



UNIVERSITÀ
DEGLI STUDI
FIRENZE

FLORE

Repository istituzionale dell'Università degli Studi di Firenze

IBM DSD Report: Part B: Triplet Correlations in NCC water

Questa è la Versione finale referata (Post print/Accepted manuscript) della seguente pubblicazione:

Original Citation:

IBM DSD Report: Part B: Triplet Correlations in NCC water / P. Procacci; D. M. Soumpasis; G. Corongiu. - STAMPA. - (1991), pp. 16-43.

Availability:

This version is available at: 2158/777087 since:

Terms of use:

Open Access

La pubblicazione è resa disponibile sotto le norme e i termini della licenza di deposito, secondo quanto stabilito dalla Policy per l'accesso aperto dell'Università degli Studi di Firenze (<https://www.sba.unifi.it/upload/policy-oa-2016-1.pdf>)

Publisher copyright claim:

(Article begins on next page)



DSD

**PART A: ON THE COMPUTATION OF N-PARTICLE CORRELATIONS
IN CLASSICAL FLUIDS VIA COMPUTER SIMULATIONS.
PART B: TRIPLET CORRELATIONS IN NCC WATER.**

Part A: D. M. Soumpasis

Part B: P. Procacci, D. M. Soumpasis & G. Corongiu

October 30, 1991

Kingston, New York

Limited Distribution Notice

This report has been submitted for publication elsewhere and has been issued as a Research Report for early dissemination of its contents. As a courtesy to the intended publisher, it should not be widely distributed until after the date of outside publication.

Copies may be requested from:
IBM, Department 48B, Building 963
Neighborhood Road
Kingston, New York 12401

Part A

**On the Computation of n-Particle Correlations in Classical Fluids
via Computer Simulations**

D. M. Soumpasis*

IBM Corporation
Center for Scientific and Engineering Computations
Department 48B/ MS 428
Neighborhood Road
Kingston, New York 12401, USA

Part B

Triplet Correlations in NCC Water

P. Procacci, D. M. Soumpasis* and G. Corongiu

IBM Corporation
Center for Scientific and Engineering Computations
Department 48B/ MS 428
Neighborhood Road
Kingston, New York 12401, USA

* Permanent Address:
Max-Planck-Institut für Biophysikalische Chemie
Abteilung Molekulare Biologie
Postfach 2841
3400 Göttingen, FRG

Triplet Correlations in NCC Water

P. Procacci, D. M. Soumpasis* and G. Corongiu

IBM Corporation
Center for Scientific and Engineering Computations
Department 48B/ MS 428
Neighborhood Road
Kingston, New York 12401, USA

* Permanent Address:

Max-Planck-Institut für Biophysikalische Chemie
Abteilung Molekulare Biologie
Postfach 2841
3400 Göttingen, FRG

1. Introduction

Quantitative understanding of the properties of water and aqueous ions is an essential prerequisite for understanding and modelling a large number of physicochemical phenomena, including structural stabilities and functions of biomolecules.¹ In the course of the last two decades Clementi's group at IBM has developed a series of increasingly sophisticated water models based on *ab initio* computations,²⁻⁹ which have been extensively used in computer simulations (Monte Carlo and Molecular Dynamics) of liquid water,^{2-9, 17-19} ionic¹⁰⁻¹³ and DNA¹⁴⁻¹⁶ hydration. The most recent development is the NCC model¹⁷⁻¹⁹ which has a better short range behavior, incorporates many body polarization effects and has been shown to yield excellent results for many structural and dynamic properties of liquid water (e.g., pair correlations, scattering intensities, IR spectra, sound modes, specific heat, NMR relaxation timer, self-diffusion coefficient, etc.). The most complete characterization of the equilibrium structure in a fluid model consists in determining the hierarchy of site correlations i.e., the pair, triplet, quadruplet etc., correlation functions. However until now, the only correlation functions computed via simulation techniques for liquid water both in the studies mentioned above and in all simulations employing totally empirical water potentials are only the OO, OH and HH pair correlation functions. There are several reasons for this. Firstly, using the data of the current experimental techniques (x-rays and neutron scattering) it is not possible to obtain the correlations beyond the pair level. Secondly, triplet and higher order correlations are much more difficult to compute, parameterize and interpret. Thirdly, in our opinion, their importance for benchmarking proposed theoretical models, constructing semi-analytical theories for the description of complex systems (e.g., biomolecules in solution) and computing thermodynamic derivatives of the pair correlations (which involve integrals over higher order correlation functions) has not been adequately recognized yet.

In particular, high order correlations are required in order to construct refined versions of the potentials of mean force (PMF) approach²⁰⁻²⁷ which provides a unique, computationally feasible way to treat solvent effects on biomolecular structures in aqueous electrolytic environments. In "Part A" of this report, we have discussed the theoretical background for computing many particle correlations via computer simulations. Here, we report for the first time results for triplet correlations in NCC water.

2. Computational Procedure

The NCC potential and its application to liquid water simulation have been extensively discussed elsewhere.¹⁷⁻¹⁹ A sample of 512 water molecules at the experimental density of 0.997 gr/cm³, enclosed in a cubic box of sidelength 24.856 Å and subject to periodic boundary conditions, has been first equilibrated using the Molecular Dynamics method. The Ewald sums technique has been used for the long range interaction energy corrections. After equilibration, the Newtonian trajectory of the system has been computed for 32 ps with a time step of 0.5 femtoseconds and configurational parameters have been saved every four femtoseconds.

From 2000 different configuration, generated at time intervals of 16 femtoseconds, the OO, OH and HH pair correlations have been computed up to a distance of 8 Å using a discretization step $\Delta=0.2$ Å and standard techniques of counting pair distances. The OOO, HHH, OOH and HOH triplet correlations have been computed using a triple nested loop to count the number of triangles of specified vertices (e.g., OOO) and side lengths r, s, t . Each side length being in the range 0 to 8 Å with the same discretization step $\Delta=0.2$ Å as above. (Due to the periodic boundary conditions correct triplet counting is guaranteed only if

$r + s + t \leq L$ where L is the side length of the cubic simulation box, in our case $L = 24.856$ Å). As discussed in "Part A" of this report the xyz ($x, y, z = O, H$) triplet correlation function $g_{xyz}^{(3)}(r,s,t)$ where r, s, t are the distances between sites $x-y, x-z$ and $y-z$, respectively, is computed by means of the relation

$$g_{xyz}^{(3)}(r,s,t) = \frac{\langle N_{xyz}^{(3)}(r,s,t) \rangle}{8\pi^2 \rho_x \rho_y \rho_z V r s t \Delta^3} \quad (1)$$

where $N_{xyz}^{(3)}(r,s,t)$ is the number of xyz triangles of sides r, s, t , in a microstate (particle configuration) and $\langle \dots \rangle$ denotes a simulation average. V is the volume of the cubic simulation box and ρ_x, ρ_y, ρ_z the number density of site x, y, z , respectively.

3. Results and Discussion

Each of the 2000 configurations used in the averaging process contains a total number of 3.616×10^9 triplets. They are sorted into 64×10^3 bins (r,s,t) . The total numbers of counts in the most populated bins (*i.e.*, those associated with the most probable triplet configurations) at the end of the simulation, are in the range 10,000-40,000. The present triplet correlation study is not only the first one for water but also the most extensive one for any fluid system reported to date. Previous studies of hard spheres^{28, 29} and Lennard-Jones systems³⁰⁻³⁴ have only sampled a very limited set of triplet configurations (isosceles and equilateral triangles). Due to the fact that $g^{(3)}(r,s,t)$ is a function of three distance variables, graphic representations are only possible either in the case where one of the distances is held constant or when two or three of the distances are equal. The parameterization of triplet functions is also entirely non-trivial but preliminary results with bicubic-splines in conjunction with a confocal ellipsoidal coordinates representation look promising and further work in this direction is underway.

An important point which we have also investigated in this study is how well triplet functions can be represented as functions of the pair correlations involved (closure relations). Perhaps the most widely used ansatz in this direction, is Kirkwood's Superposition Approximation, KSA,³⁵ which in our case becomes

$$g_{xyz}^{(3)}(r,s,t) = g_{xy}^{(2)}(r)g_{xz}^{(2)}(s)g_{yz}^{(2)}(t) \quad (2)$$

with $(x,y,z : O, H)$.

We next present some of the results obtained so far. Results for pair correlations are not included since they have been previously discussed.^{18,19}

OOO-Correlations

The triplet function for isosceles configurations is depicted as a 3-dimensional plot in Figure 1 and as a contour plot in Figure 2. The maximal value of this function (6.352) occurs at $r = 4.60 \text{ \AA}$, $s = 2.80 \text{ \AA}$ and reflects the most probable configuration of three water molecules. In all contour plots regions where $g^{(3)}$ is high are marked by H and regions where it is low by L.

The position of the global maximum of the OOO isosceles triplet correlation in the r, s plane corresponds to the positions of the second and first peak of the OO pair correlation respectively. The triplet correlation decays rather fast and has values close to one in the region $s \geq 3.5\text{\AA}$, $t \geq 3.5\text{\AA}$, but it is quite structured even there.

Figure 3 shows a contour plot of the triplet function predicted by the KSA. The maximal value of $g_{ooo}^{(3)}$ is overestimated (8.080) and the position of the peak ($r = 3.00\text{\AA}$, $s = 3.00\text{\AA}$) is also inaccurate. The same is also true of the boundary of the contour plot which essentially defines the region in the r, s plane where triangles can form (*i.e.*, the triangle inequalities are obeyed). Comparison of Figures 2, and 3 shows that the KSA does not describe this boundary correctly for $r \geq 6.0\text{\AA}$, although it is reasonably accurate in a large portion of the region where the triplet function is close to one.

A more global test of the accuracy of the KSA is shown in Figure 4 where a representative sample of $g^{(3)}$ values not only for isosceles configurations but also for all other configurations is plotted against the corresponding values predicted by the KSA. If the KSA were exact all points in this plot should fall on the diagonal. It is seen that many points are near the diagonal, but for configurations with $g^{(3)} \geq 2$ and also for those close to the boundary mentioned above (lower quasi-linear cluster of points on the plot) the KSA fails.

HHH Correlations

Figures 5-7 depict the analogous results for isosceles HHH triplets. A global maximum (5.541) is attained at $r = 1.60\text{\AA}$, and $s = 2.60\text{\AA}$, the former value is close to the HH distance in a NCC water molecule (1.5139 \AA) whereas the latter value is very close to the position of the first (intermolecular) peak in the HH pair correlation function.

The extent to which the KSA fails (in a similar manner as discussed above) and the regions where it is approximately valid can be judged from Figures 7 and 8.

OHH Correlations

In this case the set of isosceles configurations of course includes the individual water molecules. This gives rise to a huge peak (512) which for technical reasons has been trun-

cated in producing Figures 9, 10, and 11. In the whole region $r \geq 3, s = t \geq 4$ this correlation is very flat and attains values close to one.

The KSA is again not correct for some configurations (compare Figures 10 and 11) but as can be seen in Figure 12, it is a quite good approximation for many other triplets.

OHO Correlations. Hydrogen bonding in liquid water

One of the most important properties of liquid water is its associative nature which is due to hydrogen bonding. Pair correlation functions do not directly reflect the hydrogen bonding properties of a given water model (e.g., the NCC) but triplet correlations are ideally suited for the quantitative study of this property in the liquid phase.

Figure 13 shows the OHO triplet correlation function at a fixed OO distance ($r = 3.0 \text{ \AA}$) and variable OH distances s and t . Figure 14 is a contour plot of the same data. The particular value of r chosen is close to 2.97 \AA which is the OO distance in the most stable (lowest energy) configuration of the isolated NCC water dimer.¹⁹ The correlation function is entirely symmetric with respect to the s and t variables as it should be.

Due to the symmetry, it suffices to discuss the structure of the correlation function in the region of the s, t plane below the diagonals in Figs. 13 and 14. There are two pronounced peaks at $t = 1.0 \text{ \AA}, s = 3.40 \text{ \AA}$ and $t = 1.0 \text{ \AA}, s = 2.00 \text{ \AA}$ of heights 23.893 and 20.0, respectively. The latter is due to the H involved in the nearly linear H bond of the isolated water dimer ($r_{OO} = 2.97 \text{ \AA}, \text{H}\hat{\text{O}}\text{H} = 174.102^\circ, t_{OH} = 0.9572 \text{ \AA}, s_{OH} = 2.0162 \text{ \AA}$) whereas the former is due to the second H on the H-donor side of the water-water dimer (compare Figure 1 of Ref. 19).

In Figures 13 and 14 one also clearly recognizes two secondary peak structures beyond the one discussed, albeit of much smaller height and less localized nature. They are due to the other two H-atoms in the water dimer. In the isolated dimer all these peaks should be essentially δ -functions, the main ones retain their sharpness but the subsidiary ones become flatter and less localized due to thermal and many body effects in the liquid water phase.

Figures 15 and 16 are the analogous OHO correlation plots for a shorter OO distance ($r = 2.8 \text{ \AA}$). The structure of the correlation function is entirely similar but the height of the maximal peak is somewhat bigger (26.374) compared to 23.893 and the s -value somewhat smaller (3.20 \AA instead of 3.40 \AA)

4. Concluding Remarks

Computation of the triplet correlations in liquid water provides a wealth of information on its equilibrium structural and thermodynamic properties which cannot be obtained using pair correlations alone. We are currently studying efficient ways to parameterize the enormous amount of numerical data obtained, test other closures (e.g., the hypernetted chain closure) and compute certain integrals over the triplet correlation functions which determine experimentally known thermodynamic derivatives of the pair correlation function, (*i.e.*, the pressure derivative). These developments will provide further sensitive tests for the accuracy of the models used and enable us to use the information obtained in biophysical applications (e.g., the hydration of biomolecular structures).

We wish to thank R. Klement at the MPI, Goettingen, FRG, for his great help in producing the graphs.

References

1. Franks, F. *Water: A Comprehensive Treatise*, Plenum, New York, Vol. 1-6, (1971-1979).
2. Matsuoka, O., Clementi, E., Yoshimine, M., *J. Chem. Phys.* **64**, 1351 (1976).
3. Lie, G. C., Clementi, E., Yoshimine, M., *J. Chem. Phys.* **64**, 2314 (1976).
4. Clementi, E., Corongiu, G., *Int. J. Quantum Chem. Symp.* **10**, 31 (1983).
5. Wojcik, M., Clementi, E., *J. Chem. Phys.* **84**, 5970 (1986).
6. Wojcik, M., Clementi, E., *J. Chem. Phys.* **85**, 3544 (1986).
7. Wojcik, M., Clementi, E., *J. Chem. Phys.* **85**, 6085 (1986).
8. Detrich, J. H., Corongiu, G., Clementi, E., *Chem. Phys. Lett.* **112**, 476 (1984).
9. Lie, G. C., Clementi, E., *Phys. Rev. A* **33**, 2679 (1986).
10. Clementi, E., Barsotti, R., *Chem. Phys. Lett.* **59**, 21 (1978).
11. Mezei, M., Beveridge, D. L., *J. Chem. Phys.* **74**, 6902 (1981).
12. Impey, R. W., Madden, P.A., McDonald, I. R., *J. Phys. Chem.* **87**, 5071 (1983).
13. Migliore, M., Corongiu, G., Clementi, E., Lie, G.C., *J. Chem. Phys.* **88**, 7766 (1988);
Corongiu, G., Migliore, M., Clementi, E., *J. Chem. Phys.* **90**, 4629 (1989).
14. Clementi, E., Corongiu, G., *Biopolymers* **20**, 551 (1981).
15. Clementi, E., Corongiu, G., *Biopolymers* **20**, 2427 (1981).
16. Clementi, E., Corongiu, G., *Biopolymers* **21**, 763 (1982).
17. Niesar, U., Corongiu, G., Huang, M., Dupuis, M., Clementi, E., *Int. J. Quant. Chem., Quant. Chem. Symp.* **23**, 421 (1989).
18. Clementi, E., Ed., *Modern Techniques in Computational Chemistry MOTTECC-90*, ESCOM, Leiden, (1990).
19. Niesar, U., Corongiu, G., Clementi, E., Kneller, G. R., Bhattacharya, D. K., *J. Phys. Chem.* **94**, 7949 (1990).
20. Soumpasis, D. M., *Proc. Natl. Acad. Sci. USA*, **81**, 5116 (1984).
21. Soumpasis, D. M., Wiechen, J., Jovin, T. M., *J. Biom. Struct. Dyn.* **4**, 535 (1987).
22. Soumpasis, D. M., *J. Biom. Struct. Dyn.* **6**, 563 (1988).
23. Garcia, A., Soumpasis, D. M., *Proc. Natl. Acad. Sci. USA* **86**, 3160 (1989).
24. Klement, R., Soumpasis, D. M., von Kitzing, E., Jovin T. M., *Biopolymers* **29**, 1089 (1990).

25. Soumpasis, D. M., Garcia, A., Klement, R., Jovin, T. M., *Theoretical Biochemistry and Molecular Biophysics*, Beveridge, D. L., and Lavery, R., Eds., Adenine Press, Guilderland, N.Y., 343 (1991).
26. Klement, R., Soumpasis, D.M., Jovin, T.M., *Proc. Natl. Acad. Sci. USA* **88**, (in press), (1991).
27. Soumpasis, D.M., *Computation of Biomolecular Structures. Achievements, Problems and Perspectives*, Soumpasis, D.M., and Jovin, T.M., Eds., Springer Verlag, Heidelberg, (in press), (1991).
28. Alder, B. J., *Phys. Rev. Lett.* **12**, 317 (1964).
29. Uehara, Y., Lee, Y.T., Ree, T., Ree, F. H., *J. Chem. Phys.* **70**, 1884 (1979).
30. Rahaman, A., *Phys. Rev. Lett.* **12**, 575 (1964).
31. Krumhansl, J. A., Wang, S.S., *J. Chem. Phys.* **76**, 2034 (1972).
32. Wang, S. S., Krumhansl, J. A., *J. Chem. Phys.* **56**, 4287 (1972).
33. Raveché, H. J., Mountain, R. D., Street, W. B., *J. Chem. Phys.* **57**, 4999 (1972).
34. McNeil, W. J., Madden, W. E., Haymet, A.D., Rice, S. A., *J. Chem. Phys.* **78**, 388 (1983).
35. Kirkwood, J., *J. Chem. Phys.* **3**, 300 (1935).

Figure Captions

Figure 1. The OOO correlation function for isosceles triplets in NCC water.

Figure 2. Contour plot of the OOO correlation function (isosceles triplets).

Figure 3. Contour plot of the OOO correlation function for isosceles triplets as predicted by the KSA.

Figure 4. Global test of the accuracy of the KSA for OOO triplets.

Figure 5. The HHH correlation function for isosceles triplets in NCC water.

Figure 6. Contour plot of the HHH correlation function (isosceles triplets).

Figure 7. Contour plot of the HHH correlation function for isosceles triplets as predicted by the KSA.

Figure 8. Global test of the accuracy of the KSA for HHH triplets.

Figure 9. The OHH correlation function for isosceles triplets in NCC water.

Figure 10. Contour plot of the OHH correlation function (isosceles triplets).

Figure 11. Contour plot of the OHH correlation function for isosceles triplets as predicted by the KSA.

Figure 12. Global test of the accuracy of the KSA for OHH triplets.

Figure 13. The OHO correlation function for one OO distance $r = 3.0\text{\AA}$.

Figure 14. Contour plot of the data in Figure 13.

Figure 15. The OHO correlation function for one OO distance $r = 2.8\text{\AA}$.

Figure 16. Contour plot of the data in Figure 15.

O-O-O iso

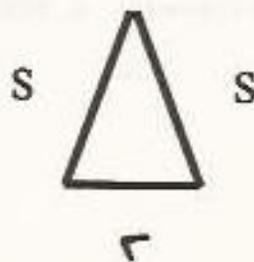
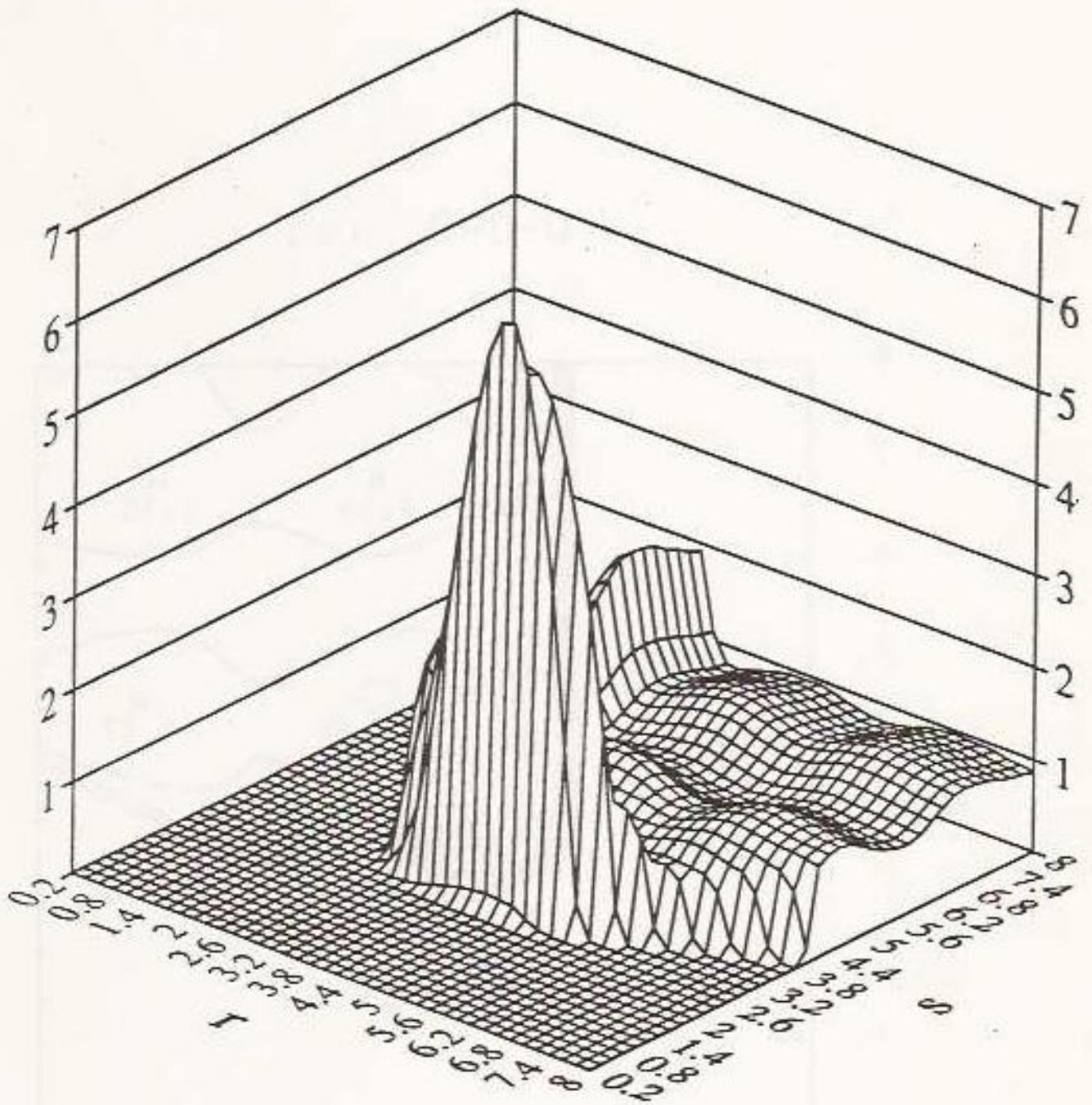
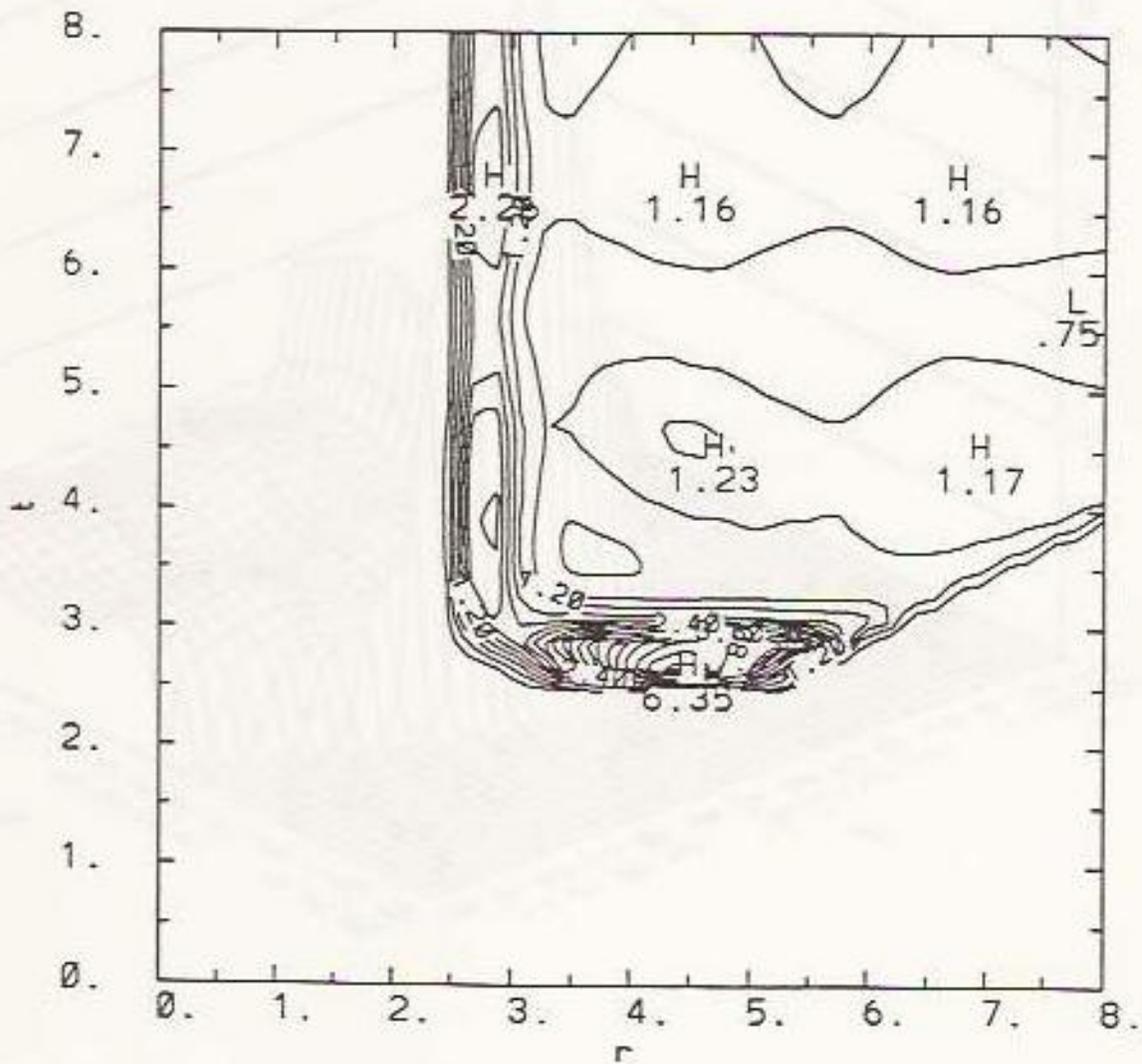


Fig. 1

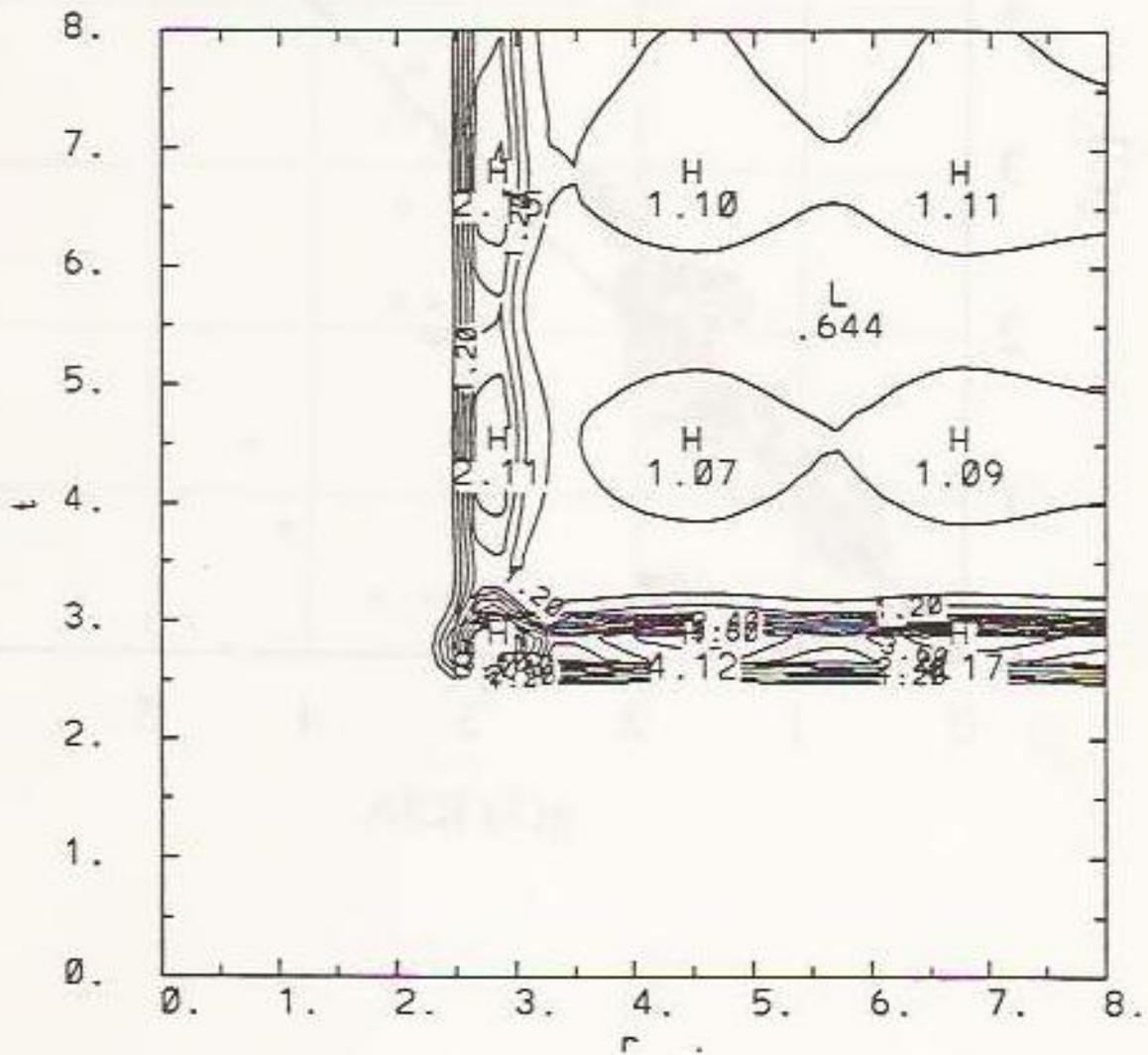
0-0-0 iso



Maximum: 6.352 ($r = 4.60, t = 2.80$)

Fig. 2

0-0-0 iso KSA



Maximum: 8.000 (r= 3.00, t= 3.00)

Fig. 3.

0-0-0

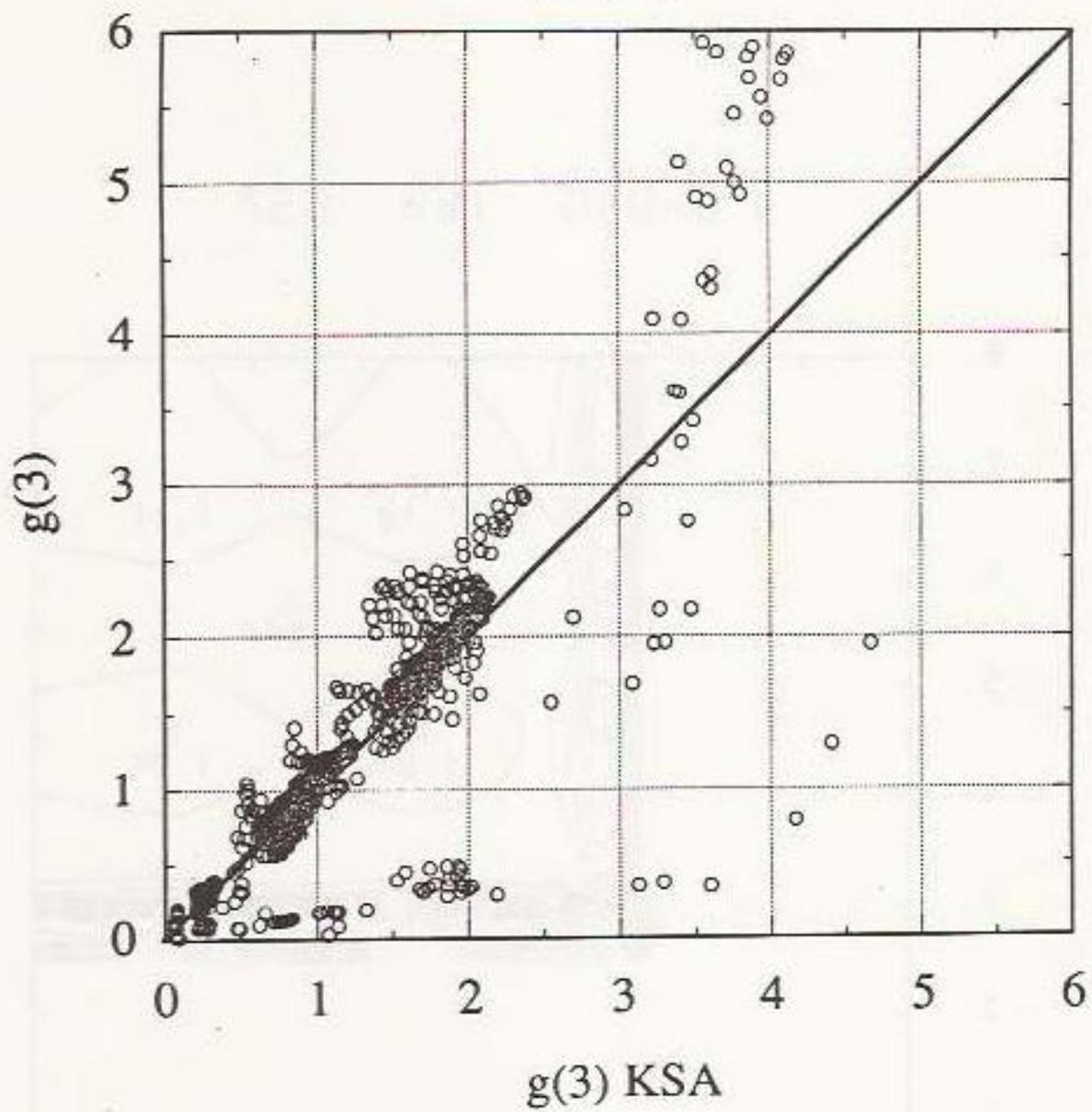


Fig. 4.

H-H-H iso

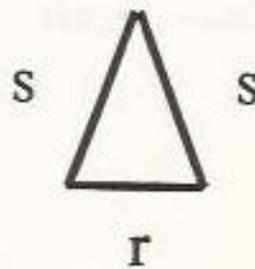
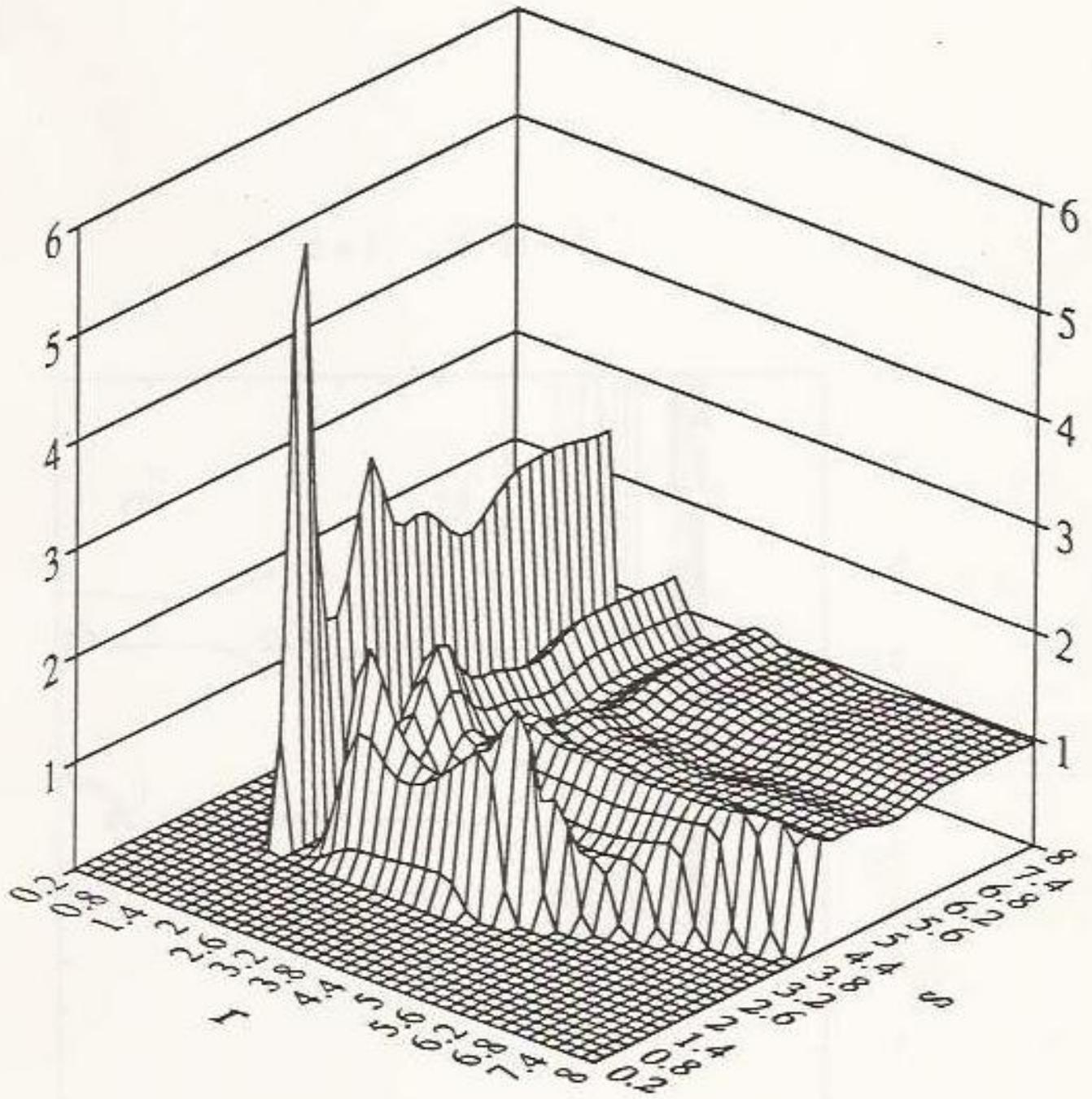
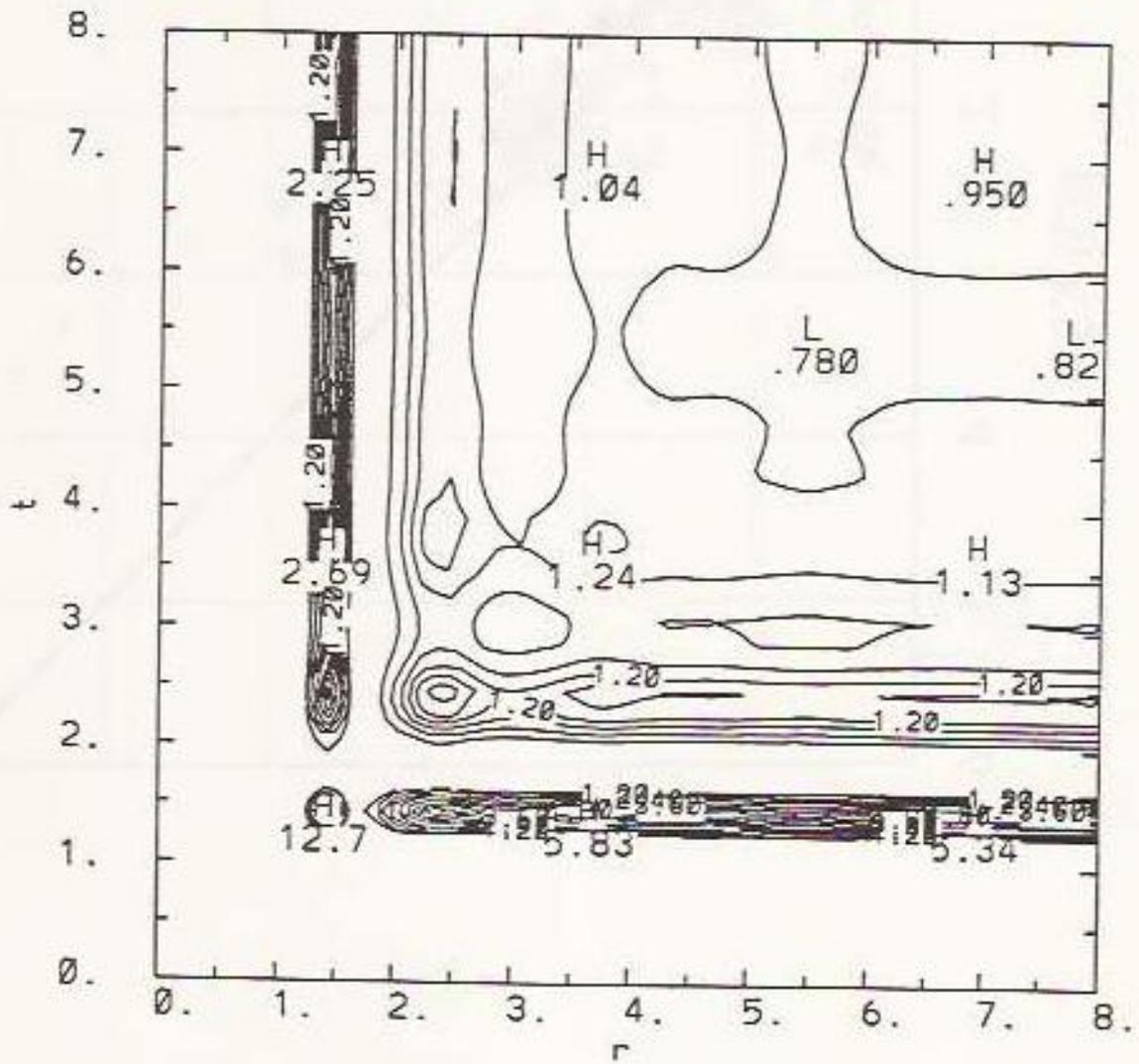


Fig. 5

H-H-H iso KSA



Maximum: 12.664 (r= 1.60, t= 1.60)

fig. 7

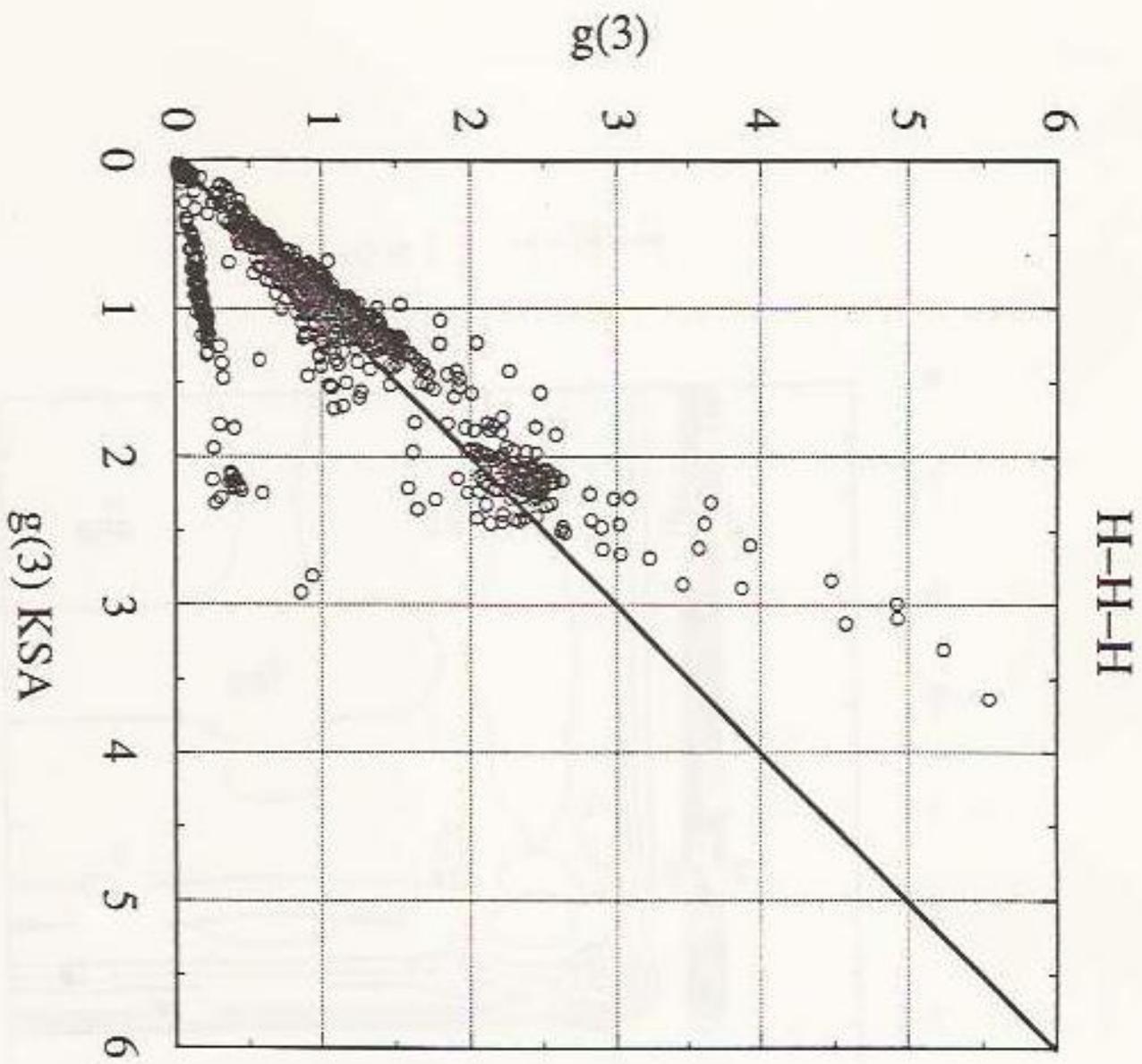


Fig. 8

O-H-H iso

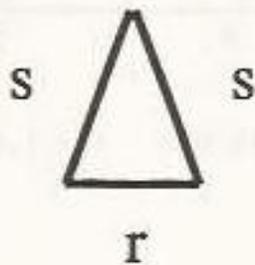
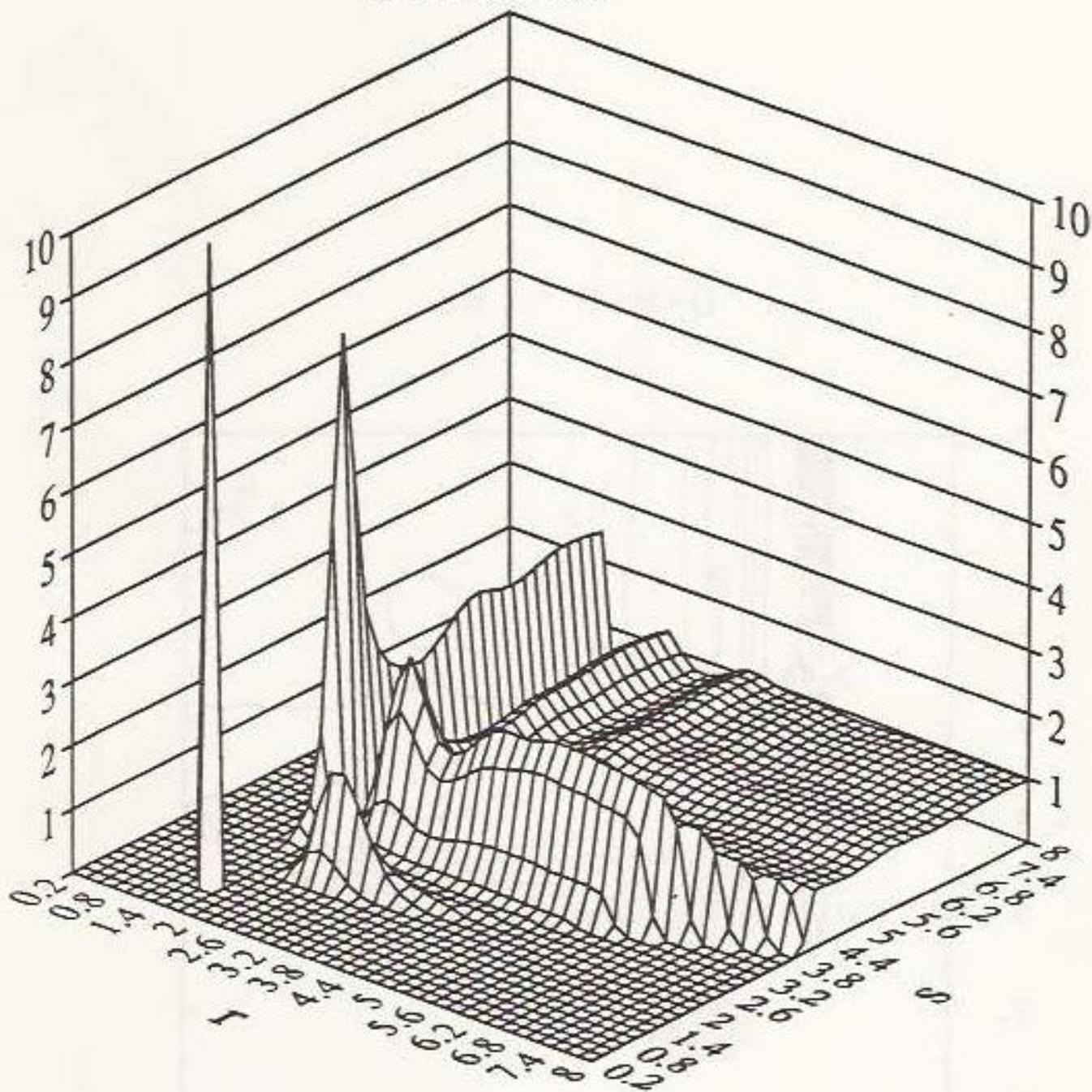
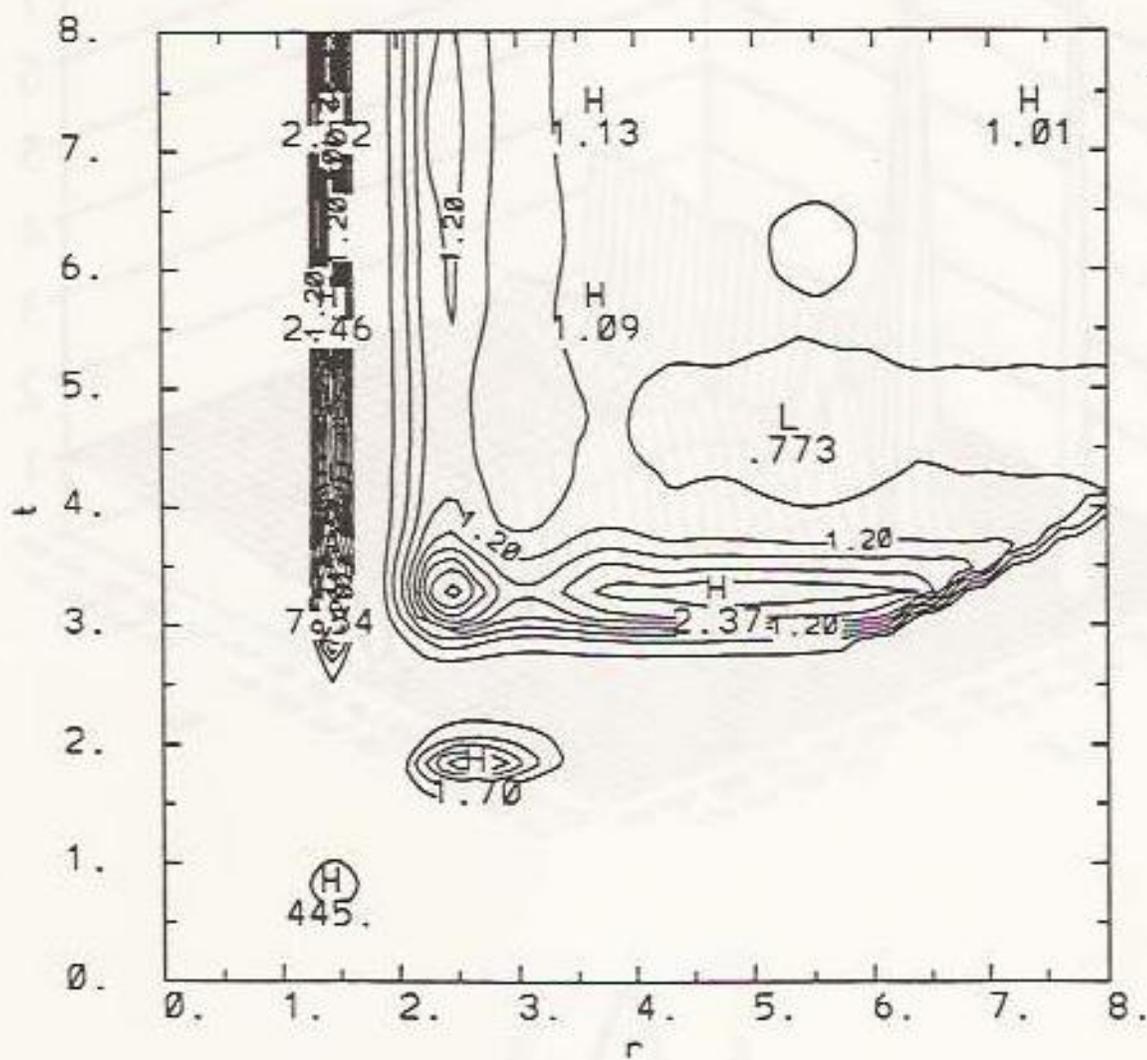


Fig 9.

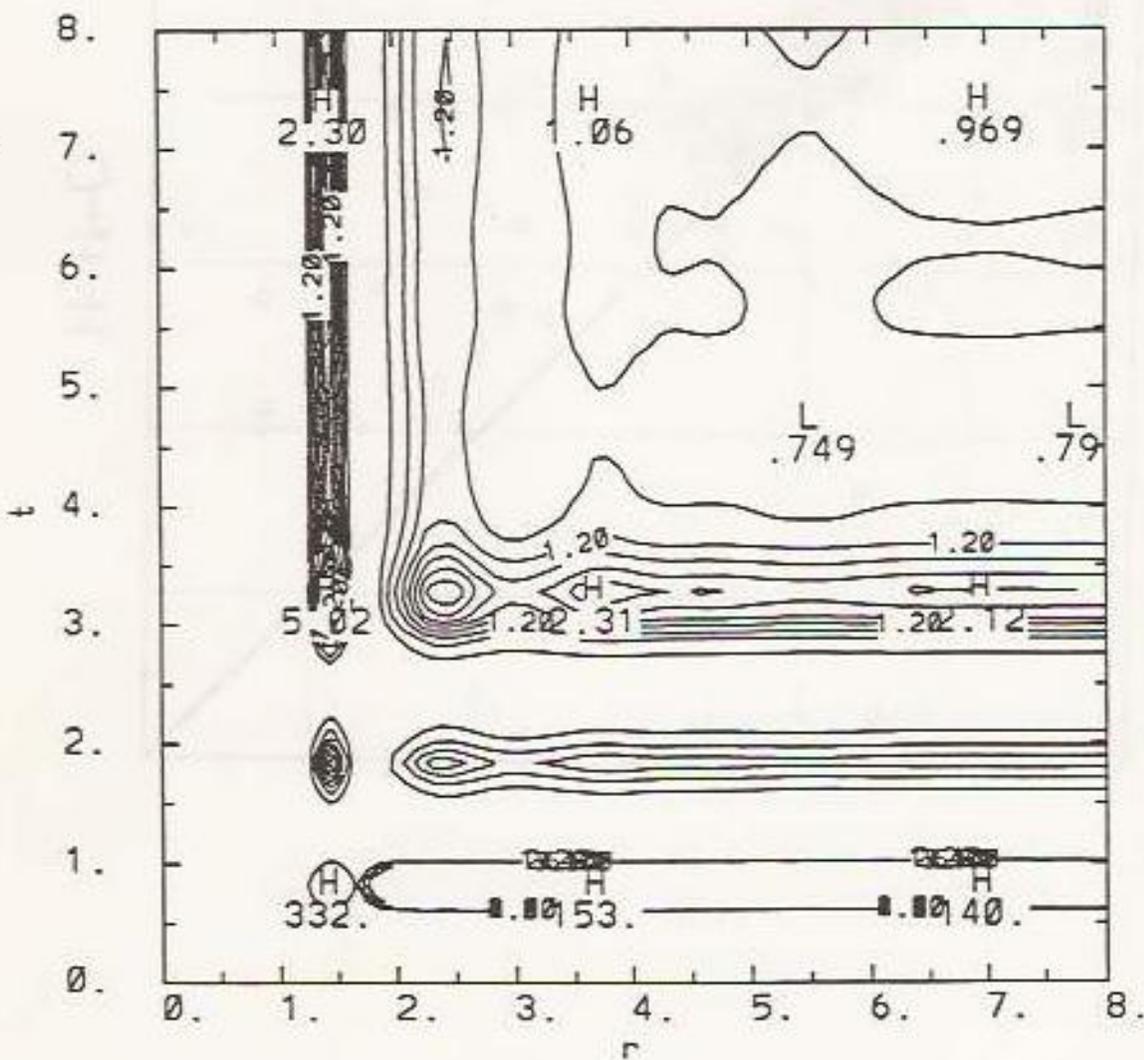
O-H-H iso



Maximum: 445.082 (r= 1.60, t= 1.00)

Fig 10.

O-H-H iso KSA



Maximum: 331.982 (r= 1.60, t= 1.00)

Fig 11.

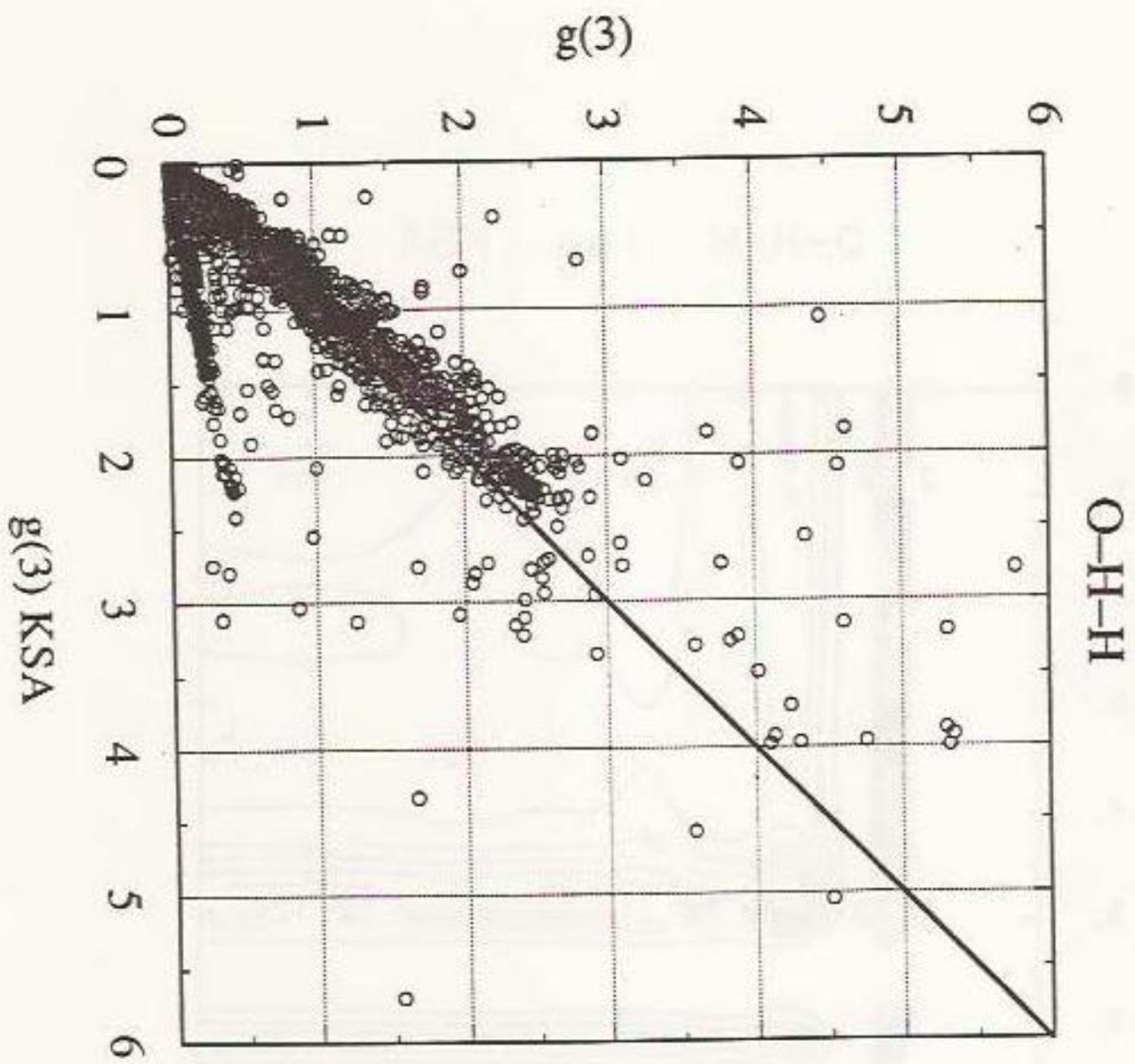


Fig 12.

OHO H-Bonding $r=3.0\text{\AA}$

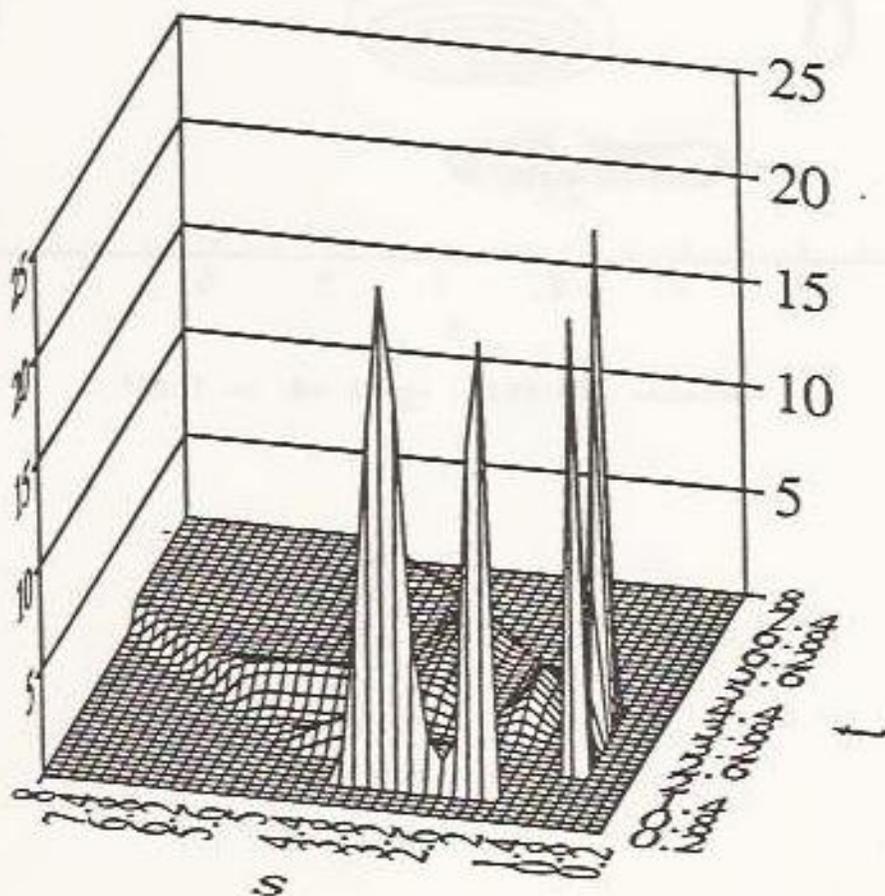
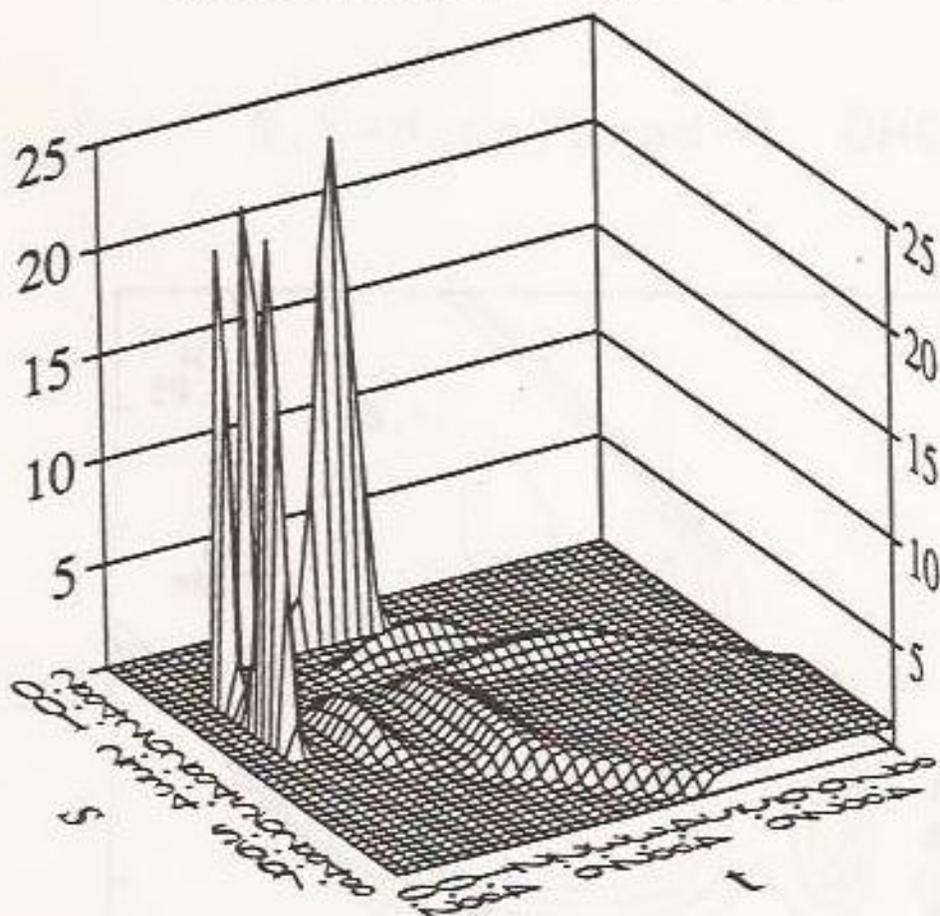


Fig 13

OHO H-bonding R=3.0

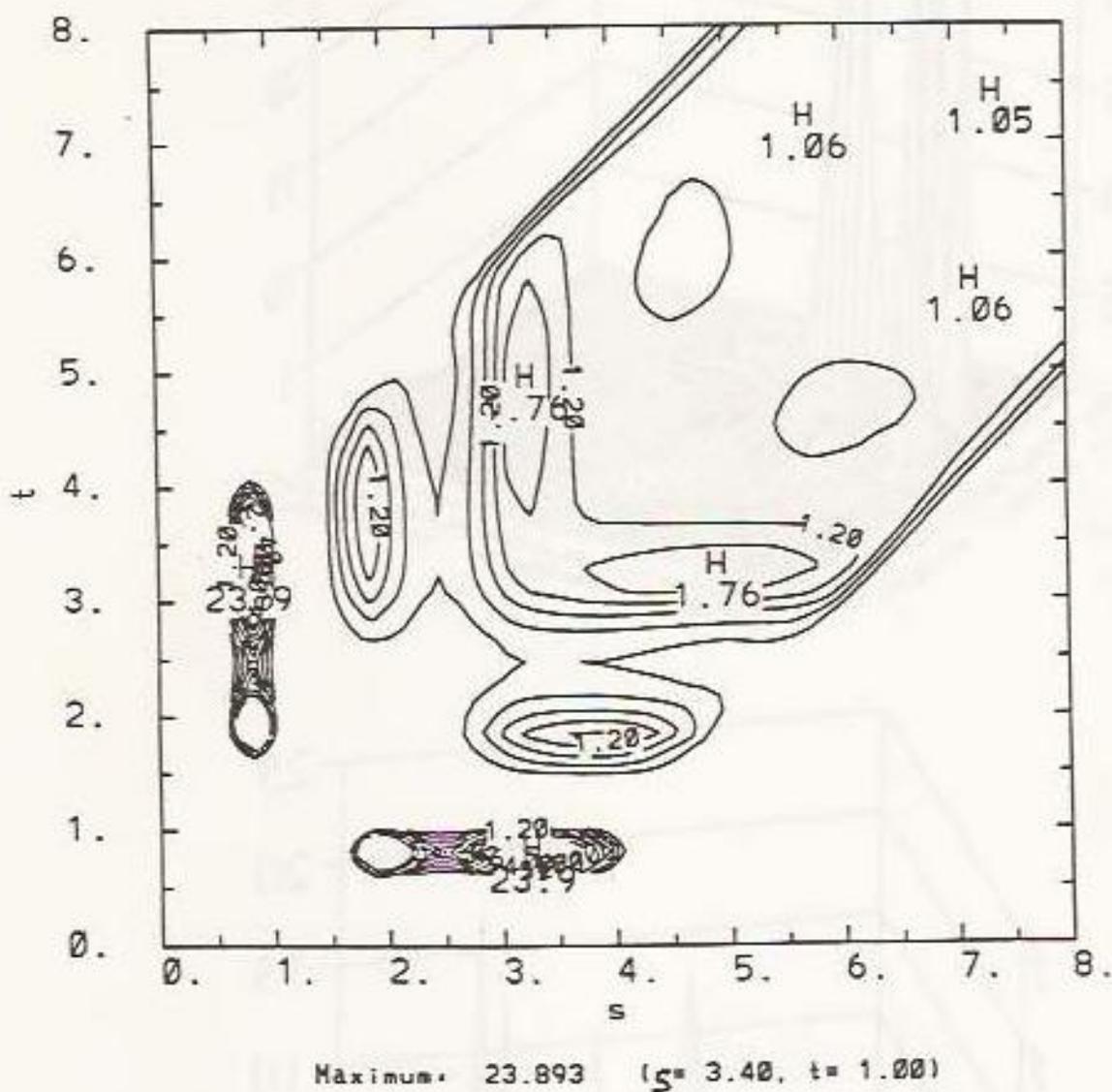


Fig. 14

OHO H-Bonding $r=2.8\text{\AA}$

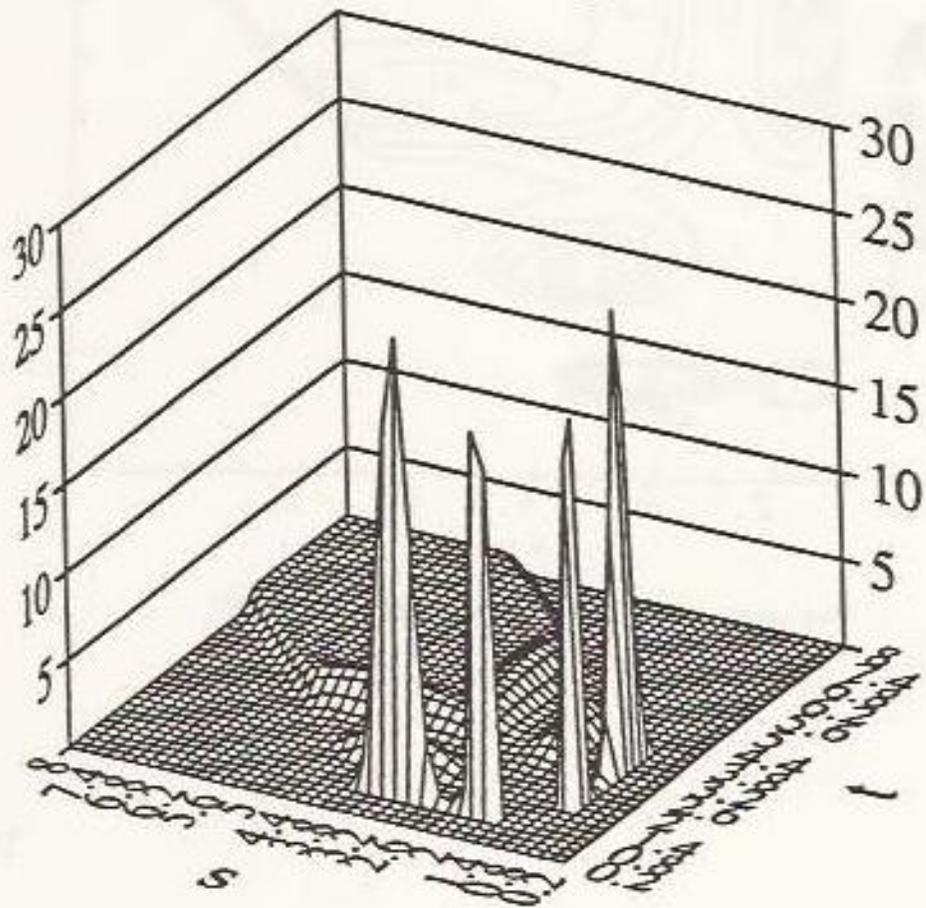
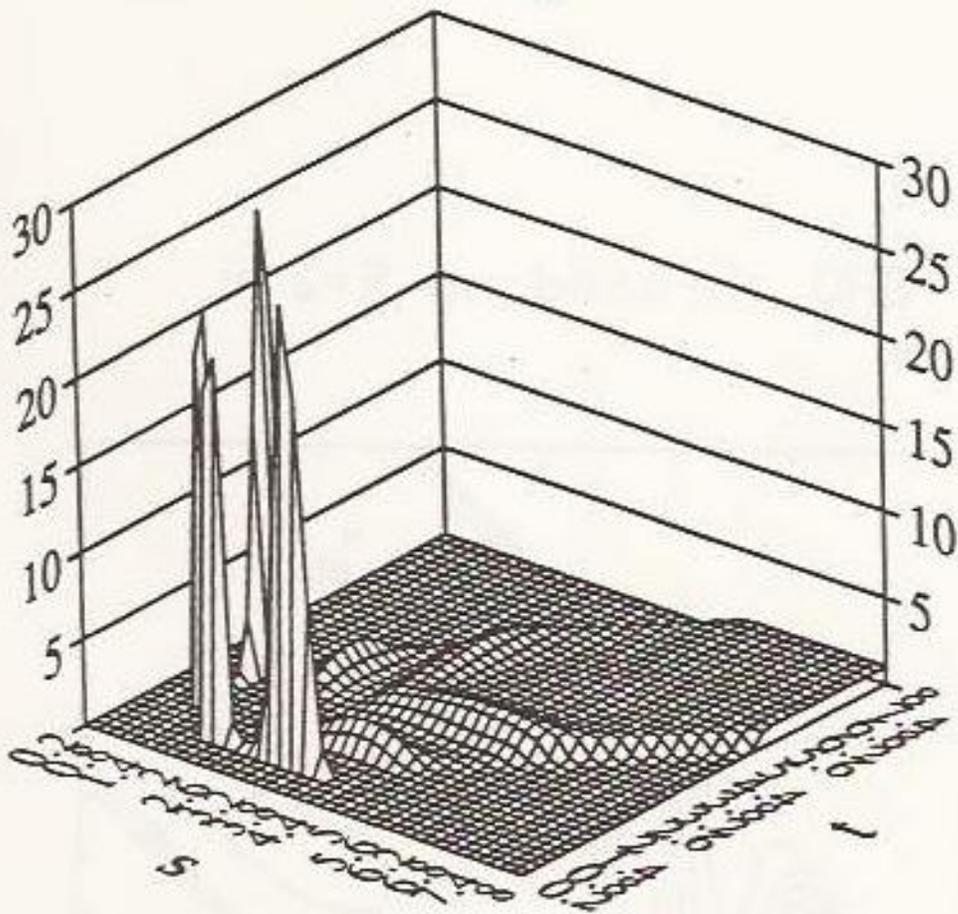
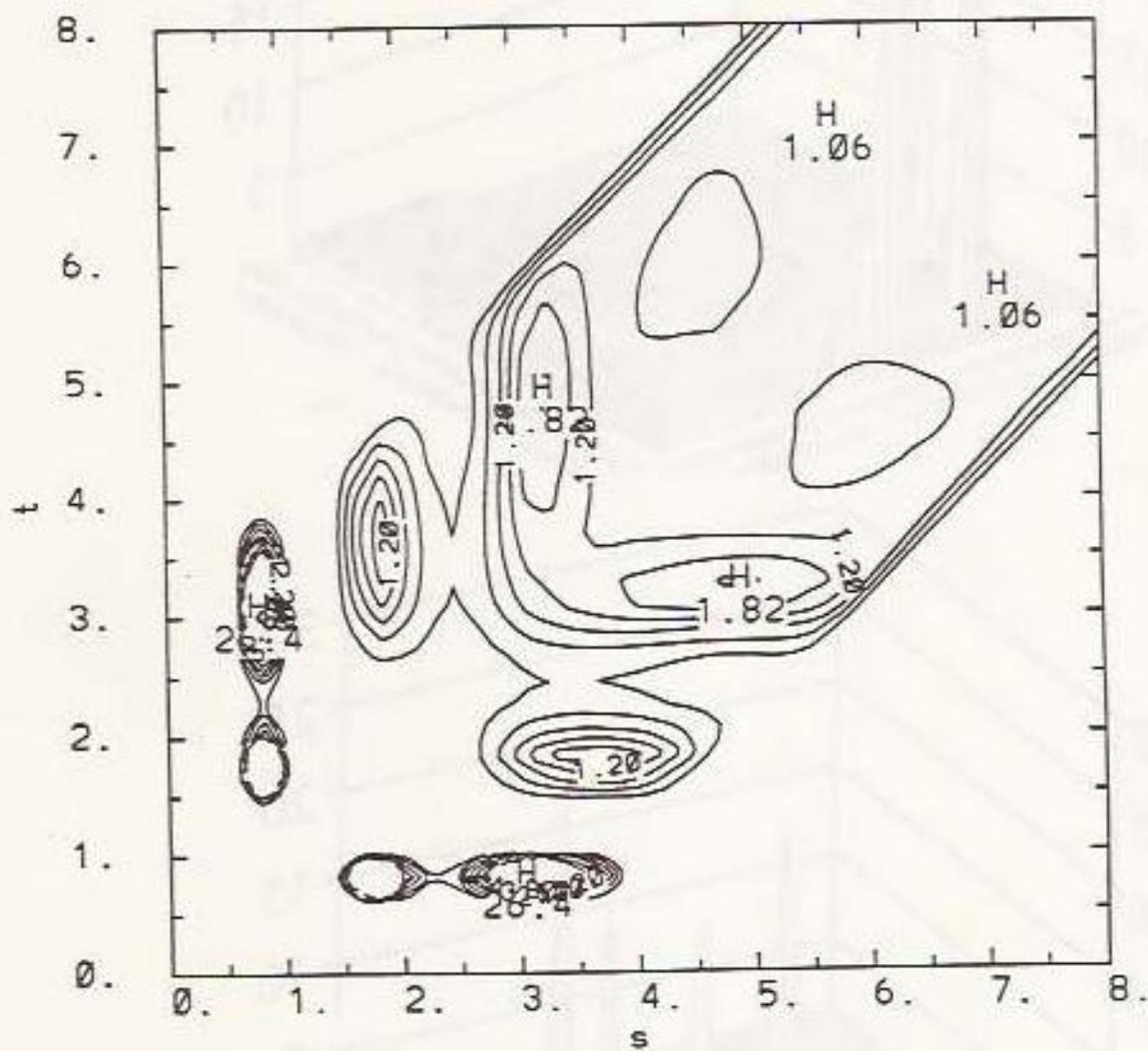


Fig. 15

OHO H-bonding $R=2.8$



Maximum: 26.374 ($s=3.20$, $t=1.00$)

Fig 16