Mihut, M.S. thesis, University of Nebraska, Lincoln, 1999.
20. K. L. Johnson, *Contact Mechanics* (Cambridge University Press, Cambridge, 1985), p. 427.
21. J. Nocedal and S. J. Wright, *Numerical Optimization* (Springer, New York, 1999), p. 528.
22. S. S. Rao, *Engineering Optimization, Theory and Practice* (New Age International (P) Ltd., 1996), pp. 381–408.
23. K. Schittkowski, "NLPQLP: A New FORTRAN Implementation of Sequential Quadratic Programming Algorithm", Users Guide, Department of Mathematics, University of Bayreuth.
24. Enrico Rigoni, "NBI-NLPQLP Scheduler", *ESTECO Technical Report 2004–003.*
25. ESTECO s.r.l, Trieste, Italy; http://www.esteco.com
26. Kalyanmoy Deb, *Multi-Objective Optimization using Evolutionary Algorithms* (Wiley, New York, 2002), pp. 200–209.
27. D. E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning* (Addison-Wesley, Reading, MA,