

TRUST: A Deterministic Algorithm for Global Optimization

Jacob Barhen, Vladimir Protopopescu, David Reister

An approach to solving continuous global optimization problems was developed. It builds on two innovative concepts, subenergy tunneling and non-Lipschitzian terminal repellers, to ensure escape from local minima in a fast, reliable, and computationally efficient manner. The generally applicable methodology is embodied in the TRUST (terminal repeller unconstrained subenergy tunneling) algorithm, which is deterministic, scalable, and easy to implement. Benchmark results show that TRUST is faster and more accurate than previously reported global optimization techniques. An application of TRUST to a large-scale exploratory seismology problem of substantial computational complexity (that is, residual statics corrections) is also reported.

Examples of the role played by nonlinear optimization in the application of mathematics to practical problems abound in every field of scientific, technologic, economic, or social interest, including physical sciences (elasticity theory, hydrodynamics, celestial mechanics), computer science (design of VLSI circuits), geophysics (determination of unknown geologic parameters from subsurface or underwater measurements), biology (protein folding), industrial technology (optimal control of regulation systems, optimization of industrial flow through the factory line), and economics (optimal warehouse maintenance, optimal transportation routes).

Typically, the overall performance of a system is described by a multivariate function called the objective function. Optimality is achieved when the objective function attains its global extremum, which can be a maximum or a minimum, depending on the problem. For most real-life applications, the objective function depends on a large number of state variables or parameters. Finding the extrema, and in particular, the absolute extremum, of such a function turns out to be painstakingly tedious. The primary difficulty is that the global extremum (for example, a minimum) of a real function is—despite its name—a local property. Significant alteration of the location and magnitude of the global minimum can be carried out without affecting the locations and magnitudes of the other minima. Thus, short of exhaustive search, it would appear to be very difficult to design unfallible methods to locate the absolute minimum for an arbitrary function.

In recent years there has been a surge of interest in global optimization (1–6), and substantial progress has been achieved in breaking new theoretical ground (7–15). Nonetheless, many of the difficulties men-

tioned above have persisted.

The generic global optimization problem considered here can be stated as follows. Let $f(\mathbf{x}) : \mathcal{D} \rightarrow \mathcal{R}$ be a bounded function with a finite number of discontinuities, and \mathbf{x} be an n -dimensional state vector. For $n = 1$, we denote the state variable by x . At any discontinuity point \mathbf{x}^δ , the function f is required to satisfy the inequality $\lim_{\mathbf{x} \rightarrow \mathbf{x}^\delta} \inf f(\mathbf{x}) \geq f(\mathbf{x}^\delta)$ (lower semicontinuity condition). Hereafter, \mathcal{R} will denote the real numbers, $f(\mathbf{x})$ will be referred to as the objective function, and the set \mathcal{D} will be referred to as the set of feasible solutions (or search space). The goal is to find the state vector \mathbf{x}^{gm} that minimizes $f(\mathbf{x})$ in \mathcal{D} . The index gm stands for global minimum. Without loss of generality, we shall take $\mathcal{D} = \{x_i | \beta_i^- \leq x_i \leq \beta_i^+; i = 1, 2, \dots, n\}$, where β_i^- and β_i^+ denote, respectively, the lower and upper bound of the i th state variable.

We shall address the unconstrained global optimization problem in terms of the evolution of a deterministic dynamical system that combines subenergy tunneling and non-Lipschitzian “terminal” repellers. We define the subenergy tunneling transformation of the function $f(\mathbf{x})$ by the following nonlinear monotonic mapping:

$$E_{\text{sub}}(\mathbf{x}, \mathbf{x}^*) = \log(1/[1 + \exp(-\hat{f}(\mathbf{x}) - a)]) \quad (1)$$

In Eq. 1, $\hat{f}(\mathbf{x}) = f(\mathbf{x}) - f(\mathbf{x}^*)$, a is a constant that affects the asymptotic behavior but not the monotonicity of the transformation, and \mathbf{x}^* is a fixed value of \mathbf{x} , whose selection will be discussed below. $E_{\text{sub}}(\mathbf{x}, \mathbf{x}^*)$ has the same discontinuity and extremal points as $f(\mathbf{x})$ and the same relative ordering of the local and global minima, that is, $E_{\text{sub}}(\mathbf{x}, \mathbf{x}^*)$ is a transformation of $f(\mathbf{x})$ that preserves all properties relevant for optimization. In addition, this transformation is designed to ensure that (i) $E_{\text{sub}}(\mathbf{x}, \mathbf{x}^*)$ quickly approaches zero for large, positive $\hat{f}(\mathbf{x})$; and

(ii) $E_{\text{sub}}(\mathbf{x}, \mathbf{x}^*)$ rapidly tends toward $\hat{f}(\mathbf{x})$ whenever $\hat{f}(\mathbf{x}) \ll 0$.

The second concept we have built on is terminal repellers. Suppose that \mathbf{x} obeys an evolution given by the dynamical system $d\mathbf{x}/dt \equiv \dot{\mathbf{x}} = \mathbf{g}(\mathbf{x})$. The solutions of the corresponding stationary system, $\mathbf{g}(\mathbf{x}) = 0$, are called equilibrium points. An equilibrium point \mathbf{x}_{eq} of the dynamical system $\dot{\mathbf{x}} = \mathbf{g}(\mathbf{x})$ is termed an attractor (repeller) if no (at least one) eigenvalue of the $n \times n$ matrix \mathcal{M} , $\mathcal{M} = \partial \mathbf{g}(\mathbf{x}_{\text{eq}})/\partial \mathbf{x}$, has a positive real part. Typically, a certain amount of regularity (Lipschitz condition) is required to guarantee the existence of a unique solution for each initial condition $\mathbf{x}(0)$, and in those cases, the system’s relaxation time to an attractor, or escape time from a repeller, is theoretically infinite. If the regularity condition at equilibrium points is violated, singular solutions are induced such that each solution approaches the terminal attractor or escapes from the terminal repeller in finite time (16). For example, the dynamical system $\dot{x} = x^{1/3}$ has a repelling unstable equilibrium point at $x = 0$, which violates the Lipschitz condition. Any initial condition, no matter how close to the repelling point $x = 0$, will escape the repeller to reach any finite point x_0 in a finite time, $t_0 \sim x_0^{2/3}$. Terminal repellers, in conjunction with the subenergy tunneling, are the foundation of our global optimization algorithm.

We now assemble the above concepts into the TRUST (terminal repeller unconstrained subenergy tunneling) global optimization scheme. Let $f(\mathbf{x})$ be a function one wishes to globally minimize over \mathcal{D} . We define the TRUST virtual objective function

$$E(\mathbf{x}, \mathbf{x}^*) = \log(1/[1 + \exp(-\hat{f}(\mathbf{x}) - a)]) - (3/4)\rho(\mathbf{x} - \mathbf{x}^*)^{4/3}\theta[\hat{f}(\mathbf{x})] \quad (2) = E_{\text{sub}}(\mathbf{x}, \mathbf{x}^*) + E_{\text{rep}}(\mathbf{x}, \mathbf{x}^*)$$

In the above expression θ is the Heaviside function, which is equal to 1 for positive values of the argument and zero otherwise. The first term on the right-hand side of Eq. 2 corresponds to the subenergy tunneling function; the second term is the repeller energy term. The parameter $\rho > 0$ quantifies the strength of the repeller. Application of gradient descent to $E(\mathbf{x}, \mathbf{x}^*)$ results in the dynamical system ($i = 1, \dots, n$)

$$\dot{x}_i = -(\partial f(\mathbf{x})/\partial x_i)(1/[1 + \exp(\hat{f}(\mathbf{x}) + a)]) + \rho(x_i - x_i^*)^{1/3}\theta(\hat{f}(\mathbf{x})) \quad (3)$$

By its very structure, $E(\mathbf{x}, \mathbf{x}^*)$ transforms the current local minimum of $f(\mathbf{x})$ into a global maximum but preserves all

Center for Engineering Systems Advanced Research, Oak Ridge National Laboratory, Oak Ridge, TN 37831-6355, USA.

lower local minima. Thus, when gradient descent is applied to the function $E(\mathbf{x}, \mathbf{x}^*)$, the new dynamics, initialized at a small perturbation from the local minimum of $f(\mathbf{x})$, will escape this critical point to a lower valley of $f(\mathbf{x})$ with a lower local minimum. Hence, application of gradient descent to $E(\mathbf{x}, \mathbf{x}^*)$ defined in Eq. 2, as opposed to the original function $f(\mathbf{x})$, results in a system that has a global descent property. This is the main idea behind constructing the TRUST virtual objective function.

The basic algorithm proceeds as follows. Initially, \mathbf{x}^* is chosen to be one corner of the hyperparallelepiped \mathcal{D} ; for example, $x_i^* = \beta_i^-$, for $i = 1, \dots, n$. A repeller is placed at \mathbf{x}^* . It should be noted that the repelling terms in the multidimensional case can be interpreted as hyperplane repellers and are active whenever $\hat{f}(\mathbf{x}) \geq 0$. The initial state of the system is set to $\mathbf{x}^* + \mathbf{d}$, where \mathbf{d} is a small perturbation that drives the system into \mathcal{D} . We assume \mathbf{d} has a uniform sign during the basic stage of the optimization. Depending on the relative values of $f(\mathbf{x}^*)$ and $f(\mathbf{x}^* + \mathbf{d})$, the dynamical system will initially be in a tunneling phase or a gradient descent phase. Subsequent switchings between the two phases are autonomous.

Our tunneling idea is different in character and implementation from other tunneling methods based on classical diffusion, quantum effects, or metastability. Instead of the state of the system jumping over or diffusing through a barrier, the subenergy tunneling eliminates the barrier and flattens the landscapes (ideally to a plane). Then, instead of propagation of the state through that plane by diffusion, one uses the extremely fast engine of the terminal repeller to move the state on the downward slope of the virtual surface this repeller induces.

A first enhancement to this basic paradigm allows for N_ϕ componentwise direction reversals of the solution flow at each boundary of the domain. The number N_ϕ is specified. Because with each lower minimum identified a larger portion of the dynamical flow will be in a repeller rather than gradient descent mode, successive traversals of the domain become significantly less expensive.

A second enhancement is instantiated immediately after the occurrence of N_ϕ reflections. It involves replacing the constant direction of multidimensional repelling with an ensemble of one-dimensional (1D) tunneling paths. As an illustration of this concept, assume that the system is descending toward the μ -th local minimum, ${}^\mu\mathbf{x}$. The solution evolves on the hypersurface obtained by collapsing $f(\mathbf{x})$

by means of the subenergy tunneling transformation. Specifically, we integrate

$$\dot{x}_i = -\frac{\partial f(\mathbf{x})}{\partial x_i} / [1 + \exp[\mu^{-1}\hat{f}(\mathbf{x}) + a]] \quad (4)$$

where

$$\mu^{-1}\hat{f}(\mathbf{x}) = f(\mathbf{x}) - f(\mu^{-1}\mathbf{x}^*) \quad (5)$$

After convergence to ${}^\mu\mathbf{x}$, a terminal repeller is positioned there, that is, ${}^\mu\mathbf{x}^*$ is set equal to ${}^\mu\mathbf{x}$. Several schemes can be considered for implementing the 1D tunneling away from repeller ${}^\mu\mathbf{x}^*$. The simplest scheme consists in evaluating $f_{[i]}(\mathbf{x})$ on a uniform grid. Alternately, one can integrate the differential equation

$$\dot{x}_i = \rho(x_i - {}^\mu x_i^*)^{1/3} \theta(\mu \hat{f}_{[i]}(\mathbf{x})) \quad (6)$$

from the initial condition ${}^\mu x_i^* + d_i$. In the above expressions, the notations $f_{[i]}(\mathbf{x})$ and $\mu \hat{f}_{[i]}(\mathbf{x})$ are to indicate that all components of \mathbf{x} except x_i are kept fixed. Then, the first value x_i for which the argument of the Heaviside function $\theta(\mu \hat{f}_{[i]})$ becomes negative yields the initial condition, denoted $\mathbf{x}^{(\mu+1)}$, for reaching the next local minimum ${}^{\mu+1}\mathbf{x}$ by means of Eq. 4. If $\theta(\mu \hat{f}_{[i]})$ remains positive over both domains $[\beta_i^-, {}^\mu x_i^*]$ and $[{}^\mu x_i^*, \beta_i^+]$, the next parameter x_{i+1} is explored with, for example, Eq. 6. The choice of the dimension to be explored first (that is, the choice of i) can be random.

The successive tunneling and descent processes continue until a suitable stopping criterion is satisfied. For the 1D case when the state flows out of the domain boundary [for example, when $x > \beta^+$ (assuming positive flow)], the last local minimum found is taken as the global minimum. A formal proof is given in (8, 13). The multidimensional stopping criterion is quite similar to the 1D case. When no new initial condition is obtained [when $\theta(\hat{f}_{[i]}) > 0$, for $i = 1, \dots, n$], the last local minimum found is taken as the global minimum.

For 1D problems, the TRUST algorithm is designed to sweep the whole search space and thus find the global minimum for any bounded lower semicontinuous function with a finite number of discontinuities (13). Indeed, in an ideal implementation, the replacement of the function $f(x)$ with the function $E(x, x^*)$ ensures that (i) the flow follows the field lines in the negative gradient regions; (ii) when the objective function presents a discontinuity, the dynamical system (Eq. 3) can only flow down (like in a waterfall) or tunnel; (iii) once a (local) minimum is found, the regions above this minimum are flattened—ideally to a horizontal line; and (iv) the repeller then bends the horizontal portion down from the local minimum, and the system moves away from it,

until a new region with negative gradient or waterfall is encountered and the process restarts.

A direct extension of the 1D scheme to multidimensional global optimization problems does not guarantee that the global optimum will always be found. This is due primarily to the tacit assumption (13) of constant direction of repelling, \mathbf{d} ($\mathbf{d} \in \mathcal{R}^n$), from a local minimum \mathbf{x}^* . The resulting trajectory of the solution in \mathcal{D} during tunneling is then inflected only by the presence of very strong surface gradients. Consequently, a global minimum may, for example, not capture a trajectory that tunnels near the periphery of its basin of attraction if the surface gradients there are weak.

A formal convergence proof for the multidimensional case has not yet been obtained. In practice, however, because of its global descent property, the system dynamics always escapes local minima valleys with the help of the repeller effect and flows into lower valleys of the objective function using the information it gets from the tunneling process.

In practice, solving multidimensional optimization problems is a trade-off problem—namely, to find the optimal balance between accuracy and cost. In a strategy that realizes measurable trade-offs between these requirements, the multidimensional problem is reduced to the 1D case for which a formal convergence proof exists (8, 13). The first step is to transform a function of n variables into a function of one variable. In the second step, one applies the algorithm of Eq. 3, with $n = 1$, to this new function.

The idea is to construct a 1D curve that “covers” the n -dimensional phase space of the problem. An approximate realization has been proposed (7). Instead of actually “filling” the n -dimensional domain, the curve is only required to pass within a given distance (for example, $\pi\alpha$) from any point in the domain. The parameter α is related to the available or desired accuracy, and it ensures that the curve has finite length and is differentiable. An explicit formula for this transformation can be found in (7), but other recipes can be devised, depending on the shape of the n -dimensional domain, optimization function, or other extraneous criteria. Such a reduction implies a worst case complexity that is proportional to the length of the 1D space and exponential in problem dimension, that is, $O[(\beta/\alpha)^n]$. Once the problem is reduced, one can apply the 1D version of TRUST, which is guaranteed to converge, within an accuracy determined by α .

To assess the TRUST algorithm, we carried out benchmark problems using several standard multidimensional test functions

(12, 13). In Tables 1 and 2, the performance of TRUST is compared with the best competing global optimization methods, where "best" indicates the best widely reported reproducible results we could find for the particular test function. The criterion for comparison is the number of function evaluations. The improvement over our previously published results (13) for such benchmark problems reflects the current advances in the TRUST algorithm. The results in Tables 1 and 2 demonstrate that TRUST is substantially faster than these state-of-the-art methods and produces consistent and accurate results.

One of the primary limitations of conventional global optimization algorithms is their lack of stopping criteria. This limitation is circumvented in benchmark problems, where the value and coordinates of the global minima are known in advance. The achievement of a desired accuracy (such as, $\epsilon = 10^{-6}$) is then considered as a suitable termination condition (2). For consistent comparisons, this condition was also used in TRUST, rather than its general stopping criterion. For each function, corners of the domain were taken as initial conditions; each result then represents the average number of evaluations required for convergence to the global minimum of the particular

function. The TRUST calculations were performed with the value $a = 2$, for which the subenergy tunneling transformation achieves its most desirable asymptotic behavior (8). The dynamical equations were integrated with an adaptive scheme, that, within the basin of attraction of a local minimum, considers the local minimum as a terminal attractor. Typical base values for the key parameters Δ_t (time step used in integrating Eq. 3) and ρ were 0.05 and 10, respectively.

To assess the performance of TRUST for a large-scale practical application, we selected the problem of residual statics corrections for seismic data. In many geophysical tasks seismic energy is detected by receivers that are regularly spaced along a grid that covers the domain being explored. A source is positioned at some grid location to produce a shot. Time series data are collected from the detectors for each shot; then the source is moved to another grid node for the next shot.

A major degradation of seismic signals usually arises from near-surface geologic irregularities (17–19). These include uneven soil densities, topography, and significant lateral variations in the velocity of seismic waves. The most important consequence of such irregularities is a distorted

image of the subsurface structure that is due to misalignment of signals caused by unpredictable delays in recorded travel times of seismic waves in the vertical neighborhood of every source and receiver. The quality of the seismic analysis is improved with timing adjustments (statics corrections). One typically distinguishes between field statics, which correspond to corrections that can be derived directly from topographic and well measurements, and residual statics, which incorporate adjustments that must be inferred statistically from the seismic data. The common occurrence of severe residual statics (where the dominant period of the recorded data is significantly exceeded) and the significant noise contamination render the automatic identification of large static shifts very difficult. This problem has generally been formulated in terms of global optimization, and to date, Monte-Carlo techniques (11) (such as simulated annealing and genetic algorithms) have provided the primary tools for seeking a potential solution. Such an approach, however, is extremely expensive.

The statics correction problem can be summarized as follows. Acoustic signals are shot from N_s source locations and received by N_r sensors. Each signal reflects at a midpoint k , $k = 1, \dots, N_k$. A trace t corresponds to seismic energy traveling from a source s_t to a receiver r_t through a midpoint

Table 1. Number of function evaluations required by different methods to reach a global minimum of standard test functions. Abbreviations for the functions: BR, Branin; CA, Camelback; GP, Goldstein-Price; RA, Rastrigin; SH, Shubert; and H3, Hartman. Abbreviations for the methods: SDE is the stochastic method of Aluffi-Pentini *et al.* (22); EA denotes the evolution algorithms of Yong *et al.* (23) or Schneider *et al.* (24); MLSL is the multiple-level single-linkage method of Kan and Timmer (25); IA is the interval arithmetic technique of Ratschek and Rokne (7); TUN is the tunneling method of Levy and Montalvo (26); and TS refers to the taboo search scheme of Cvijovic and Klinowski (12).

Method	Test function					
	BR	CA	GP	RA	SH	H3
SDE	2,700	10,822	5,439	—	241,215	3,416
EA	430	—	460	2,048	—	—
MLSL	206	—	148	—	—	197
IA	1,354	326	—	—	7,424	—
TUN	—	1,469	—	—	12,160	—
TS	492	—	486	540	727	508
TRUST	55	31	103	59	72	58

Table 2. Comparison of the number of function evaluations (given in parentheses), and precision of each variable (x_1 to x_5) for TRUST, the fast simulated annealing (FSA) algorithm of Szu and Hartley (27), and the stochastic approximation paradigm (SAS) of Styblinski and Tang (28). Global minimum FSA and SAS results are from (28).

Variable	Method			Exact
	FSA (100,000)	SAS (3,710)	TRUST (89)	
x_1	-2.702844	-2.903506	-2.90353	-2.903534
x_2	-3.148829	-2.903527	-2.90353	-2.903534
x_3	1.099552	1.000241	1.00004	1
x_4	1.355916	0.999855	0.99997	1
x_5	1.485936	1.000194	0.99997	1

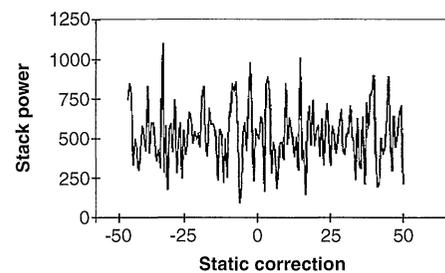


Fig. 1. One-dimensional slice through a 154-dimensional objective function associated with a residual statics problem.

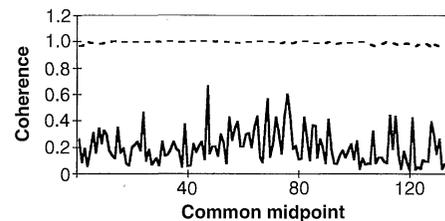


Fig. 2. The coherence factors, that is, the dimensionless ratios E_k/B_k , are plotted for each common gather by using the initial and the optimal time shifts (residual statics). Ideally, at the global optimum, these ratios should be equal to 1. Our results, represented by the dotted line, are indeed quite close to unity.

k_c . We denote by D_{ft} the (complex) Fourier coefficient of frequency f ($f = 1, \dots, N_f$) for trace t , $t = 1, \dots, N_t \leq N_s N_r$. Ideally, after they have been corrected for normal move-out (18), all traces corresponding to the same midpoint carry coherent information. If there were no need for statics corrections, all signals, stacked by their common midpoint, should be in phase and yield a maximum for the total power

$$E(\mathbf{S}, \mathbf{R}) = \sum_k \sum_f \left| \sum_t \exp[2\pi i f(S_{st} + R_{rt})] D_{ft} \delta_{kk} \right|^2 \quad (7)$$

In Eq. 7, the statics corrections $\mathbf{S} = (S_1, \dots, S_{N_s})$ and $\mathbf{R} = (R_1, \dots, R_{N_r})$ are now considered independent variables. Their optimum values are found by maximizing the power E . Equation 7 highlights the multimodal nature of E , which, even for relatively low-dimensional \mathbf{S} and \mathbf{R} , exhibits a very large number of local minima. This is illustrated in Fig. 1.

To assess the performance of TRUST, we considered a problem involving $N_s = 77$ shots and $N_r = 77$ receivers. A data set consisting of $N_t = 1462$ synthetic seismic traces folded over $N_k = 133$ common midpoint gathers was obtained from CogniSeis Corporation (20). The data set uses $N_f = 49$ Fourier components for data representation. Even though this set is somewhat smaller than typical collections obtained during seismic surveys by the oil industry, it is representative of the extreme complexity underlying residual statics problems. To derive a quantitative estimate of TRUST's performance, let E_k denote the total contribution to the stack power arising from midpoint k , and let B_k refer to the upper bound of E_k in terms of \mathbf{S} and \mathbf{R} . Using a polar coordinates representation for the trace data D_{ft} , that is, $D_{ft} = \alpha_{ft} \exp(iw_{ft})$, we can prove (21) that

$$B_k = \sum_f \left(\sum_t \alpha_{ft} \delta_{kk} \right)^2 \quad (8)$$

The TRUST results, illustrated in Fig. 2, show the significant improvement in the coherence factor of each common gather. This factor is the ratio E_k/B_k and characterizes the overall quality of the seismic image.

In conclusion, the TRUST methodology for solving unconstrained global function optimization problems proves to be a powerful tool not only for academic problems; it has the robustness and consistency required by large-scale, real-life applications.

REFERENCES AND NOTES

1. H. Ratschek and J. Rokne, *New Computer Methods for Global Optimization* (Ellis Horwood, Chichester, UK, 1988).
2. A. Törn and A. Ziliinskas, *Global Optimization* (Springer-Verlag, Berlin, 1989).

3. R. Horst and H. Tuy, *Global Optimization* (Springer-Verlag, Berlin, ed. 2, 1993).
4. R. Horst and P. M. Pardalos, *Handbook of Global Optimization* (Kluwer, Dordrecht, Netherlands, 1995).
5. N. V. Thoai, *Introduction to Global Optimization* (Kluwer, Dordrecht, Netherlands, 1996).
6. C. A. Floudas and P. M. Pardalos, *State of the Art in Global Optimization: Computational Methods and Applications* (Kluwer, Dordrecht, Netherlands, 1996).
7. H. Ammar and Y. Cherruault, *Math. Comput. Modeling* **18**, 17 (1993).
8. B. Cetin, J. Barhen, J. Burdick, *J. Optimization Theory Appl.* **77**, 97 (1993).
9. G. L. Bilbro, *IEEE Trans. Syst. Man. Cyber.* **24** (no. 4), 684 (1994).
10. I. P. Androulakis, C. D. Maranas, C. A. Floudas, *J. Global Optimization* **7**, 337 (1995).
11. M. Sen and P. L. Stoffa, *Global Optimization Methods in Geophysical Inversion* (Elsevier, Amsterdam, 1995).
12. D. Cvijovic and J. Klinowski, *Science* **267**, 664 (1995).
13. J. Barhen and V. Protopopescu, in *State of the Art in Global Optimization*, C. A. Floudas and P. M. Pardalos, Eds. (Kluwer, Dordrecht, Netherlands, 1996), pp. 163–180.
14. C. S. Adjiman and C. A. Floudas, *Global Optimization* **9**, 23 (1996).
15. P. M. Pardalos, D. Shalloway, G. Xue, Eds., *Global Minimization of Nonconvex Energy Functions: Molecular Conformation and Protein Folding*, vol. 23 of *Series in Discrete Mathematics and Theoretical Computer Science* (DIMACS) (American Mathematical Society, Providence, RI, 1996).
16. J. Barhen, M. Zak, N. Toomarian, in *Advanced Neural Computers*, R. Eckmiller, Ed. (North-Holland, Amsterdam, 1990), pp. 102–112.
17. D. Rothman, *Geophysics* **50**, 2784 (1985).
18. O. Yilmaz, *Seismic Data Processing* (Society of Exploration Geophysicists, Tulsa, OK, 1988).
19. J. DuBose, *Geophysics* **58**, 399 (1993).
20. The benchmark problem was provided by J. DuBose.
21. J. Barhen, D. Reister, J. DuBose, in preparation.
22. F. Aluffi-Pentini, V. Parisi, F. Zirilli, *J. Optimization Theory Appl.* **47**, 1 (1985).
23. L. Yong, K. Lishan, D. J. Evans, *Parallel Comput.* **21**, 389 (1995).
24. G. Schneider, J. Schuchhardt, P. Wrede, *Biol. Cybern.* **74**, 203 (1996).
25. A. H. G. R. Kan and G. T. Timmer, in *Numerical Optimization*, P. T. Boggs, R. H. Byrd, R. B. Schnabel, Eds. [Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1985], pp. 245–262.
26. A. Levy and A. Montalvo, *SIAM J. Sci. Stat. Comput.* **6**, 15 (1985).
27. H. Szu and R. Hartley, *Phys. Lett.* **A122**, 157 (1987).
28. M. A. Styblinski and T. S. Tang, *Neural Networks* **3**, 467 (1990).
29. This research was performed at the Center for Engineering Systems Advanced Research, Oak Ridge National Laboratory. Funding provided by the Engineering Research Program of the Office of Basic Energy Sciences, and by the Advanced Computing Technology Initiative Program of the U.S. Department of Energy, under contract DE-AC05-96OR22464 with Lockheed Martin Energy Research Corporation. We thank Y. Y. Azmy, R. C. Mann, and T. H. Rowan for their comments and three unknown referees for their critical suggestions.

6 November 1996; accepted 11 March 1997

A Saturated X-ray Laser Beam at 7 Nanometers

J. Zhang,* A. G. MacPhee, J. Lin, E. Wolfrum, R. Smith, C. Danson, M. H. Key, C. L. S. Lewis, D. Neely, J. Nilsen, G. J. Pert, G. J. Tallents, J. S. Wark

A saturated nickel-like samarium x-ray laser beam at 7 nanometers has been demonstrated with an output energy of 0.3 millijoule in 50-picosecond pulses, demonstrating that saturated operation of a laser at wavelengths shorter than 10 nanometers can be achieved. The narrow divergence, short wavelength, short pulse duration, high efficiency, and high brightness of this samarium laser make it an ideal candidate for many x-ray laser applications.

An important objective in the development of x-ray lasers is to deliver a coherent, saturated output at wavelengths shorter

than 10 nm (1). Such saturated x-ray lasers are required for holography (2) and microscopy (3) of biological specimens and for deflectometry (4), interferometry (5), and radiography (6) of dense plasmas relevant to inertial confinement fusion and laboratory astrophysics (5). Saturated operation is important because it means that the maximum power possible for a given volume of excited plasma is extracted by the stimulated emission. Saturated x-ray lasers ensure an output energy sufficient for most applications and tend to produce a consistent output with little variation from shot to shot. Lasers are characterized by the product of the laser gain coefficient and the length of the laser region (gain-length product), and saturation requires a gain-

J. Zhang, E. Wolfrum, J. S. Wark, Clarendon Laboratory, Department of Physics, University of Oxford, Oxford, OX1 3PU, UK.

A. G. MacPhee and C. L. S. Lewis, School of Mathematics and Physics, Queen's University, Belfast, BT7 1NN, UK.

J. Lin, R. Smith, G. J. Tallents, Department of Physics, University of Essex, Colchester, CO4 3SQ, UK.
C. Danson and D. Neely, Central Laser Facility, Rutherford Appleton Laboratory, Chilton, Oxon, OX11 0QX, UK.
M. H. Key and J. Nilsen, Lawrence Livermore National Laboratory, Livermore, CA 94550, USA.
G. J. Pert, Department of Physics, University of York, York, YO1 5DD, UK.

*To whom correspondence should be addressed at Rutherford Appleton Laboratory, Chilton, OX11 0QX, UK. E-mail: j.zhang@rl.ac.uk