

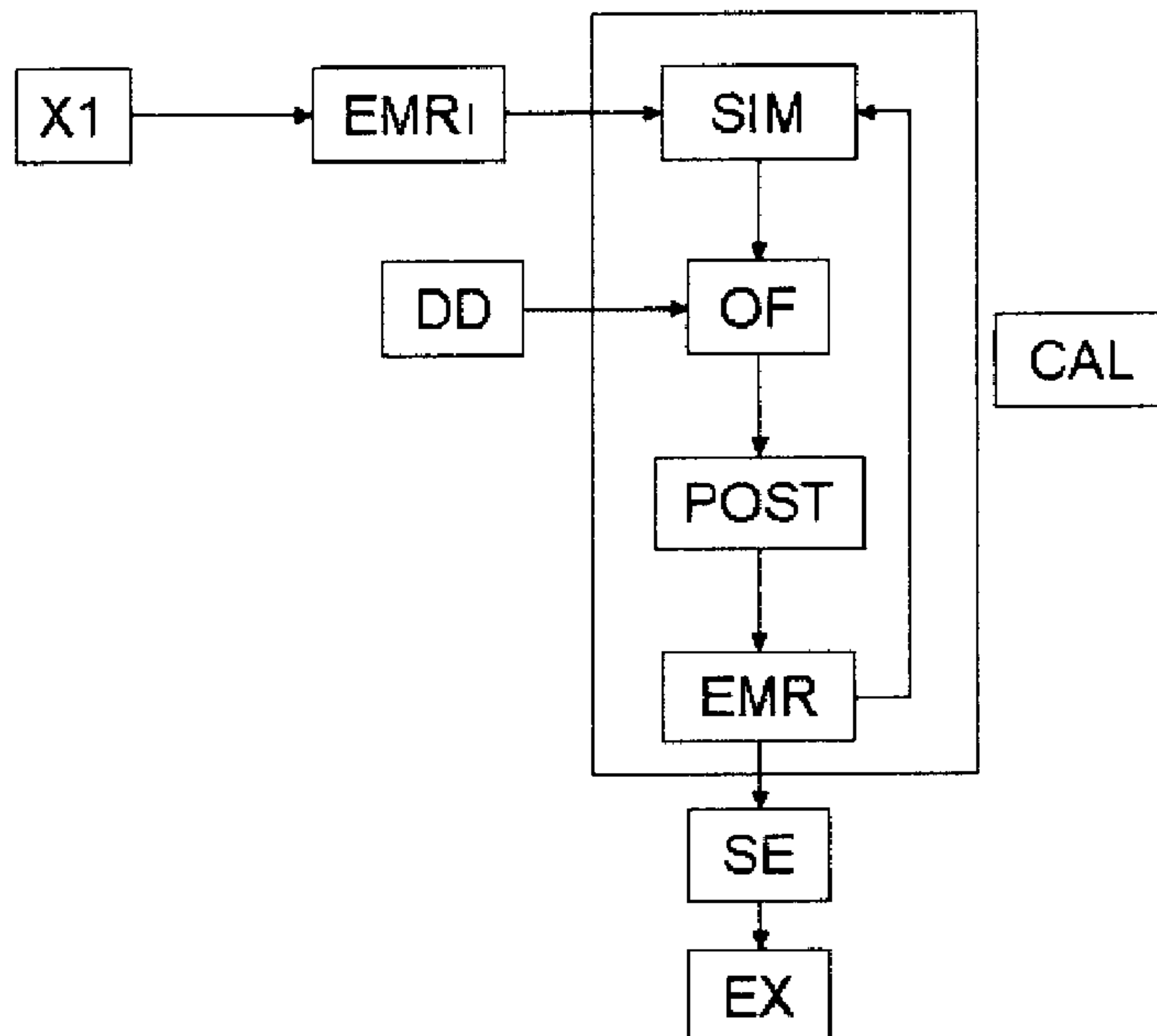


(22) Date de dépôt/Filing Date: 2013/07/12  
(41) Mise à la disp. pub./Open to Public Insp.: 2014/01/13  
(30) Priorité/Priority: 2012/07/13 (FR12 02 005)

(51) Cl.Int./Int.Cl. *E21B 43/00* (2006.01),  
*G06F 17/10* (2006.01), *G06F 19/00* (2011.01)  
(71) Demandeur/Applicant:  
IFP ENERGIES NOUVELLES, FR  
(72) Inventeurs/Inventors:  
DA VEIGA, SEBASTIEN, FR;  
GERVAIS-COUPLET, VERONIQUE, FR;  
FERAILLE, MATHIEU, FR  
(74) Agent: ROBIC

(54) Titre : PROCÉDE D'EXPLOITATION D'UN RESERVOIR GEOLOGIQUE A PARTIR D'UN MODELE DE RESERVOIR  
CALE PAR LE CALCUL D'UNE LOI ANALYTIQUE DE DISTRIBUTION CONDITIONNELLE DE PARAMETRES  
INCERTAINS DU MODELE

(54) Title: PRODUCTION PROCESS FOR A GEOLOGICAL RESERVOIR BASED ON A STOCK RESERVOIR BY  
CALCULATING AN ANALYTICAL CONDITIONAL DISTRIBUTION LAW OF UNCERTAIN PARAMETERS OF THE  
MODEL



(57) **Abrégé/Abstract:**

L'invention concerne un procédé d'exploitation d'un réservoir géologique qui peut consister à l'exploitation d'un réservoir d'hydrocarbures ou d'un réservoir de stockage de gaz (par exemple de CO<sub>2</sub>). L'exploitation (EX) du réservoir géologique est réalisée selon un schéma d'exploitation (SE) défini à partir d'un modèle de réservoir (MR), ledit modèle de réservoir étant calé (CAL) par rapport à des données dynamiques (DD). Selon l'invention, le calage (CAL) est réalisé en calculant une fonction objectif (OF) pour un ensemble de modèles de réservoir. Ladite fonction objectif (OF) permet de déterminer une loi analytique de distribution conditionnelle (POST) des paramètres incertains dudit modèle, cette loi est utilisée ensuite pour générer de nouveaux



(57) **Abrégé(suite)/Abstract(continued):**

modèles de réservoir ajoutés audit ensemble de modèles de réservoir.

### **Abrégé descriptif**

L'invention concerne un procédé d'exploitation d'un réservoir géologique qui peut consister à l'exploitation d'un réservoir d'hydrocarbures ou d'un réservoir de stockage de gaz (par exemple de CO<sub>2</sub>). L'exploitation (EX) du réservoir géologique est réalisée selon un schéma d'exploitation (SE) défini à partir d'un modèle de réservoir (MR), ledit modèle de réservoir étant calé (CAL) par rapport à des données dynamiques (DD). Selon l'invention, le calage (CAL) est réalisé en calculant une fonction objectif (OF) pour un ensemble de modèles de réservoir. Ladite fonction objectif (OF) permet de déterminer une loi analytique de distribution conditionnelle (POST) des paramètres incertains dudit modèle, cette loi est utilisée ensuite pour générer de nouveaux modèles de réservoir ajoutés audit ensemble de modèles de réservoir.

**PROCÉDÉ D'EXPLOITATION D'UN RÉSERVOIR GÉOLOGIQUE À PARTIR D'UN  
MODÈLE DE RÉSERVOIR CALÉ PAR LE CALCUL D'UNE LOI ANALYTIQUE DE  
DISTRIBUTION CONDITIONNELLE DE PARAMÈTRES INCERTAINS DU MODÈLE**

La présente invention concerne le domaine technique de l'industrie pétrolière, et plus  
5 particulièrement l'exploitation de réservoirs souterrains, tels que des réservoirs pétroliers ou  
des sites de stockage de gaz.

L'optimisation et l'exploitation d'un gisement pétrolier reposent sur une description  
aussi précise que possible de la structure, des propriétés pétrophysiques, des propriétés des  
fluides, etc., du gisement étudié. Pour ce faire, les spécialistes utilisent un outil qui permet de  
10 rendre compte de ces aspects de façon approchée : le modèle de réservoir. Un tel modèle  
constitue une maquette du sous-sol, représentative à la fois de sa structure et de son  
comportement. Généralement, ce type de maquette est représenté sur un ordinateur, et l'on  
parle alors de modèle numérique. Un modèle de réservoir comporte un maillage ou grille,  
généralement tridimensionnelle, associée à une ou plusieurs cartes de propriétés  
15 pétrophysiques (porosité, perméabilité, saturation...). L'association consiste à attribuer des  
valeurs de ces propriétés pétrophysiques à chacune des mailles de la grille.

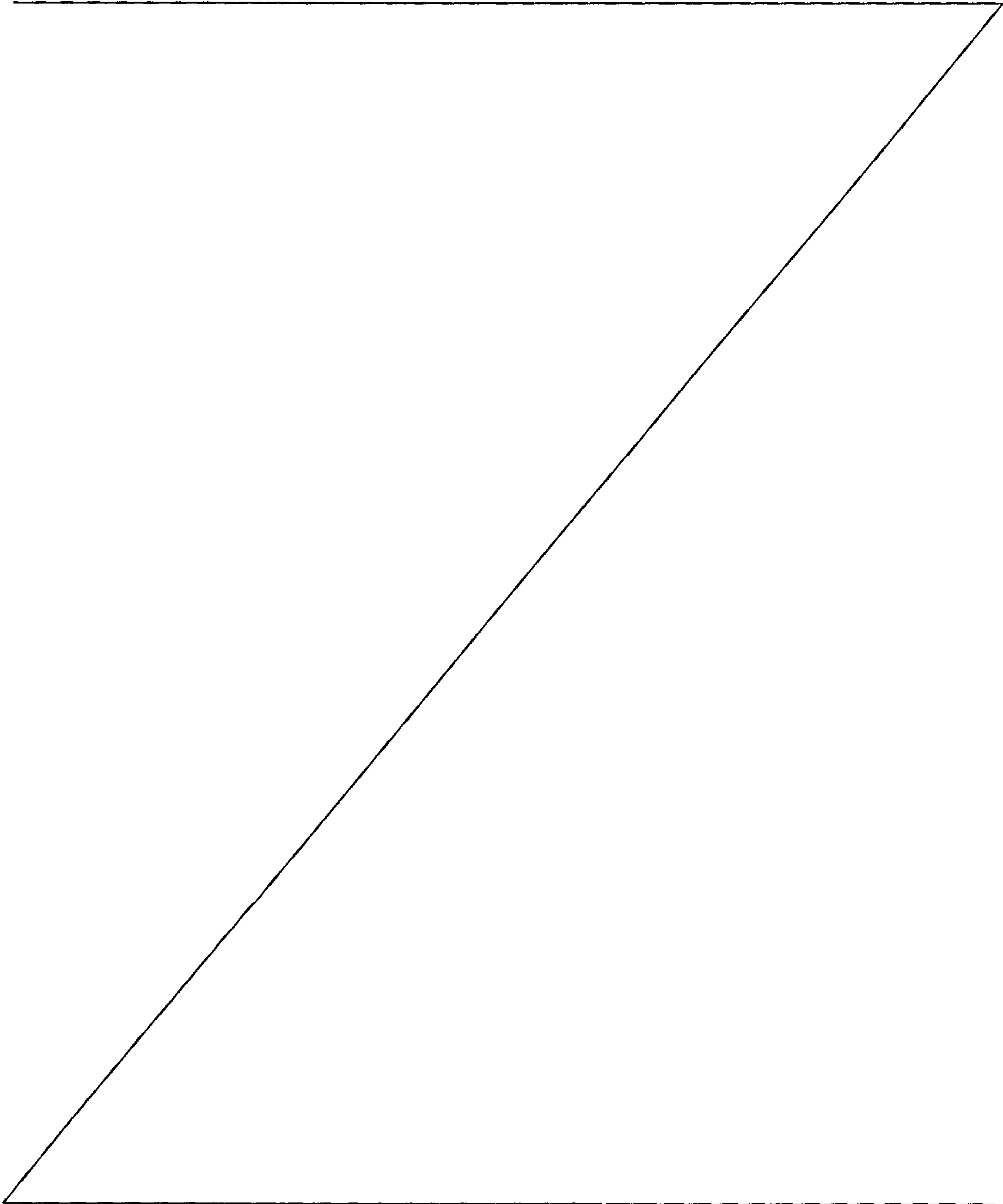
Ces modèles bien connus et largement utilisés dans l'industrie pétrolière, permettent  
de déterminer de nombreux paramètres techniques relatifs à l'étude ou l'exploitation d'un  
réservoir, d'hydrocarbures par exemple. En effet, puisque le modèle de réservoir est  
20 représentatif de la structure du réservoir et de son comportement, un spécialiste peut  
l'utiliser par exemple pour déterminer les zones qui ont le plus de chances de contenir des  
hydrocarbures, les zones dans lesquelles il peut être intéressant/nécessaire de forer un puits  
d'injection ou de production pour améliorer la récupération des hydrocarbures, le type  
d'outils à utiliser, les propriétés des fluides utilisés et récupérés.... Ces interprétations de  
25 modèles de réservoir en termes de « paramètres techniques d'exploitation » sont bien  
connues des spécialistes. De la même façon, la modélisation des sites de stockages de CO<sub>2</sub>  
permet de surveiller ces sites, de détecter des comportements inattendus et de prédire le  
déplacement du CO<sub>2</sub> injecté.

Afin de réaliser ces différentes interprétations, les spécialistes ont besoin de connaître  
30 le comportement dynamique de leur modèle de réservoir. Ceci est réalisé en simulant les  
écoulements de fluides au sein d'un réservoir au moyen d'un logiciel appelé simulateur  
d'écoulement. C'est ce qui est appelé une simulation de réservoir. Une ou plusieurs  
simulations de réservoir peuvent être nécessaires. Le logiciel PumaFlow ® (IFP Energies

1a

nouvelles, France) est un exemple de simulateur d'écoulement utilisé dans le domaine de l'exploitation des réservoirs souterrains.

Ces interprétations sont réalisées à l'aide d'un modèle de réservoir que l'on souhaite le plus représentatif possible, c'est-à-dire cohérent avec toutes les données disponibles. Ces  
5 données comprennent en général :



- des mesures en certains points de la formation géologique de la propriété modélisée, par exemple dans des puits. Ces données sont dites statiques car elles sont invariables dans le temps (à l'échelle des temps de la production du réservoir) et sont directement liées à la propriété d'intérêt, et
- 5 - des "données d'historique", comprenant des données de production, par exemple les débits de fluide mesurés aux puits ou les concentrations de traceurs et des données issues de campagnes d'acquisition sismique répétées à des temps successifs, dites de sismique 4D. Ces données sont dites dynamiques car elles évoluent en cours d'exploitation et sont indirectement liées aux propriétés attribuées aux mailles du
- 10 modèle de réservoir.

Le nombre de données statiques disponibles étant très faible par rapport au nombre de mailles du modèle de réservoir, des extrapolations sont nécessaires. Par exemple des méthodes probabilistes peuvent être utilisées pour remplir le modèle en propriétés pétrophysiques. Dans ce cas les données statiques disponibles sont utilisées pour définir

15 des fonctions aléatoires pour chaque propriété pétrophysique comme la porosité ou la perméabilité. Une représentation de la répartition spatiale d'une propriété pétrophysique est une réalisation d'une fonction aléatoire. De façon générale, une réalisation est générée à partir d'une part d'une moyenne, d'une variance et d'une fonction de covariance qui caractérise la variabilité spatiale de la propriété étudiée et d'autre part d'un germe ou d'une

20 série de nombres aléatoires. De nombreuses techniques de simulation, dites simulations géostatistiques, existent comme la méthode de simulation séquentielle Gaussienne, la méthode de Cholesky ou encore la méthode de transformée de Fourier rapide avec moyenne mobile FFT-MA. Ces techniques sont décrites notamment dans les documents suivants :

- 25 - Goovaerts, P., 1997, Geostatistics for natural resources evaluation, Oxford Press, New York, 483 p.
- Le Ravalec, M., Noëtinger B., and Hu L.-Y., 2000, The FFT moving average (FFT-MA) generator: an efficient numerical method for generating and conditioning Gaussian simulations, *Mathematical Geology*, 32(6), 701-723.

30 La construction d'un modèle de réservoir représentatif se fait par étapes. Tout d'abord, on construit un modèle de réservoir sur la base des données statiques. Puis ce modèle est mis à jour afin de reproduire au mieux les données d'historique via la simulation d'écoulement tout en conservant la cohérence avec les données statiques.

Les techniques d'intégration des données dynamiques (données de production et/ou sismique 4D) dans un modèle de réservoir sont bien connues des spécialistes : ce sont des techniques dites de "calage d'historique" ("history-matching" en anglais).

Le calage d'historique consiste à modifier les paramètres d'un modèle de réservoir, tels que les perméabilités, les porosités ou les skins de puits (représentant les endommagements autour du puits), les connections de failles..., pour minimiser les écarts entre les données d'historique mesurées et les réponses correspondantes simulées à partir du modèle de réservoir. Les paramètres peuvent être liés à des régions géographiques comme les perméabilités ou porosités autour d'un puits ou plusieurs puits. L'écart entre les données d'historique et réponses simulées forme une fonctionnelle, dite fonction objectif. Le problème du calage d'historique se résout en minimisant cette fonctionnelle.

Plusieurs types de méthodes existent pour minimiser cette fonctionnelle. Certaines ne fournissent qu'un seul modèle correspondant au minimum de la fonction objectif. D'autres cherchent à estimer la loi de distribution des paramètres incertains conditionnée aux données dynamiques, dites loi a posteriori, et ainsi estimer l'incertitude sur les paramètres. En effet, l'incertitude sur les données mesurées et le simulateur d'écoulement se propage sur l'estimation du "meilleur" modèle de réservoir. Il est alors important de connaître la loi a posteriori autour du meilleur modèle. Dans le cas général, cette loi ne peut être connue qu'en générant un échantillon de valeurs suivant cette loi. On peut utiliser pour cela des méthodes de type Markov-Chain Monte Carlo (MCMC), qui convergent vers la loi voulue. Une telle méthode est décrite dans le document :

- Geyer C J (1992) Practical Markov chain Monte Carlo (with discussion). *Statistical Science* 7:473-511. doi:10.1214/ss/1177011137.

Les méthodes MCMC nécessitent un grand nombre de simulations d'écoulement, ce qui demande un temps de calcul important. Pour limiter l'utilisation du simulateur d'écoulement et donc pour réduire le temps de calcul, d'autres solutions ont été imaginées. Par exemple, il est possible d'utiliser un modèle de la fonction objectif dans l'espace des paramètres, appelé "surface de réponse". Ce modèle est construit progressivement au cours des itérations par interpolation des points simulés. Par exemple, le document suivant décrit une telle méthode :

- Busby D, Feraille M (2008) Adaptive Design of Experiments for Bayesian Inversion – An Application to Uncertainty Quantification of a Mature Oil Field. *Journal of Physics: Conference Series* 135. doi: 10.1088/1742-6596/135/1/012026.

Cette approche présente donc deux niveaux d'approximation : le premier concerne la convergence de l'algorithme MCMC dont il est très difficile de dire s'il a convergé vers la loi voulue ; le second concerne l'approximation de la fonction objectif par la surface de réponse.

Ainsi, l'objet de l'invention concerne un procédé d'exploitation d'un réservoir géologique selon un schéma d'exploitation défini à partir d'un modèle de réservoir. Le modèle de réservoir est obtenu par une technique de "calage d'historique" probabiliste prenant en compte une loi analytique de distribution conditionnelle de paramètres incertains associée à une surface de réponse donnée. Dans le procédé selon l'invention, cette surface de réponse n'approche pas la fonction objectif elle-même mais une autre fonction déterminée à partir de la fonction objectif. Ceci permet d'obtenir de meilleures approximations de la fonction objectif et donc de la loi a posteriori. De plus, l'utilisation d'une loi analytique permet de réduire les temps de calcul, d'envisager un traitement beaucoup plus rapide des résultats et d'obtenir un modèle de réservoir représentatif en limitant les approximations.

15

#### **Le procédé selon l'invention**

L'invention concerne un procédé d'exploitation d'un réservoir géologique selon un schéma d'exploitation défini à partir d'un modèle de réservoir, ledit modèle de réservoir comportant un maillage associé à des paramètres  $\theta$  dudit réservoir. Pour ce procédé, on réalise les étapes suivantes :

20

a) on construit un modèle de réservoir calé sur des données mesurées au sein dudit réservoir au moyen des étapes suivantes :

25

i) on génère de manière stochastique un ensemble initial de modèles de réservoir à partir de lois de probabilité  $p(\theta)$  desdits paramètres  $\theta$  ;

ii) on détermine une fonction objectif  $F(\theta)$  mesurant un écart entre des données dynamiques  $y_1, \dots, y_n$  acquises en cours d'exploitation et des données dynamiques simulées au moyen d'un simulateur d'écoulement et desdits modèles de réservoir appartenant audit ensemble ;

30

iii) on détermine à partir de ladite fonction objectif  $F(\theta)$  une loi analytique  $p(\theta|y_1, \dots, y_n)$  de probabilité conditionnelle des paramètres  $\theta$  connaissant lesdites données dynamiques mesurées  $y_1, \dots, y_n$  ;

iv) on génère au moins un nouveau modèle de réservoir au moyen de ladite loi analytique  $p(\theta|y_1, \dots, y_n)$  et on ajoute ledit nouveau modèle audit ensemble de modèles de réservoir ;

- v) on réitère les étapes ii) à iv) et on détermine le modèle de réservoir qui minimise ladite fonction objectif ;
- b) on détermine un schéma d'exploitation optimal du réservoir en simulant l'exploitation du réservoir au moyen dudit modèle de réservoir calé et dudit simulateur d'écoulement ; et
- 5 c) on exploite ledit réservoir en mettant en œuvre ledit schéma d'exploitation optimal.

Selon l'invention, on détermine ladite loi analytique conditionnelle  $p(\theta|y_1, \dots, y_n)$  en déterminant une approximation d'une fonction  $G(\theta)$  telle que  $G(\theta) = \exp(-F(\theta))$ , puis en calculant ladite loi analytique à partir de la relation suivante :

$$10 \quad p(\theta|y_1, \dots, y_n) = \frac{G(\theta) \times p(\theta)}{\int G(\theta) \times p(\theta) d\theta}.$$

De préférence, on détermine l'approximation de ladite fonction  $G(\theta)$  en calculant  $G(\theta_i) = \exp(-F(\theta_i))$  pour les modèles  $i$  appartenant audit ensemble, puis par interpolation desdites valeurs de  $G(\theta_i)$ .

15 Selon un mode de réalisation de l'invention, on interpole lesdites valeurs de  $G(\theta_i)$  par une technique de krigeage.

Selon l'invention, à l'étape iv) à chaque itération on génère plusieurs nouveaux modèles de réservoir, parmi lesquels on choisit  $M_1$  modèles qu'on ajoute audit ensemble, ledit ensemble de modèles de réservoir est constitué lors de la construction de l'ensemble initial d'un nombre  $M_0$  de modèles.

20 De manière avantageuse, lesdits  $M_0$  modèles dudit ensemble initial sont échantillonnés par une technique de plan d'expériences de type hypercube latin.

De manière avantageuse, lesdits  $M_1$  modèles ajoutés dans ledit ensemble sont choisis à l'aide de l'approximation de ladite fonction  $G(\theta)$ .

25 Selon une variante de réalisation de l'invention, au moins une partie desdits  $M_1$  modèles sont les modèles pour lesquels l'approximation de la fonction  $G(\theta)$  donne les valeurs maximales.

En outre, au moins une partie desdits  $M_1$  modèles peuvent être les modèles pour lesquelles pour lesquels la qualité de l'approximation de la fonction  $G(\theta)$  possède des variances de krigeage les plus élevées parmi lesdits nouveaux modèles.

30 On peut réaliser une étude de sensibilité par le calcul de plusieurs fonctions objectif, chaque fonction objectif étant calculée pour une partie des paramètres ou pour un intervalle de temps limité par rapport aux informations disponibles.

L'invention concerne également un produit programme d'ordinateur téléchargeable depuis un réseau de communication et/ou enregistré sur un support lisible par ordinateur et/ou exécutable par un processeur. Il comprend des instructions de code de programme pour la mise en œuvre du procédé tel que décrit ci-dessus, lorsque ledit programme est exécuté sur un ordinateur.

### **Présentation succincte des figures**

D'autres caractéristiques et avantages du procédé selon l'invention, apparaîtront à la lecture de la description ci-après d'exemples non limitatifs de réalisations, en se référant aux figures annexées et décrites ci-après.

La figure 1 illustre les différentes étapes du procédé selon l'invention.

La figure 2 illustre des valeurs de la fonction objectif pour les différents modèles de réservoir considérés au cours du processus de calage du procédé selon l'invention.

Les figures 3 illustrent la distribution de différents paramètres.

Les figures 4 illustrent la distribution de différents paramètres pour différentes périodes de calage.

Les figures 5 illustrent la distribution de différents paramètres lorsque certains paramètres sont fixés.

Les figures 6 illustrent des données de production au puits.

Les figures 7 à 12 illustrent la distribution de différents paramètres pour l'exemple 2.

### **Description détaillée de l'invention**

Le procédé de l'invention est destiné à l'exploitation d'un réservoir souterrain, cette exploitation peut consister à la récupération d'hydrocarbures mais aussi au stockage de gaz, tel que du CO<sub>2</sub> dans le réservoir. Dans la suite de la description, nous allons décrire essentiellement l'utilisation du procédé dans le cas de la récupération d'hydrocarbures.

La figure 1 illustre le procédé d'exploitation d'un réservoir pétrolier selon l'invention. Le procédé comporte les étapes principales suivantes :

- 1) détermination de paramètres incertains (X1)
- 2) construction d'un ensemble initial de modèles de réservoir (EMRi)
- 3) calage d'historique de type probabiliste (CAL)
- 4) simulation de schémas d'exploitation (SE)

## 5) exploitation du gisement (EX)

Étape 1) détermination de paramètres incertains (X1)

Lors de cette étape, on détermine un ensemble de P paramètres incertains  $\theta = (\theta_1 \dots \theta_p)$  utilisés dans le modèle de réservoir ainsi que des lois de probabilité  $p(\theta_i)$  (dites lois "a priori") pour chacun de ces paramètres. Ces lois n'incorporent aucune information venant des données dynamiques. Elles sont basées sur des informations venant de la connaissance du réservoir, de données statiques (géologiques par exemple), etc. Par exemple, les lois de probabilité  $p(\theta)$  choisies peuvent être du type uniforme, Gaussien, etc.

Les paramètres incertains peuvent être liés à n'importe quelle partie du modèle de réservoir. Par exemple, il peut s'agir de paramètres permettant de remplir le modèle de réservoir en propriétés pétrophysiques (perméabilité, porosité), de paramètres caractérisant les puits (skins), les fluides en place ou injectés (saturation en huile résiduelle après balayage à l'eau ou au gaz par exemple) ou encore la structure du réservoir (connections de failles). Un jeu de valeurs des paramètres  $\theta = (\theta_1 \dots \theta_p)$ , qui correspond également à un point de l'espace des paramètres, définit donc un modèle de réservoir. Une modification de chaque paramètre entraîne donc une modification du modèle de réservoir. En particulier, si le paramètre incertain est lié à la répartition spatiale d'une propriété pétrophysique, il peut être nécessaire d'utiliser une technique de simulation géostatistique pour générer le modèle de réservoir associé.

Étape 2) construction d'un ensemble initial de modèles de réservoir (EMRi)

Les formations géologiques sont en général des milieux très hétérogènes. La modélisation d'un réservoir, c'est-à-dire la construction d'un modèle de réservoir respectant les données statiques, nécessite de recourir à des procédés de construction dits « probabilistes » du fait de la limitation de l'information disponible (nombre de puits restreints, ...). De ce fait, les modèles géologiques construits à partir de ces procédés probabilistes sont appelés « modèles stochastiques ». La construction d'un modèle stochastique de réservoir doit d'abord dépendre de l'environnement du dépôt géologique, ce qui permet de représenter les hétérogénéités majeures qui contrôlent l'écoulement des fluides. L'intégration des données statiques dans ce modèle passe par des opérations linéaires et peut se faire à partir de techniques géostatistiques bien connues des spécialistes. Pour cela, on peut utiliser avantageusement un simulateur géostatistique.

Un modèle de réservoir, pouvant être représenté sur un ordinateur, comprend une grille à N dimensions ( $N > 0$  et en général égal à deux ou trois) dont chacune des mailles se voit affecter la valeur d'une propriété caractéristique de la zone étudiée. Il peut s'agir par exemple de la porosité ou de la perméabilité distribuée dans un réservoir. Ces valeurs constituent des cartes. Ainsi un modèle comprend une grille associée à au moins une carte. D'autres paramètres caractérisant par exemple les puits ou les fluides injectés interviennent également dans la définition de ce modèle.

Lors de cette étape on génère de manière stochastique un ensemble initial de modèles de réservoir au moyen de la loi de probabilité a priori  $p(\theta)$  de chaque paramètre incertain.

Plus précisément, on génère pour chaque paramètre incertain  $\theta_i$  sélectionné lors de l'étape 1), un échantillon ( $X_1$ ) de  $M_0$  valeurs selon la loi de distribution a priori  $p(\theta_i)$ , par exemple à partir d'une méthode de plan d'expérience de type hypercube latin. On obtient ainsi un ensemble de  $M_0$  jeux de valeurs des paramètres incertains, et donc  $M_0$  modèles de réservoir différents.

15

### Étape 3) calage du modèle de réservoir (CAL)

Le but du processus de calage selon l'invention est de construire un modèle de réservoir le plus représentatif possible. Or, à ce stade, les données dynamiques n'ont pas été considérées pour construire le modèle de réservoir initial. On acquiert donc des données dynamiques en cours d'exploitation du gisement (DD). Il s'agit de données de production, d'essais de puits, de temps de percée, de sismique 4D... dont la particularité est de varier au cours du temps en fonction des écoulements de fluide dans le réservoir. Ces mesures sont réalisées au moyen d'outils de mesure tels que des débitmètres ou des capteurs sismiques. On note  $(y_1 \dots y_n)$  le vecteur des données dynamiques observées et  $(t_1 \dots t_n)$  les temps d'acquisition correspondants.

25

Ces données dynamiques sont ensuite intégrées dans le modèle de réservoir par le biais du calage probabiliste. Selon l'invention, on réalise un calage d'historique de type probabiliste (CAL) en réalisant les étapes suivantes :

30

- i. on utilise un simulateur d'écoulement  $f$  pour calculer la réponse en production simulée avec les modèles de réservoir que l'on souhaite évaluer (SIM) ;
- ii. on calcule une fonction objectif (OF) mesurant l'écart entre les résultats simulés (SIM) et les données dynamiques (DD) associées ;

- iii. on construit par krigeage une approximation de l'exponentielle de l'opposé de la fonction objectif, et on détermine pour cette surface de réponse une loi de probabilité conditionnelle analytique  $\tilde{p}(\theta|y_1, \dots, y_n)$  des paramètres incertains connaissant les données dynamiques, cette loi est également appelée loi a posteriori (POST) ;
- iv. on génère au moins un nouveau modèle de réservoir que l'on souhaite évaluer pour améliorer la qualité de la surface de réponse (EMR) ;
- v. on réitère les étapes i) à iv), pour connaître le plus précisément possible la distribution a posteriori de la fonction objectif.

5

10

#### i. simulation des données dynamiques (SIM)

Au moyen d'un simulateur d'écoulement, par exemple le logiciel PumaFlow ® (IFP Energies nouvelles, France), on simule les écoulements dans les modèles de réservoir à évaluer, c'est à dire soit les modèles de réservoir de l'ensemble initial, soit les modèles de réservoir générés lors de l'itération précédente. On peut simuler par exemple la récupération de pétrole ou les déplacements des fluides (par exemple les gaz stockés) dans le réservoir.

15

#### ii. calcul d'une fonction objectif (OF)

La fonction objectif  $F(\theta)$  représente les écarts entre les données dynamiques mesurées (DD) et les données dynamiques simulées (SIM). Selon un mode de réalisation de l'invention, on calcule la fonction objectif  $F(\theta)$  au sens des moindres carrés :

20

$$F(\theta) = \frac{1}{2} \sum_{i=1}^n \omega_i (y_i - f(t_i, \theta))^2 \text{ avec } \omega_i \text{ le poids associé à la donnée } i.$$

Selon l'invention, on calcule la fonction objectif pour chaque modèle de réservoir de l'ensemble. On rappelle que cet ensemble comporte initialement  $M = M_0$  modèles de réservoir. Cet ensemble augmente à chaque itération d'un nombre  $M_1$  de modèles. La sélection des  $M_1$  modèles supplémentaires, ou de façon équivalente des  $M_1$  jeux de valeurs des paramètres incertains, est décrite dans le paragraphe iv) de sélection de nouveaux modèles de réservoir.

25

30

iii. détermination de la loi conditionnelle analytique (POST)

Le problème inverse en ingénierie de réservoir consiste à chercher la valeur de paramètres incertains du modèle réservoir qui, appliquées au simulateur d'écoulement, fournissent une réponse simulée la plus proche possible des données dynamiques mesurées. La réponse simulée aux puits se modélise de la façon suivante :

$$y_i = f(t_i, \theta) + \varepsilon_i \quad i = 1, \dots, n$$

où  $t_i$  est le  $i$ -ième temps d'acquisition de données dynamiques,  $y_i$  le vecteur des données dynamiques observées,  $\theta$  le vecteur des paramètres incertains du modèle,  $f$  le simulateur d'écoulement qui modélise la relation entre les paramètres et les données observées,  $\varepsilon_i$  l'erreur entre le modèle et la donnée dynamique et  $n$  le nombre de données dynamiques mesurées. L'objectif est d'utiliser les données observées pour déduire des informations sur le vecteur des paramètres incertains  $\theta$ .

On suppose ici que les erreurs  $\varepsilon_i$  entre le modèle de simulation d'écoulement  $f(t_i, \theta)$  et les données observées  $y_i$  suivent une loi donnée, estimée en fonction des erreurs de mesure. La loi de distribution conditionnelle du vecteur  $\theta$  connaissant les données mesurées, ou loi a posteriori, est donnée par la formule de Bayes :

$$p(\theta | y_1, \dots, y_n) = \frac{p(y_1, \dots, y_n | \theta) \times p(\theta)}{p(y_1, \dots, y_n)}$$

où  $p(\theta)$  est la loi de distribution a priori,  $p(y_1, \dots, y_n)$  est la loi de distribution des données,  $p(y_1, \dots, y_n | \theta)$  est la loi de distribution conditionnelle des données dynamiques sachant  $\theta$  et  $p(\theta | y_1, \dots, y_n)$  est la loi de distribution conditionnelle a posteriori recherchée. Enfin, en supposant les erreurs  $\varepsilon_i$  Gaussiennes et indépendantes, la loi a posteriori peut se réécrire :

$$p(\theta | y_1, \dots, y_n) = \frac{\exp(-F(\theta)) \times p(\theta)}{\int \exp(-F(\theta)) \times p(\theta) d\theta}$$

Et dans ce cas, le poids  $\omega_i$  utilisé pour le calcul de la fonction objectif est défini par  $\omega_i = 1/\sigma_i^2$  avec  $\sigma_i^2$  la variance de l'erreur  $\varepsilon_i$ .

Selon l'invention, on détermine une approximation de la loi  $p(\theta|y_1, \dots, y_n)$  de distribution conditionnelle des paramètres incertains connaissant les données dynamiques mesurées par une loi analytique  $\tilde{p}(\theta|y_1, \dots, y_n)$  dépendant de la fonction objectif  $F(\theta)$ .

On détermine ladite loi analytique conditionnelle  $\tilde{p}(\theta|y_1, \dots, y_n)$  en considérant une  
 5 fonction  $G(\theta)$  dépendante de ladite fonction objectif  $F(\theta)$ , définie par  $G(\theta) = \exp(-F(\theta))$ .

La loi a posteriori se réécrit alors :  $p(\theta|y_1, \dots, y_n) = \frac{G(\theta) \times p(\theta)}{\int G(\theta) \times p(\theta) d\theta}$ . Cette relation découle

de la relation que nous avons vu précédemment  $p(\theta|y_1, \dots, y_n) = \frac{\exp(-F(\theta)) \times p(\theta)}{\int \exp(-F(\theta)) \times p(\theta) d\theta}$ .

La fonction  $G(\theta)$  n'étant pas connue de manière analytique, on en construit une approximation, appelée également surface de réponse, par interpolation des valeurs  
 10 connues  $G(\theta_j) = \exp(-F(\theta_j))$  calculées pour chacun des modèles  $j$  de réservoir au cours de l'étape ii). Contrairement à l'art antérieur, la surface de réponse utilisée par l'invention approche non pas la fonction objectif, mais une fonction dépendant de cette fonction objectif. La fonction  $G(\theta)$  est comprise entre 0 et 1, et varie donc dans un intervalle plus restreint que la fonction  $F(\theta)$ . Par conséquent, il est plus facile de construire une bonne approximation de  
 15 la fonction  $G(\theta)$  que de la fonction  $F(\theta)$ .

De préférence, la méthode d'interpolation utilisée est une technique de krigeage ou toute autre méthode de régression appartenant à la famille de la minimisation du risque empirique pénalisé dans des espaces de Hilbert à noyau reproduisant (RKHS de l'anglais Reproducing Kernel Hilbert Space). Ces méthodes d'interpolation permettent d'écrire  
 20 l'approximation de la fonction (ou surface de réponse)  $\tilde{G}(\theta)$  sous la forme d'une combinaison linéaire de fonctions connues. Le calcul de la loi  $\tilde{p}(\theta|y_1, \dots, y_n)$  de probabilité conditionnelle associée à  $\tilde{G}(\theta)$  fait alors intervenir différentes intégrales de ces fonctions connues, qui sont calculables analytiquement.

#### 25 iv. sélection de nouveaux modèles de réservoir (EMR)

On génère de nouveaux modèles de réservoir, c'est à dire de nouveaux jeux de valeurs des paramètres incertains, en fonction de la loi analytique  $\tilde{p}(\theta|y_1, \dots, y_n)$  de probabilité conditionnelle obtenue à l'étape précédente. On sélectionne  $M_1$  modèles parmi

cet échantillon pour lesquels on va calculer la fonction objectif afin d'améliorer notre connaissance de cette fonction, et donc la qualité de la surface de réponse.

Selon l'invention, l'ensemble des modèles de réservoir pour lesquels on calcule la réponse en production est augmenté à chaque itération du processus de calage par l'ajout  
5 de  $M_1$  nouveaux modèles. Ces modèles sont choisis selon plusieurs critères (de préférence avec un nombre de modèles identique pour chacun des critères), notamment :

- (i) on sélectionne les points de l'espace des paramètres, c'est à dire les jeux de valeur des paramètres incertains, pour lesquels la qualité de l'approximation de la fonction  $G(\theta)$  est faible.
- 10 (ii) on sélectionne les points de l'espace des paramètres pour lesquels la surface de réponse  $\tilde{G}(\theta)$  prédit de fortes valeurs, c'est-à-dire de faibles valeurs de la fonction objectif. On peut par exemple utiliser la technique de l'Expected Improvement (amélioration attendue) décrite par exemple dans le document : Schonlau M (1997) Computer Experiments and Global  
15 Optimization, Dissertation, University of Waterloo, Canada.

D'autres critères peuvent être utilisés notamment pour choisir des points où la variance de krigeage est élevée, les points qui si on les ajoute font diminuer le plus l'intégrale de l'erreur de modèle, ou les points d'entropie maximale, ....

On génère ensuite les modèles de réservoir correspondant à chaque nouveau jeu de  
20 valeurs des paramètres sélectionnés précédemment. Ceci peut nécessiter d'utiliser un simulateur géostatistique pour générer les propriétés pétrophysiques du modèle de réservoir.

#### v. répétition des étapes

Tant que le nombre maximum de points que l'on peut évaluer (simuler) n'a pas été  
25 atteint, on réitère les étapes du calage avec un nombre de modèles de réservoir appartenant à l'ensemble de plus en plus important.

A chaque construction par krigeage de  $\tilde{G}(\theta)$ , on calcule le critère de qualité  $Q_2$  de cette surface de réponse à partir des modèles de réservoir de l'ensemble utilisé pour la

définir :  $Q2 = 1 - \frac{\sum_{i=1}^M (G(\theta_i) - \tilde{G}_{-i}(\theta_i))^2}{\sum_{i=1}^M (G(\theta_i) - \bar{G})^2}$ . Dans cette équation, on note  $\bar{G}$  la moyenne des

valeurs  $G(\theta_i), i=1 \dots M$  et  $\tilde{G}_{-i}(\theta_i)$  la prédiction pour les valeurs  $\theta_i$  de la surface de réponse obtenue par krigeage en utilisant les  $M-1$  jeux de valeurs  $\theta_j, j=1 \dots M, j \neq i$ . Selon un mode de réalisation de l'invention, il est possible de définir un critère Q2 en enlevant des groupes de mailles plutôt qu'une seule maille. D'autres types de critères de qualité peuvent être également utilisés. Le critère de qualité Q2 est meilleur si les fortes valeurs de  $G$  sont bien prédites. Dans notre cas, ces fortes valeurs correspondent à des valeurs de paramètres pour lesquelles la fonction objectif est faible, c'est-à-dire à des valeurs où l'adéquation avec les données d'historique est bonne. Ainsi, un critère Q2 élevé signifie que l'on approche correctement la fonction objectif pour des paramètres correspondant à un bon calage d'historique.

Dans le cas déterministe, un modèle de réservoir calé avec les données dynamiques revient à associer aux paramètres incertains (ou de calage du modèle) la valeur correspondant à la plus faible valeur de la fonction objectif. Dans notre cas, un modèle de réservoir calé avec les données dynamiques et selon la méthode probabiliste décrite précédemment revient à associer aux paramètres incertains (ou de calage du modèle) leur loi de probabilité a posteriori.

Selon un mode de réalisation de l'invention, on utilise une fonction objectif normalisée  $F_s(\theta) = F(\theta)/s$  où  $s$  est une constante de normalisation. Ceci implique de construire une approximation de  $G_s(\theta) = \exp\left(-F(\theta)/s\right)$ . Les étapes précédentes ii à iv sont conservées pour choisir de nouveaux modèles à simuler à l'aide de l'approximation de  $G_s(\theta)$ . Ensuite, on échantillonne la loi de probabilité a posteriori finale de la fonction objectif par une méthode MCMC, après évaluation d'un modèle de  $F(\theta)$  par krigeage.

25

#### Étape 4) Simulation de schémas d'exploitation (SE)

A partir d'au moins un modèle de réservoir calé de manière probabiliste, les spécialistes peuvent déterminer plusieurs schémas d'exploitation (SE) correspondant à différentes options possibles d'exploitation du réservoir souterrain : emplacement des puits

producteur et/ou injecteur, valeurs cibles pour les débits par puits et/ou pour le réservoir, le type d'outils utilisés, les fluides utilisés et récupérés... Pour chacun de ces schémas, il convient de déterminer leurs prévisions de production après la période de calage. Ces prévisions de production probabilistes sont obtenues au moyen d'un logiciel de simulateur d'écoulement (de préférence le même que celui utilisé auparavant) ainsi qu'au moyen du modèle numérique de réservoir calé.

On définit un ou plusieurs schémas d'exploitation (SE) possibles adaptés au modèle de réservoir. Pour chacun de ces schémas, on propage l'incertitude a posteriori obtenue après le calage probabiliste.

La propagation de la loi a posteriori des paramètres incertains peut être réalisée en construisant des surfaces de réponses sur les résultats d'intérêts de la simulation (débit cumulé d'huile ou valeur économique du réservoir, par exemple), ceci afin d'éviter le lancement d'un trop grand nombre de simulations numériques. La description de cette étape est donnée par exemple dans le document :

- Feraille M, Marrel A (2012) Prediction under uncertainty on a mature field. Oil & Gas Science and Technology. doi: 10.2516/ogst/2011172.

A partir des prévisions de productions probabilistes définies pour chaque schéma d'exploitation (étape précédente), les spécialistes peuvent par comparaison choisir le schéma d'exploitation qui leur semble le plus pertinent. Par exemple :

- en comparant le maximum du volume d'huile récupéré, on peut déterminer le schéma de production susceptible de fournir le maximum de récupération ou d'être le plus rentable.
- en comparant l'écart type du volume d'huile récupéré, on peut déterminer le schéma de production le moins risqué.

#### Étape 5) Exploitation du réservoir souterrain (EX)

On exploite alors le réservoir selon le schéma d'exploitation choisi (EX) par exemple en forant de nouveaux puits (producteur ou injecteur), en modifiant les débits et/ou les fluides injectés ...

#### Exploitation des résultats

Le calcul analytique, donc quasi-instantané, des lois de probabilité a posteriori des paramètres et de la fonction objectif pour une surface de réponse donnée permet

d'envisager de nouvelles voies d'exploitation des résultats du calage probabiliste permettant de mieux comprendre l'impact de chaque paramètre incertain sur le comportement du réservoir.

Par exemple, on peut décider de fixer la valeur d'un ou plusieurs des paramètres et de regarder l'impact sur l'incertitude des autres paramètres ainsi que sur la ou les fonctions objectif choisies. Dans ce cas, il suffit de recalculer analytiquement la loi de distribution conditionnelle des paramètres  $\theta_j, j = 1 \dots N, j \neq i$  conditionnée à la valeur choisie pour  $\theta_i$ .

On peut également envisager de considérer plusieurs fonctions objectif selon l'information recherchée. Ceci se fait en estimant la distribution a posteriori des paramètres pour chacune de ces fonctions via la construction d'une surface de réponse. La fonction objectif  $F$  mesure l'écart global entre les résultats de simulations et les données dynamiques mesurées. Cette fonction est classiquement construite au sens des moindres carrés. D'un point de vue métier, une fonction objectif calcule la somme des écarts entre des propriétés d'historique liées à des objets métiers (puits P1 à P10, groupe de puits, piège, ...), eux-mêmes liés à des propriétés comme le débit d'huile QOS, d'eau QWS, de gaz QGS, ..., qui sont elles-mêmes fonction du temps où ont été effectuées les mesures (par exemple tous les mois de 1980 à 2000). A partir des données d'historique, il est donc possible de calculer une fonction objectif pour n'importe quel sous-ensemble constitué d'une partie des objets, des propriétés ou d'un certain nombre d'intervalles de temps. Par exemple, on peut déterminer une fonction objectif liée :

- au puits P1 pour toutes les propriétés et tous les temps, ou
- à la période 1980-1990 pour tous les objets et toutes les propriétés.

Il peut donc être intéressant de connaître l'influence des paramètres sur différentes fonctions objectif, par exemple :

- pour identifier à partir de quel moment les différents paramètres deviennent influents, il suffit de considérer, pour plusieurs temps, la fonction objectif incluant toutes les données jusqu'à chacun de ces temps, puis de construire une surface de réponse pour chacune des fonctions obtenues. On a ainsi accès à l'évolution de la distribution a posteriori en fonction du temps de calage de manière interactive.
- pour connaître l'influence des paramètres sur chaque puits, il suffit d'estimer les distributions a posteriori des fonctions objectif associées aux propriétés d'un puits donné.

Connaître l'influence des paramètres incertains sur des fonctions objectif "métiers" définies par le spécialiste lui permet de mieux appréhender le fonctionnement du réservoir et de prendre des décisions. Par exemple, il est possible

- 5 - de fixer fictivement la valeur des paramètres vus comme influents, le spécialiste peut alors s'assurer que la réduction de l'incertitude sur le modèle de réservoir qui en découle (sur les autres paramètres et sur une ou plusieurs fonctions objectif définies par l'ingénieur) est suffisante pour justifier une meilleure caractérisation de ces paramètres influents. Dans ce cas, de nouvelles mesures sur le champ pétrolier peuvent être réalisées (tests de puits, diagrapies, ...).
- 10 - de savoir que des paramètres liés à une partie du réservoir commencent à avoir une influence à partir d'une date sur une ou des fonctions objectif liées à une autre partie du réservoir, ce qui permet de mieux comprendre le fonctionnement dynamique du réservoir. Cette compréhension peut se traduire par des schémas d'exploitation différents et adaptés pour chacune des parties du réservoir.
- 15 Enfin, on peut envisager des études de sensibilité aux poids de la fonction objectif, ce qui est extrêmement coûteux si l'on doit utiliser la méthode MCMC de l'art antérieur.

### Exemples d'application

20 Afin d'illustrer le procédé selon l'invention, le demandeur a réalisé deux séries d'expérimentations. La première série concerne l'obtention des distributions a posteriori. La deuxième série illustre la bonne corrélation entre la loi analytique associée aux surfaces de réponse et la loi de distribution conditionnelle exacte.

#### Exemple 1

25 Pour illustrer la méthode, on décrit un exemple d'application synthétique. Le réservoir considéré contient initialement du gaz, de l'huile et de l'eau. Il est fermé par des failles sur deux côtés, et par un aquifère sur les autres côtés. La production se fait par déplétion à partir de six puits (PRO1, PRO4, PRO5, PRO11, PRO12, PRO15) avec un débit de production d'huile de 150 m<sup>3</sup>/jour imposé à chaque puits producteur pendant 10 ans. Au cours de cette

30 période de production, le cumulé d'huile (V) et le ratio gaz/huile (R) sont collectés pour chaque puits producteur, ainsi que le cumulé d'huile total dans le réservoir. Ces valeurs sont connues avec un bruit relatif de 4%. Ces données sont illustrées par les courbes de la figure 6. La figure 6a) correspond au volume (V) du cumulé d'huile pour chaque puits producteur en

fonction du temps  $t$  exprimé en années. La figure 6b) correspond au ratio gaz/huile ( $R$ ) pour chaque puits producteur en fonction du temps  $t$  exprimé en années.

On suppose également connue dans le réservoir la distribution des propriétés pétrophysiques (porosité, perméabilité). Les paramètres incertains sont les sept paramètres  
5 suivants : quatre multiplicateurs de transmissivité ( $MPH2$ ,  $MPH1$ ,  $MPV2$ ,  $MPV1$ ), la saturation en huile résiduelle après balayage à l'eau ( $SORW$ ), la saturation en huile résiduelle après balayage au gaz ( $SORG$ ) et un paramètre caractérisant l'aquifère ( $AQUI$ ).

Il s'agit alors d'estimer, en utilisant le procédé selon l'invention, la distribution des paramètres incertains et la distribution de la fonction objectif conditionnées aux données  
10 dynamiques.

#### *Estimation des distributions a posteriori*

Le calage probabiliste est réalisé avec un plan d'expérience initial de type hypercube latin LHS définissant un ensemble constitué de 70 modèles de réservoir, puis avec un ajout de modèles par séries de 10 à chaque itération du processus de calage. Les 5 premiers  
15 modèles ajoutés sont choisis là où la surface de réponse  $\tilde{G}(\theta)$  prédit de fortes valeurs, et les 5 suivants où la qualité de l'approximation de  $G(\theta)$  est faible au sens de la variance de krigeage élevée. L'évolution de la fonction objectif au cours de la procédure d'ajout de modèles dans le processus de calage du modèle numérique du réservoir est donnée sur la Figure 2. Les barres verticales plus foncées marquent la dernière simulation du LHS et de  
20 chaque série d'ajout de modèles. On peut donc voir que, après les 70 premiers modèles d'exploration (LHS), les points ajoutés par séries de 10 présentent bien 5 premières valeurs globalement faibles de la fonction objectif, puis 5 valeurs globalement plus élevées.

Les figures 3 représentent les courbes de la distribution a priori, de la distribution a posteriori déterminée après le plan d'expérience initial et de la distribution a posteriori à la fin  
25 du processus de calage pour chaque paramètre incertain ( $MPH2$ ,  $MPH1$ ,  $MPV2$ ,  $MPV1$ ,  $SORW$ ,  $SORG$ ,  $AQUI$ ) et pour la fonction objectif ( $FO$ ). Sur ces figures, le point noir indique la valeur de référence, la courbe discontinue en gris correspond à la loi de probabilité a priori, la courbe continue en noir correspond à la loi de probabilité a posteriori avec 70 modèles et la courbe discontinue en noir correspond à la loi de probabilité a posteriori à la fin  
30 du processus de calage avec 210 modèles.

#### *Exploitation des résultats*

Grâce à la mise à jour quasi-immédiate (car analytique) de la distribution conditionnelle a posteriori pour un modèle de krigeage donné, on peut envisager de fixer la valeur d'un ou plusieurs des paramètres incertains et de regarder l'impact sur l'incertitude des autres

paramètres. Ceci est illustré par exemple sur les figures 5 où sont représentées les distributions conditionnelles a posteriori à la fin du processus de calage des paramètres MPH2, MPH1, MPV2, MPV1 et AQU1 en fixant ou non la valeur des paramètres SORG et SORW à leur valeur la plus probable. Sur ces figures, le point noir correspond à la valeur de référence, la courbe discontinue correspond à la loi de probabilité a posteriori lorsqu'aucun paramètre n'est fixé, alors que la courbe continue correspond à la loi de probabilité lorsqu'on fixe SORG et SORW. On voit ici que cela engendre une modification de certaines distributions.

On peut également envisager des études de sensibilité aux poids de la fonction objectif. Par exemple, on peut s'intéresser à l'évolution des lois conditionnelles a posteriori des paramètres lorsque l'on considère un intervalle de calage plus ou moins long dans le calcul de la fonction objectif. Dans le cas présenté ici, on a considéré toutes les simulations effectuées lors du processus probabiliste et reconstruit le modèle de krigeage associé respectivement à un historique de 10, 8, 6, 4 et 2 ans. Les nouvelles lois conditionnelles a posteriori des paramètres sont données sur les figures 4. Sur ces figures, le point noir désigne la valeur de référence, la courbe discontinue en gris qui reste horizontale (uniforme) correspond à la loi de probabilité a priori, la courbe discontinue constituée d'une série de petits points correspond à la loi de probabilité a posteriori calculée pour deux ans, la courbe discontinue constituée d'une série de tirets correspond à la loi de probabilité a posteriori calculée pour quatre ans, la courbe continue correspond à la loi de probabilité a posteriori calculée pour six ans, la courbe discontinue constituée d'une série de losanges correspond à la loi de probabilité a posteriori calculée pour huit ans, et la courbe discontinue constituée d'une série de gros points correspond à la loi de probabilité a posteriori calculée pour dix ans. On remarque que les principaux changements interviennent après 6 ans.

25

### Exemple 2

Cet exemple permet de détailler la méthodologie proposée sur un exemple analytique, qui nous permet de valider et d'étudier de manière approfondie les formes analytiques possibles de la loi a posteriori pour différents cas. Tous les exemples développés ici ont été traités avec le logiciel Matlab ® (MathWorks, États-Unis) et la boîte à outils "Dacefit" pour la construction du krigeage.

Pour cet exemple on remplace le simulateur d'écoulement  $f$  par un modèle linéaire  $f(t_i, \theta) = t_i^T \theta$  où  $\theta = (\theta_1 \dots \theta_D)$  est un vecteur de dimension D.

Pour tester la méthode de calage du procédé selon l'invention sur ce modèle linéaire, on répète la procédure suivante pour  $D = 1$ ,  $D = 2$  et  $D = 3$  :

1. On fixe la valeur de référence de  $\theta$

2. On choisit  $n = 10$  points dans l'intervalle  $[0,1]^D$  selon un hypercube latin, on note ces points  $t_i$ ,  $i = 1, \dots, n$  et on génère les observations  $y_i = t_i^T \theta + \varepsilon_i$  en générant aléatoirement des variables Gaussiennes  $\varepsilon_i$  en se fixant une valeur  $\sigma_\varepsilon^2$ . D'une part, cette étape permet de générer les "fausses" données observées, ce qui permet de construire une fonction objectif. D'autre part elle permet de déduire (dans ce cas précis uniquement) les lois a posteriori exactes, pour les comparer aux lois approchées.

3. On génère un ensemble  $(\theta_i)_{i=1, \dots, N}$  de  $N$  paramètres avec  $N = 10D$ ,  $20D$ , ou  $50D$  selon un hypercube latin adapté à la loi a priori considérée et on calcule  $(G_i)_{i=1, \dots, N}$  où  $G(\theta_i) = \exp(-F(\theta_i))$  pour  $i = 1, \dots, N$ . Les choix  $N = 10D$  et  $N = 20D$  correspondent à des situations communes en pratique. Le choix  $N = 50D$  permet généralement d'avoir un modèle de krigeage extrêmement prédictif et sert ici de référence.

4. Pour chaque taille de l'échantillon des paramètres, on détermine la loi de distribution conditionnelle à partir des résultats de l'étape 3, que l'on compare aux lois a posteriori exactes.

- *Test pour  $\theta$  à une dimension ( $D = 1$ )*

On choisit  $\theta = 1$ .

Dans le cas d'une loi a priori Gaussienne, on fixe la moyenne  $\mu = 1.2$  et un écart-type de 0.4. On représente en figure 7 la loi a priori (trait discontinu gris), la loi a posteriori approchée (constitué de croix) et la loi a posteriori exacte (constitué de tirets) pour  $N = 10D$  et pour une covariance Gaussienne (figure 7a)) et de Matern (figure 7b)). On remarque que dans les deux cas l'approximation est bonne : en effet, les lois de distribution a posteriori sont confondues. Par conséquent, ajouter des points supplémentaires ( $N=20D$  ou  $N=50D$ ) n'est pas nécessaire.

Dans le cas d'une loi a priori uniforme, on fixe  $[a_1, b_1] = [0,2]$ . On représente en figure 8 la loi a priori (trait continu gris), la loi a posteriori approchée (constitué de croix) et la loi a posteriori exacte (constitué de tirets) pour  $N = 10D$  et une covariance Gaussienne (figure 8a)) et de Matern (figure 8b)). Comme précédemment, l'approximation est de bonne qualité.

- *Test pour  $\theta$  à deux dimensions ( $D = 2$ )*

On choisit  $\theta = [1 \ 1]$ .

Dans le cas d'une loi a priori Gaussienne, on fixe la moyenne  $\mu = [1,2 ; 1,2]$  et un écart-type de 0.4 pour chaque composante, toutes les deux étant indépendantes. On représente en figure 9 la loi a priori (trait continu en noir), la loi a posteriori approchée (constitué de croix) et la loi a posteriori exacte (constitué de tirets) pour  $N = 10D$  et pour une covariance Gaussienne (figures 9a) et 9c)) et de Matern (figure 9b) et 9d)). La première ligne (figures 9a) et 9b)) correspond à  $\theta_1$  et la seconde (figures 9c) et d)) à  $\theta_2$ . On remarque que la loi approchée pour le premier paramètre est légèrement décentrée par rapport à la loi exacte. En revanche, la loi approchée pour le second paramètre est proche de la loi exacte.

Dans le cas d'une loi a priori uniforme, on fixe  $[a_1, b_1] = [a_2, b_2] = [0,2]$ . Dans ce cas, c'est la loi du premier paramètre qui est mieux approchée, alors que la loi du second présente des oscillations.

En terme de loi a posteriori, on obtient des approximations raisonnables pour seulement  $N = 10D$ .

Lorsque l'on augmente le nombre d'échantillons pour atteindre  $N = 20D$ , les approximations sont meilleures: le cas de la loi a priori Gaussienne est nettement amélioré (Figure 10a) à 10d)) et celui de la loi a priori uniforme légèrement. Les figures 10a) à 10d) correspondent aux figures 9a) à 9d) pour  $N = 20D$ .

Enfin, pour  $N = 50D$ , on obtient de très bonnes approximations. Les meilleurs résultats sont obtenus avec la loi a priori Gaussienne, le premier paramètre du cas uniforme présentant encore quelques oscillations.

- *Test pour  $\theta$  à trois dimensions ( $D = 3$ )*

On choisit  $\theta = [1 \ 1 \ 1]$ .

Dans le cas d'une loi a priori Gaussienne, on fixe la moyenne  $\mu = [1,2 \ 1,2 \ 1,2]$  et un écart-type de 0,4 pour chaque composante, toutes les trois étant indépendantes. On représente en figure 11 la loi a priori (trait continu gris), la loi a posteriori approchée (trait continu en noir) et la loi a posteriori exacte (constitué de tirets) pour  $N = 10D$  et pour une covariance Gaussienne (figures 11a), c) et e)) et de Matern (figure 11b),d) et f)). La première ligne (figures 11a) et b)) correspond à  $\theta_1$ , la seconde (figures 11c) et d)) à  $\theta_2$ , la troisième (figures 11e) et f)) à  $\theta_3$ .

Dans le cas d'une loi a priori uniforme, on fixe  $[a_1, b_1] = [a_2, b_2] = [a_3, b_3] = [0,2]$ . Dans ce cas, les approximations sont moins précises.

Lorsque l'on augmente le nombre d'échantillons pour atteindre  $N = 20D$ , les approximations sont meilleures : le cas de la loi a priori Gaussienne est nettement amélioré et celui de la loi a priori uniforme légèrement.

5 Enfin, pour  $N = 50D$ , on obtient de très bonnes approximations. Les figures 12a) à f) correspondent aux figures 11a) à f) pour  $N = 50D$ .

10 Les différents tests montrent que lorsque la dimension  $D$  augmente, les résultats se dégradent. C'est un résultat attendu : plus l'espace des paramètres est grand, plus l'approximation par krigeage perd en précision. Cependant, les choix classiques du nombre de points dans l'espace des paramètres  $N = 10D$  et  $N = 20D$  fournissent malgré tout des approximations correctes. Pour cet exemple, les tests ne font pas appel à une procédure itérative pour rajouter des points. En pratique, lorsque l'on augmente la taille de l'ensemble, les points sont placés dans les régions d'intérêt et les approximations deviennent moins sensibles à la dimension.

15 On note que pour les simulations entreprises pour ces exemples, la loi analytique a posteriori obtenue grâce au procédé selon l'invention requiert un temps de calcul de l'ordre du centième de seconde, alors qu'un échantillonnage MCMC de l'art antérieur nécessite plusieurs minutes.

### Revendications

- 1) Procédé d'exploitation d'un réservoir géologique selon un schéma d'exploitation défini à partir d'un modèle de réservoir, ledit modèle de réservoir comportant un maillage associé à des paramètres  $\theta$  dudit réservoir, caractérisé en ce qu'on réalise les étapes suivantes :
- 5 a) on construit un modèle de réservoir calé sur des données mesurées au sein dudit réservoir au moyen des étapes suivantes :
- i) on génère de manière stochastique un ensemble initial de modèles de réservoir à partir de lois de probabilité  $p(\theta)$  desdits paramètres  $\theta$  ;
- 10 ii) on détermine une fonction objectif  $F(\theta)$  mesurant un écart entre des données dynamiques  $y_1, \dots, y_n$  acquises en cours d'exploitation et des données dynamiques simulées au moyen d'un simulateur d'écoulement et desdits modèles de réservoir appartenant audit ensemble ;
- 15 iii) on détermine à partir de ladite fonction objectif  $F(\theta)$  une loi analytique  $p(\theta|y_1, \dots, y_n)$  de probabilité conditionnelle des paramètres  $\theta$  connaissant lesdites données dynamiques mesurées  $y_1, \dots, y_n$  ;
- iv) on génère au moins un nouveau modèle de réservoir au moyen de ladite loi analytique  $p(\theta|y_1, \dots, y_n)$  et on ajoute ledit nouveau modèle audit ensemble de modèles de réservoir ;
- 20 v) on réitère les étapes ii) à iv) et on détermine le modèle de réservoir qui minimise ladite fonction objectif ;
- b) on détermine un schéma d'exploitation optimal du réservoir en simulant l'exploitation du réservoir au moyen dudit modèle de réservoir calé et dudit simulateur d'écoulement ; et
- c) on exploite ledit réservoir en mettant en œuvre ledit schéma d'exploitation optimal.
- 25
- 2) Procédé selon la revendication 1, dans lequel on détermine ladite loi analytique conditionnelle  $p(\theta|y_1, \dots, y_n)$  en déterminant une approximation d'une fonction  $G(\theta)$  telle que  $G(\theta) = \exp(-F(\theta))$ , puis en calculant ladite loi analytique à partir de la relation
- 30 suivante : 
$$p(\theta|y_1, \dots, y_n) = \frac{G(\theta) \times p(\theta)}{\int G(\theta) \times p(\theta) d\theta}.$$

- 3) Procédé selon la revendication 2, dans lequel on détermine l'approximation de ladite fonction  $G(\theta)$  en calculant  $G(\theta_i) = \exp(-F(\theta_i))$  pour les modèles  $i$  appartenant audit ensemble, puis par interpolation desdites valeurs de  $G(\theta_i)$ .
- 5 4) Procédé selon la revendication 3, dans lequel on interpole lesdites valeurs de  $G(\theta_i)$  par une technique de krigeage.
- 10 5) Procédé selon l'une des revendications précédentes, dans lequel à l'étape iv) à chaque itération on génère plusieurs nouveaux modèles de réservoir, parmi lesquels on choisit  $M_1$  modèles qu'on ajoute audit ensemble, ledit ensemble de modèles de réservoir est constitué lors de la construction de l'ensemble initial d'un nombre  $M_0$  de modèles.
- 15 6) Procédé selon la revendication 5, dans lequel lesdits  $M_0$  modèles dudit ensemble initial sont échantillonnés par une technique de plan d'expériences de type hypercube latin.
- 7) Procédé selon l'une des revendications 5 et 6, dans lequel lesdits  $M_1$  modèles ajoutés dans ledit ensemble sont choisis à l'aide de l'approximation de ladite fonction  $G(\theta)$ .
- 20 8) Procédé selon la revendication 7, dans lequel au moins une partie desdits  $M_1$  modèles sont les modèles pour lesquels l'approximation de la fonction  $G(\theta)$  donne les valeurs maximales.
- 25 9) Procédé selon l'une des revendications 7 ou 8, dans lequel au moins une partie desdits  $M_1$  modèles sont les modèles pour lesquels la qualité de l'approximation de la fonction  $G(\theta)$  possède des variances de krigeage les plus élevées parmi lesdits nouveaux modèles.
- 30 10) Procédé selon l'une des revendications précédentes, dans lequel on réalise une étude de sensibilité par le calcul de plusieurs fonctions objectif, chaque fonction objectif étant calculée pour une partie des paramètres ou pour un intervalle de temps limité par rapport aux informations disponibles.
- 11) Produit programme d'ordinateur téléchargeable depuis un réseau de communication et/ou enregistré sur un support lisible par ordinateur et/ou exécutable par un processeur,

comprenant des instructions de code de programme pour la mise en œuvre du procédé selon l'une quelconque des revendications précédentes, lorsque ledit programme est exécuté sur un ordinateur.

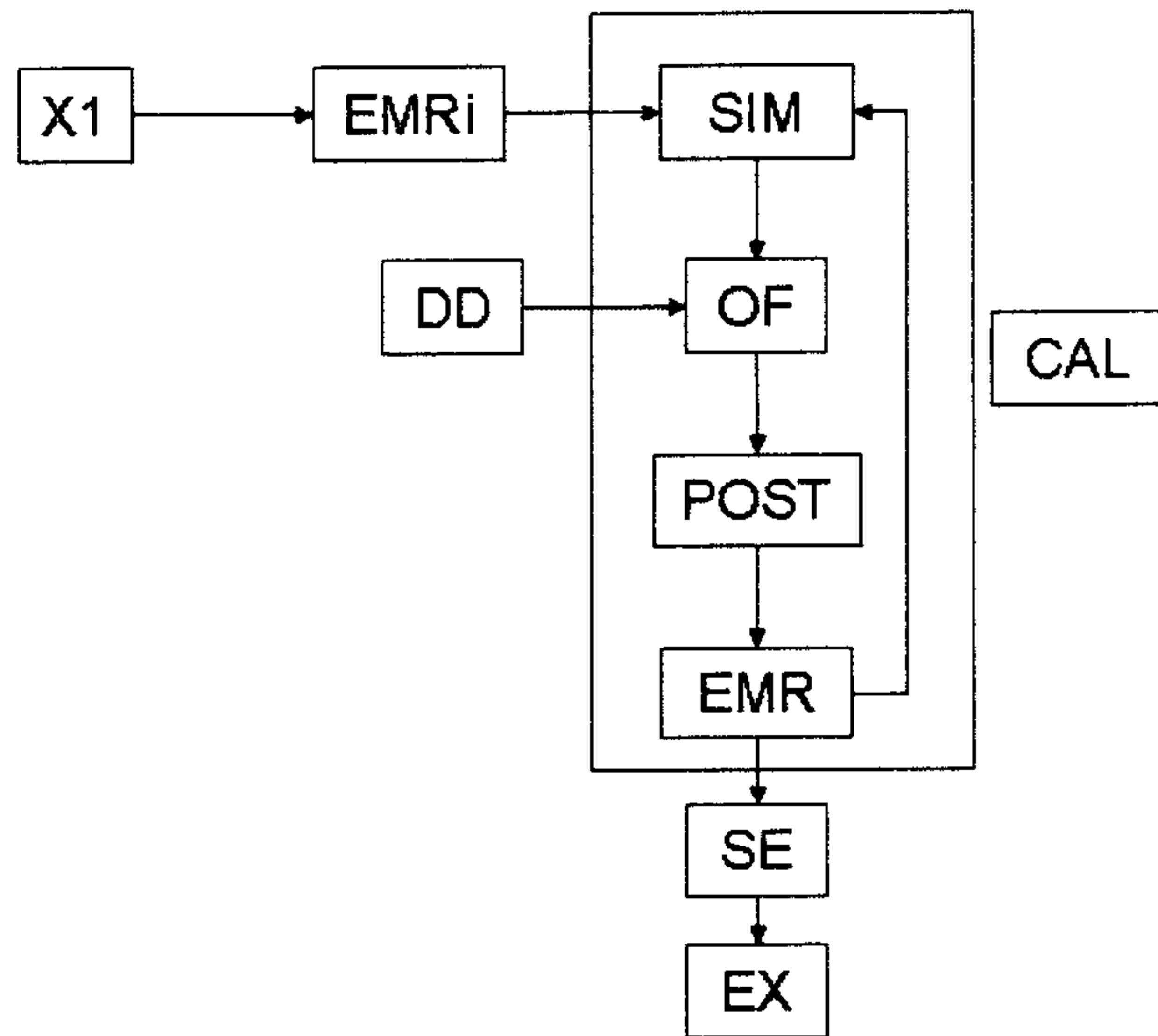


Figure 1

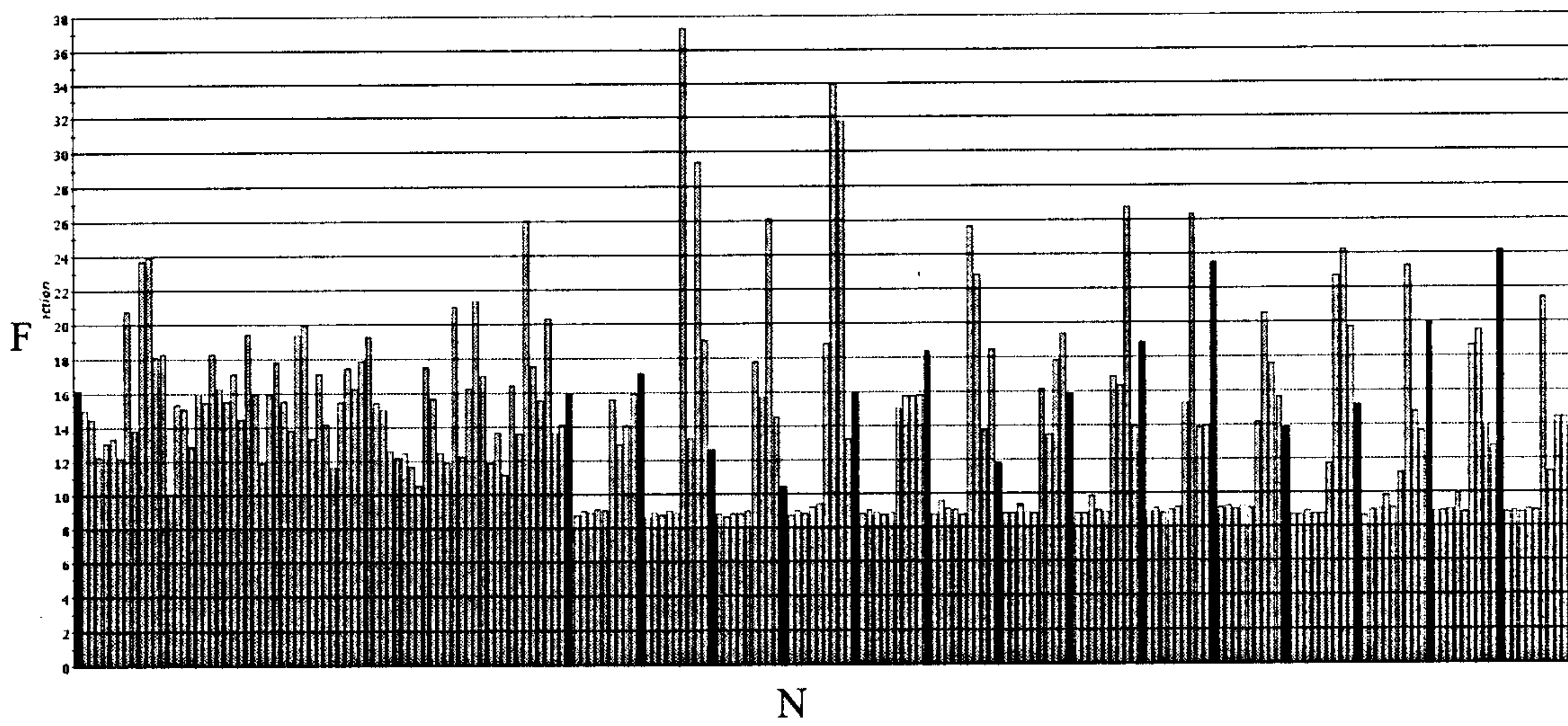


Figure 2

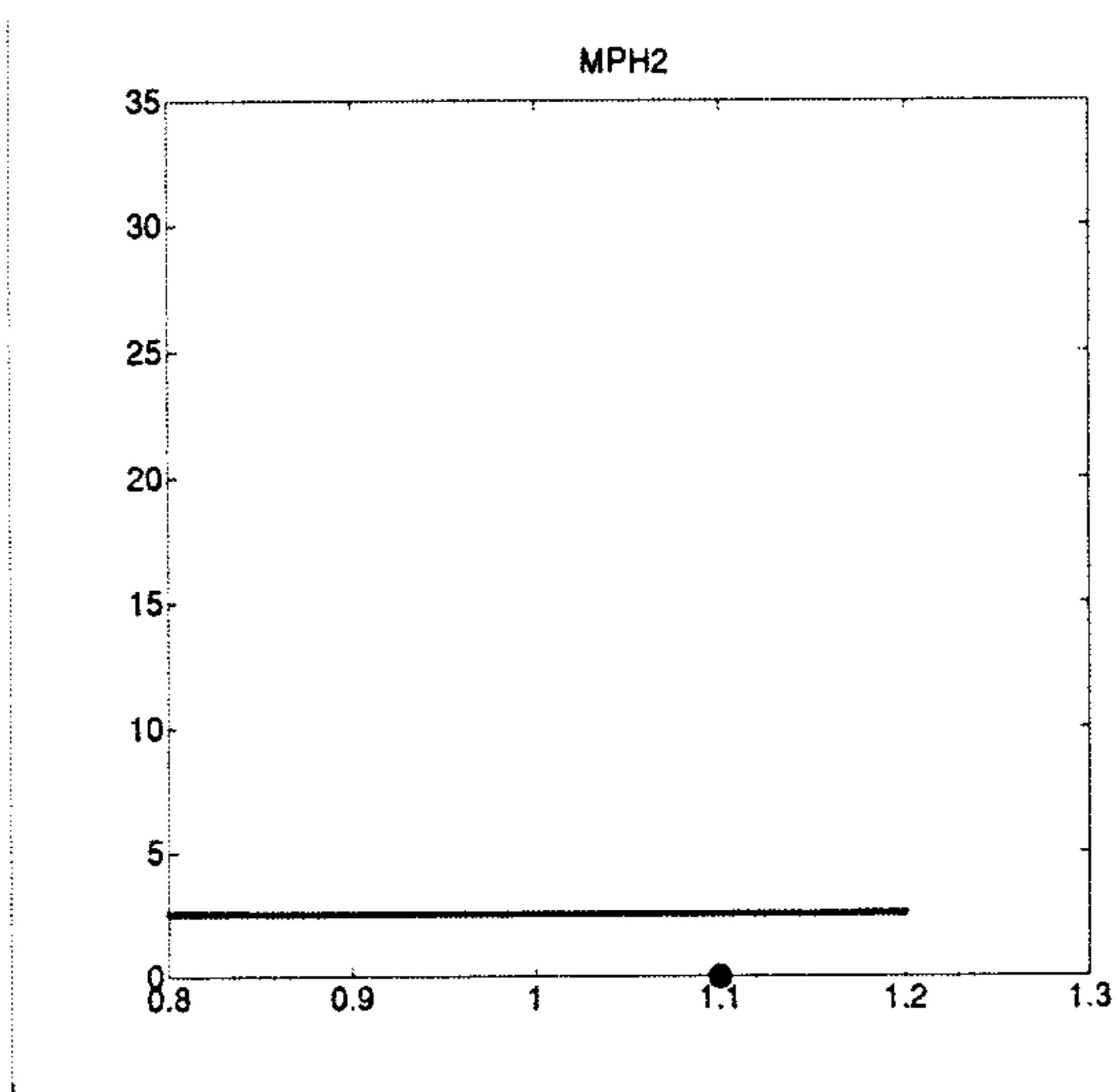


Figure 3a)

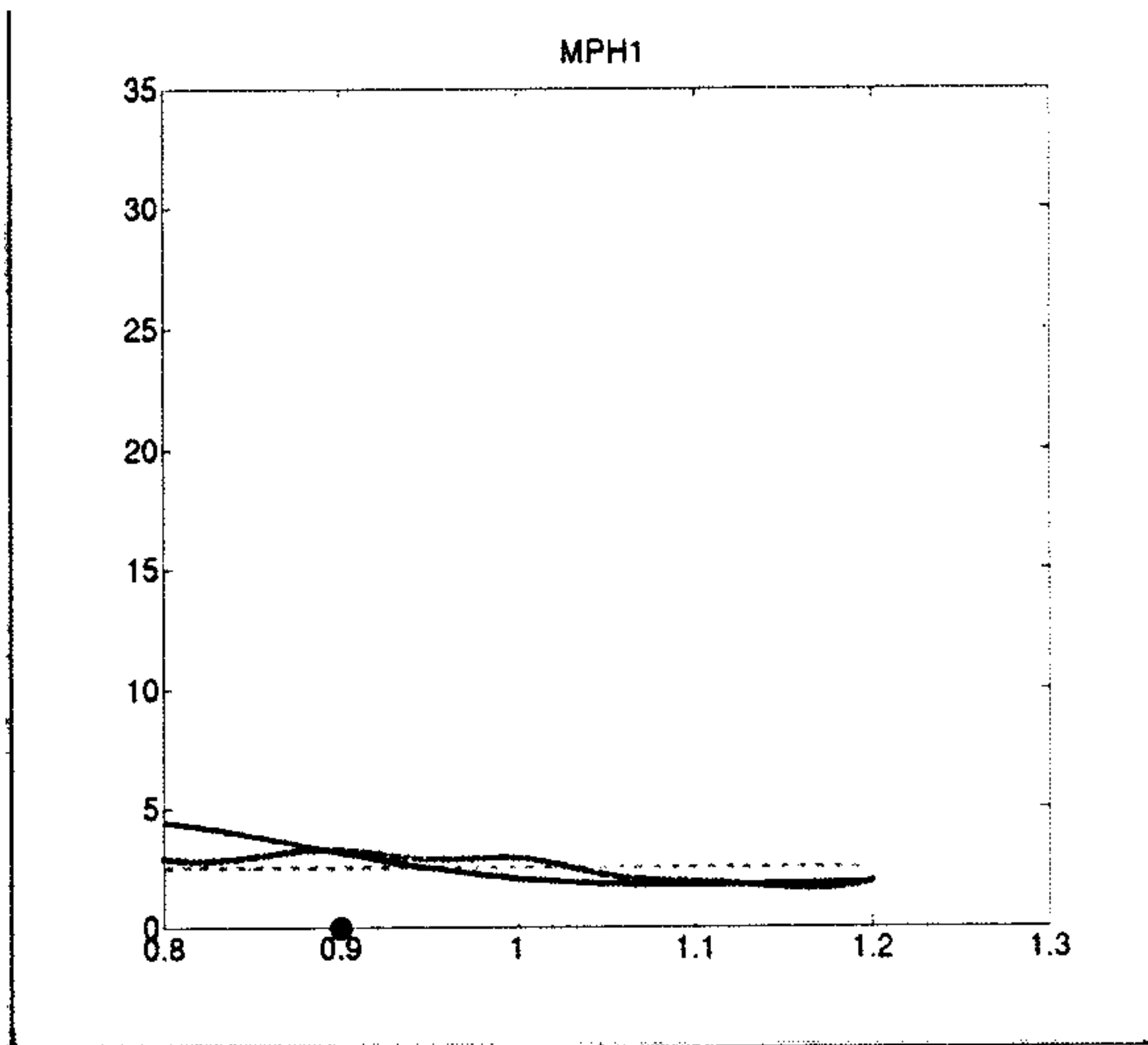


Figure 3b)

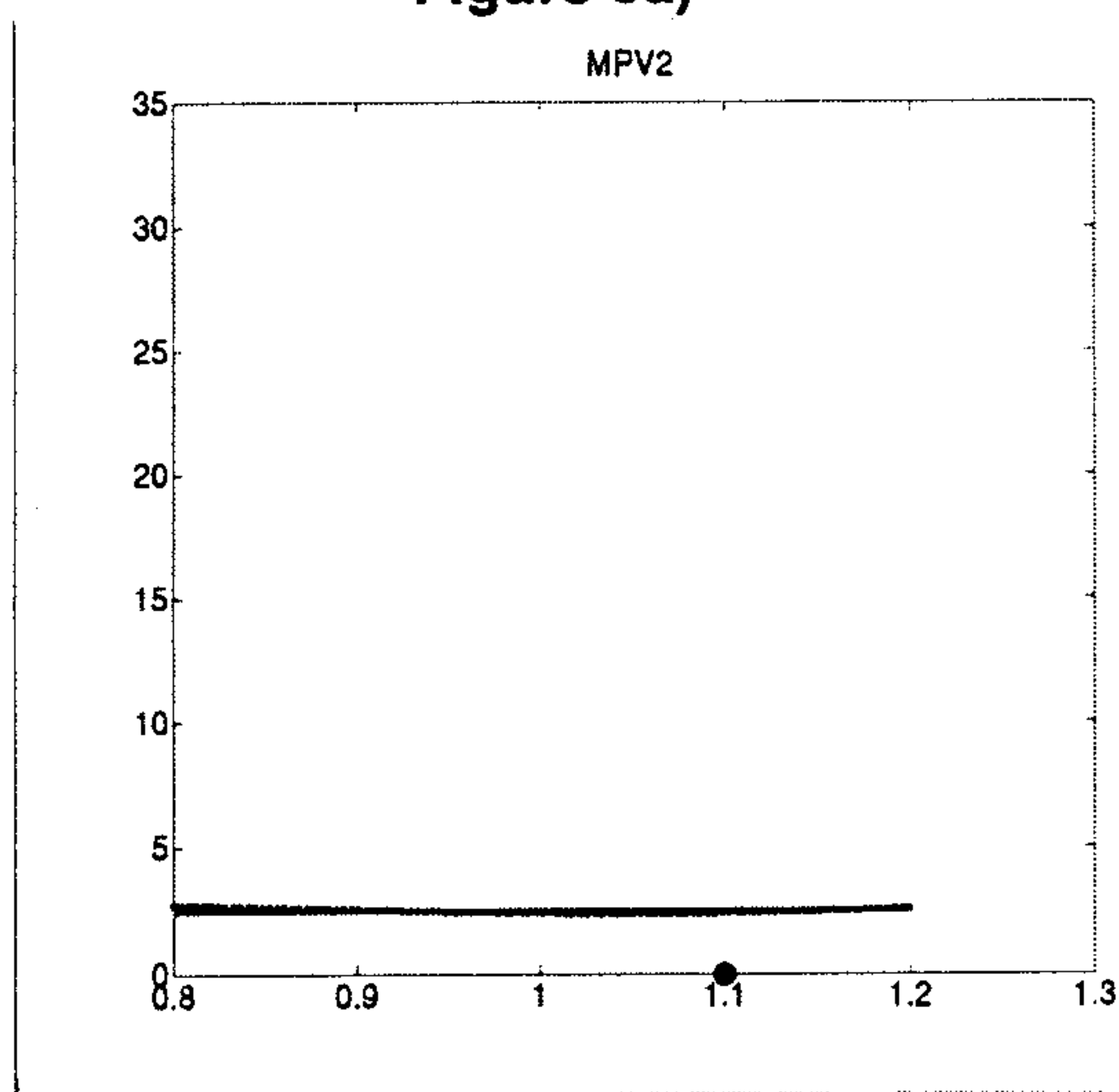


Figure 3c)

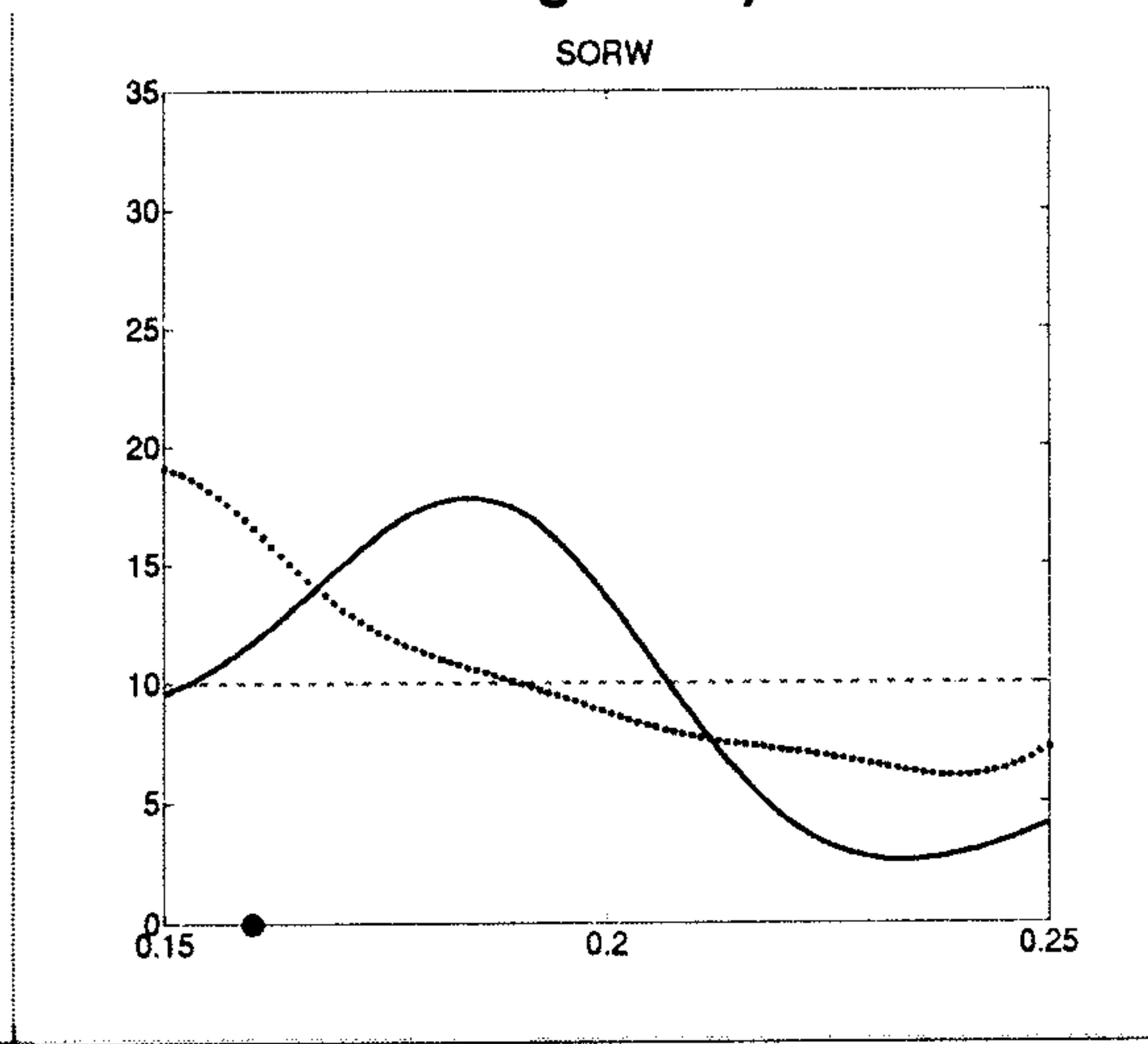


Figure 3d)

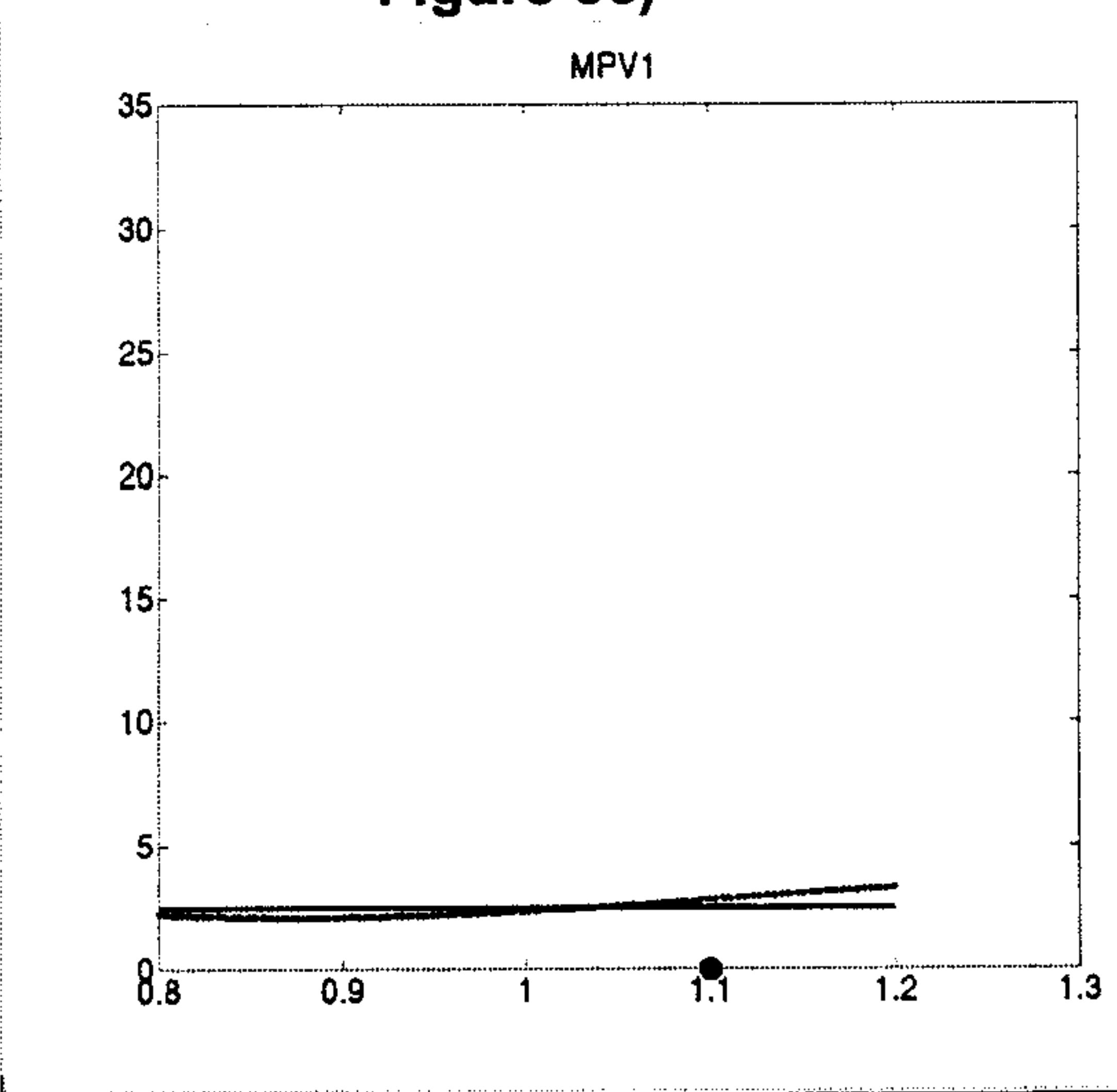


Figure 3e)

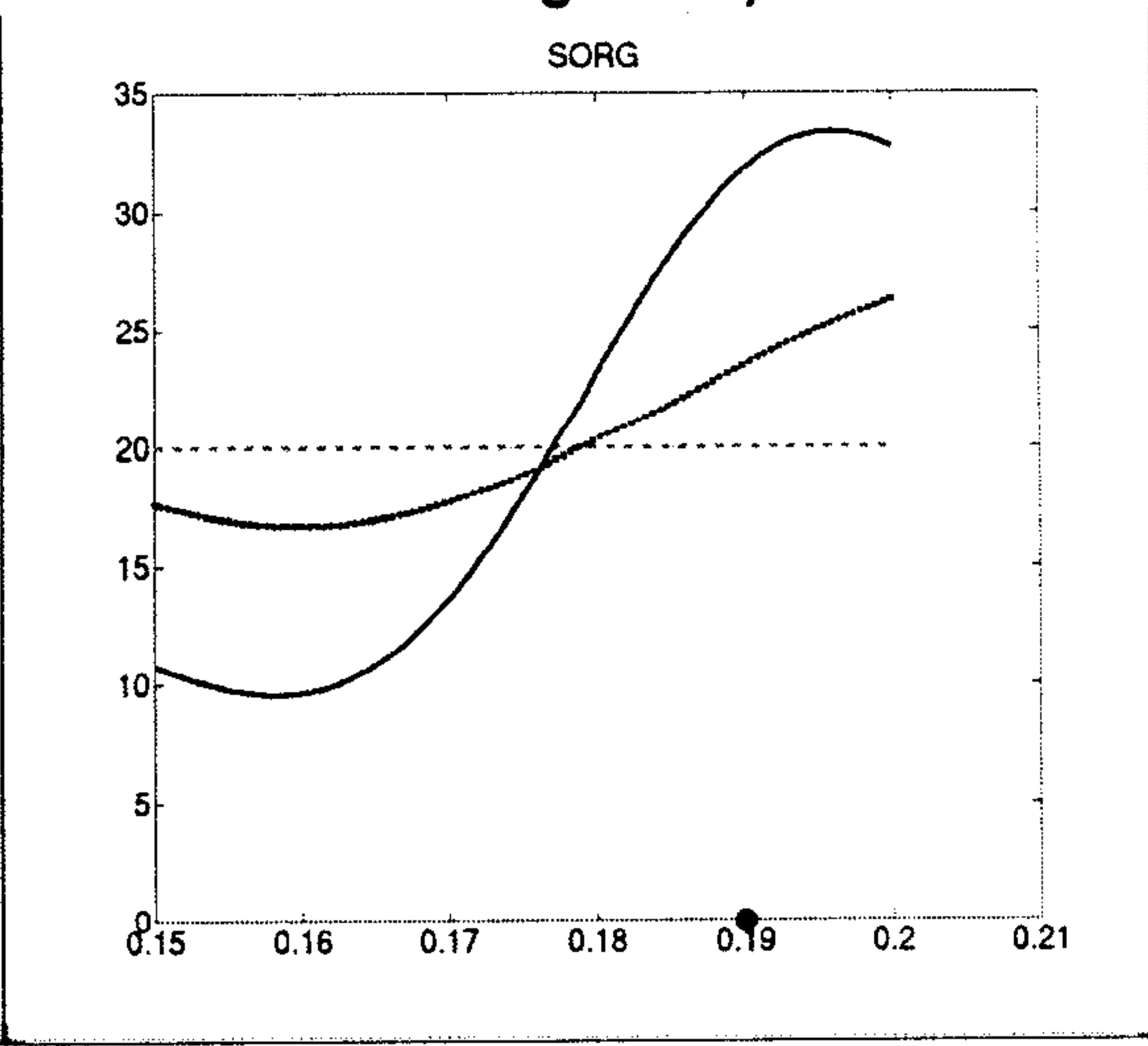


Figure 3f)

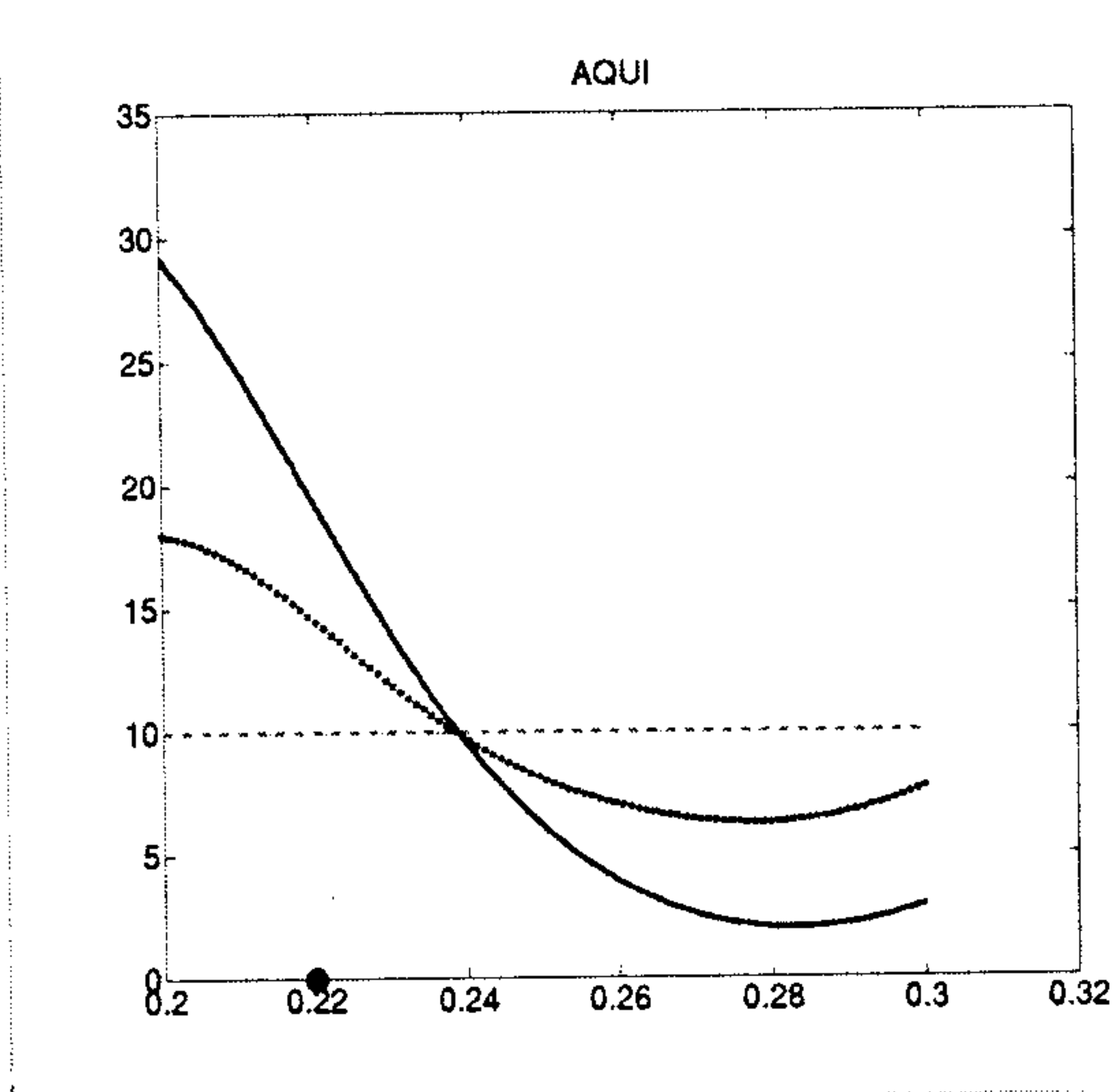


Figure 3g)

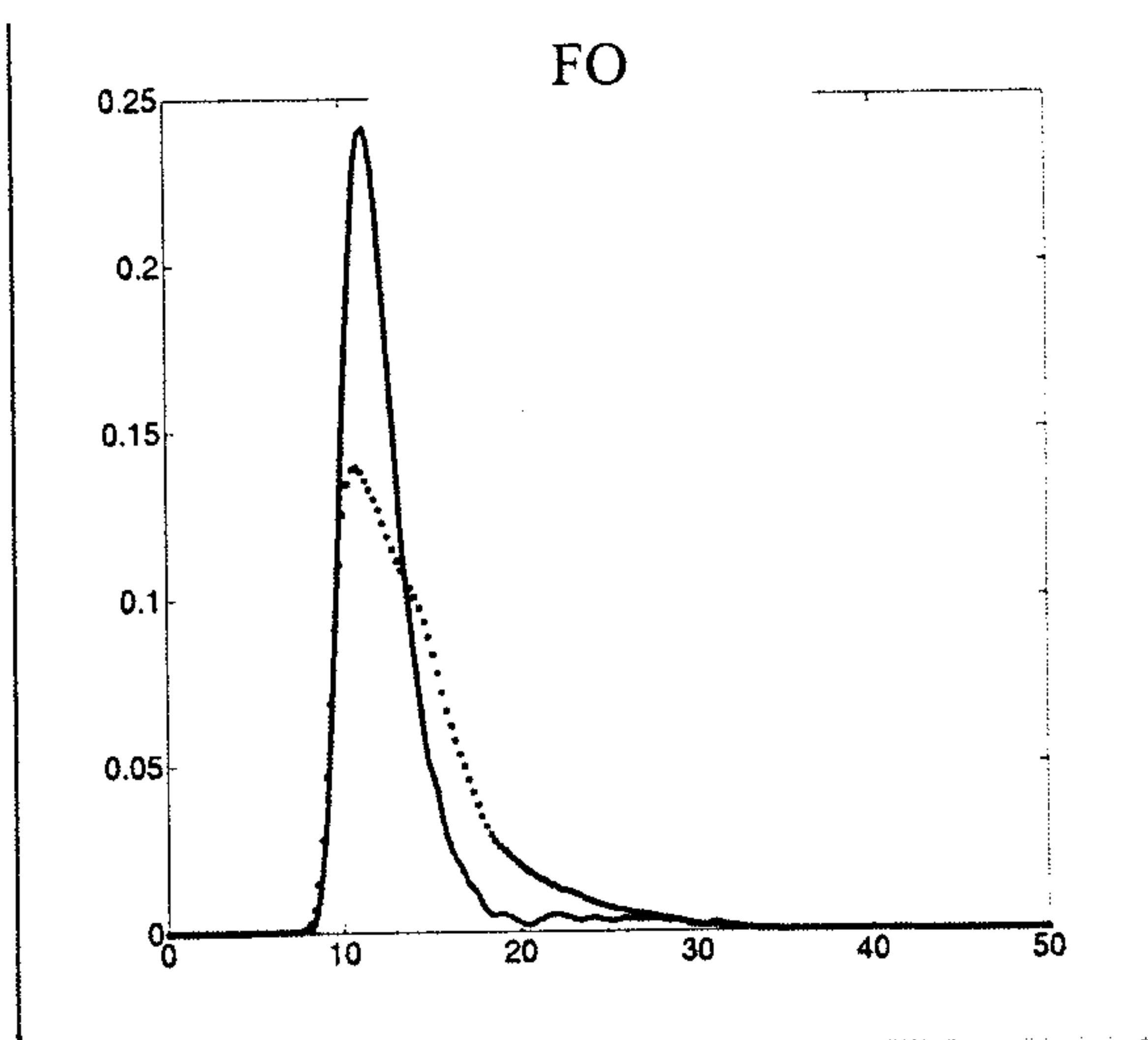


Figure 3h)

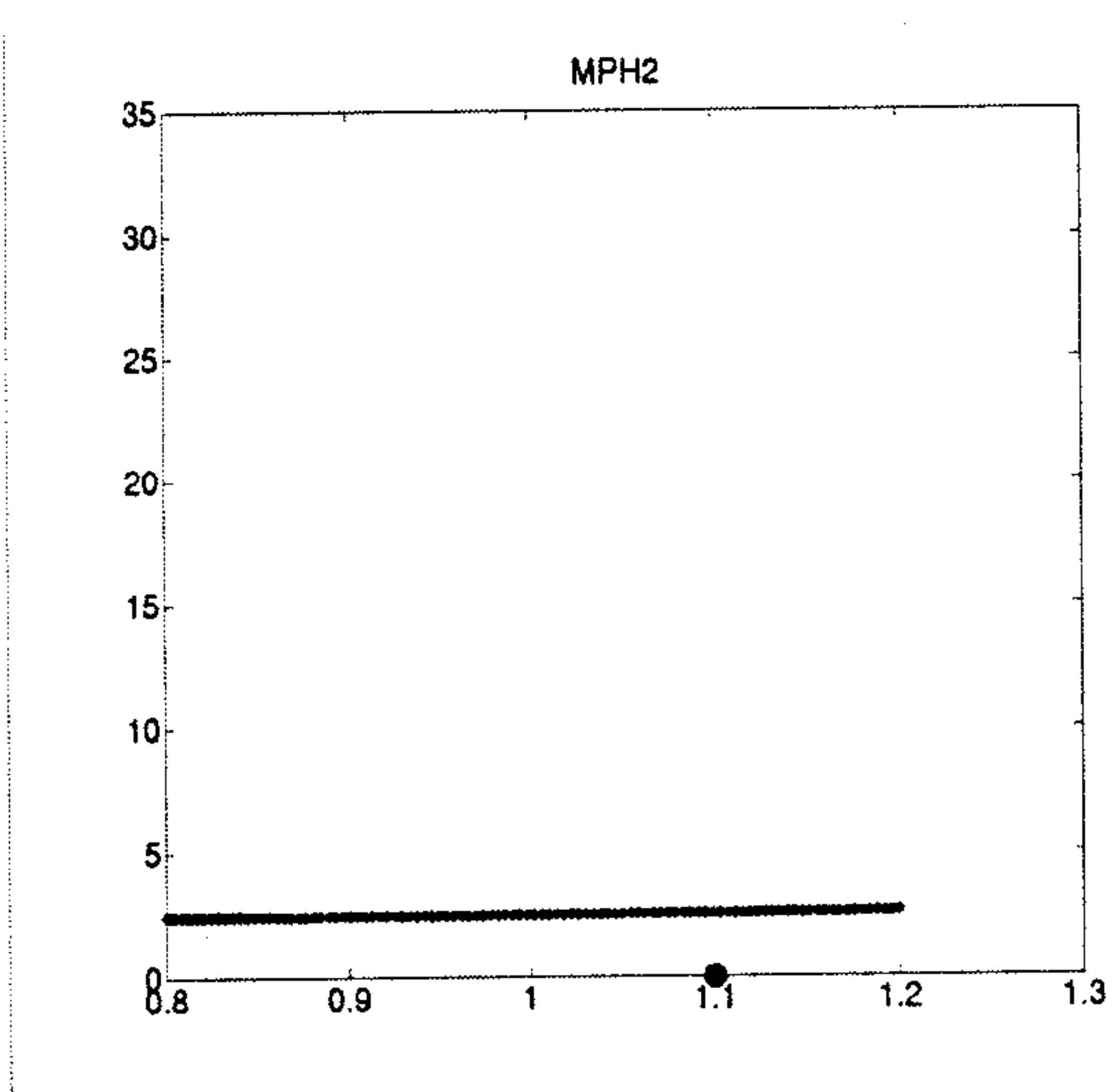


Figure 4a)

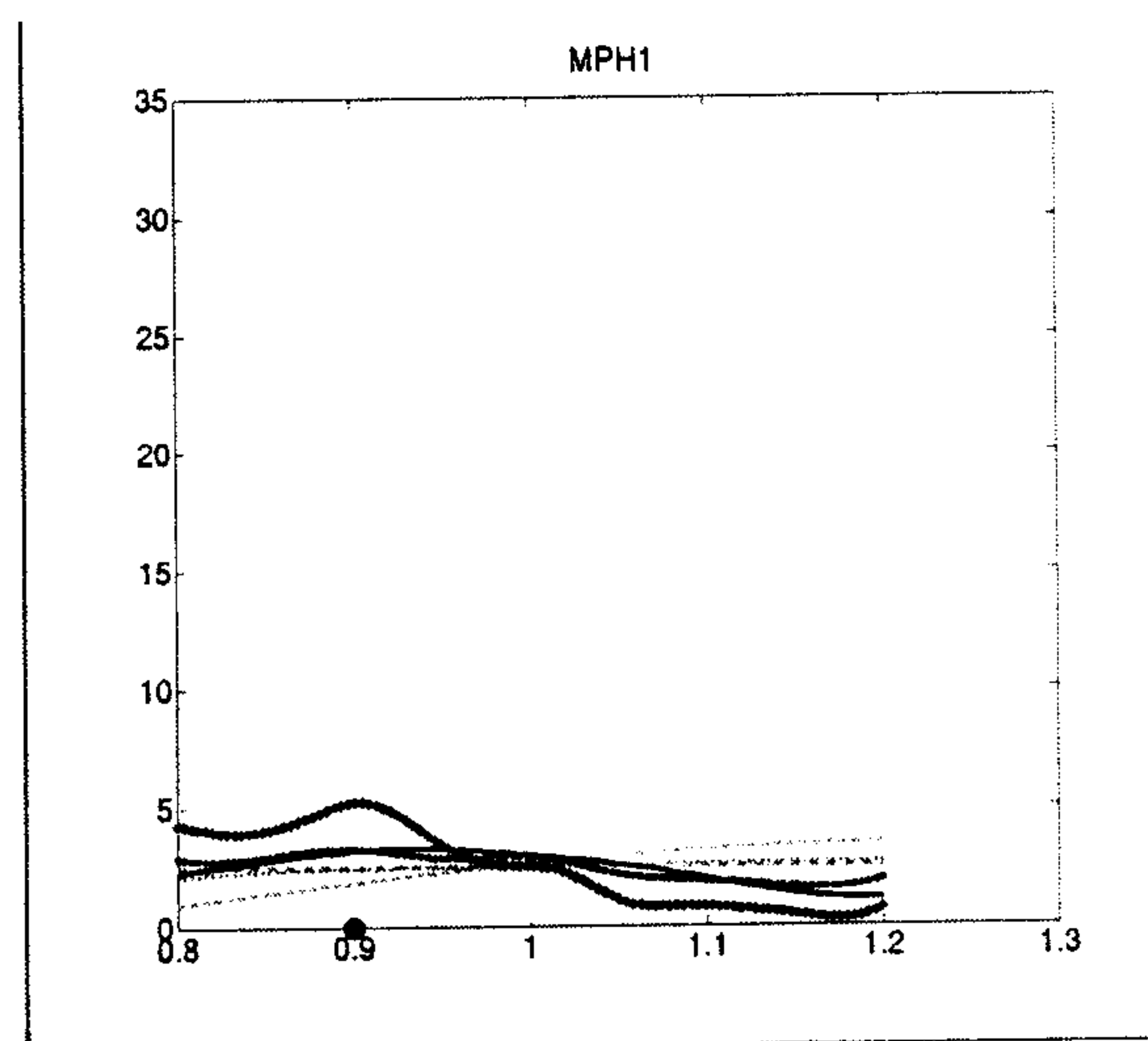


Figure 4b)

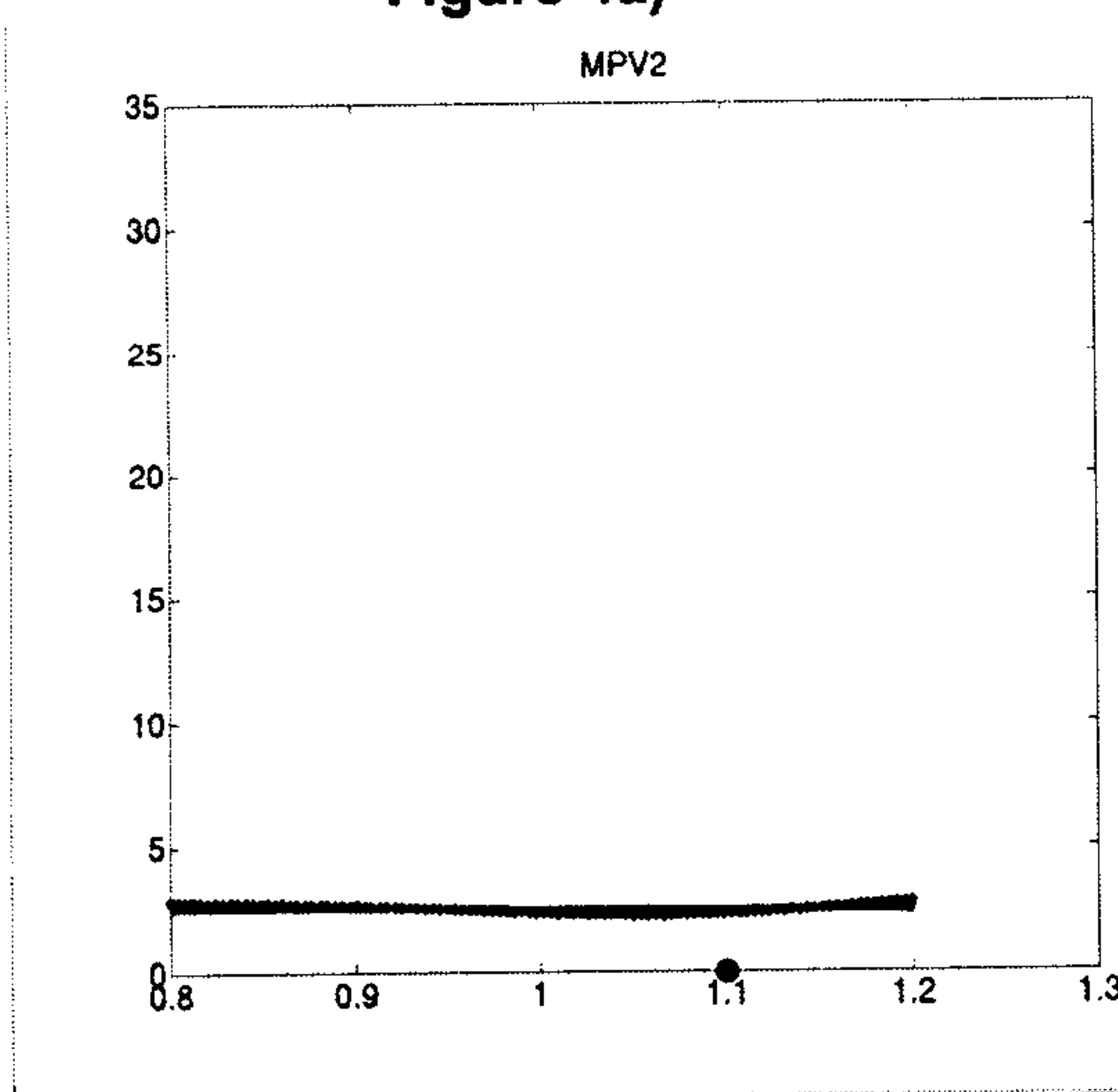


Figure 4c)

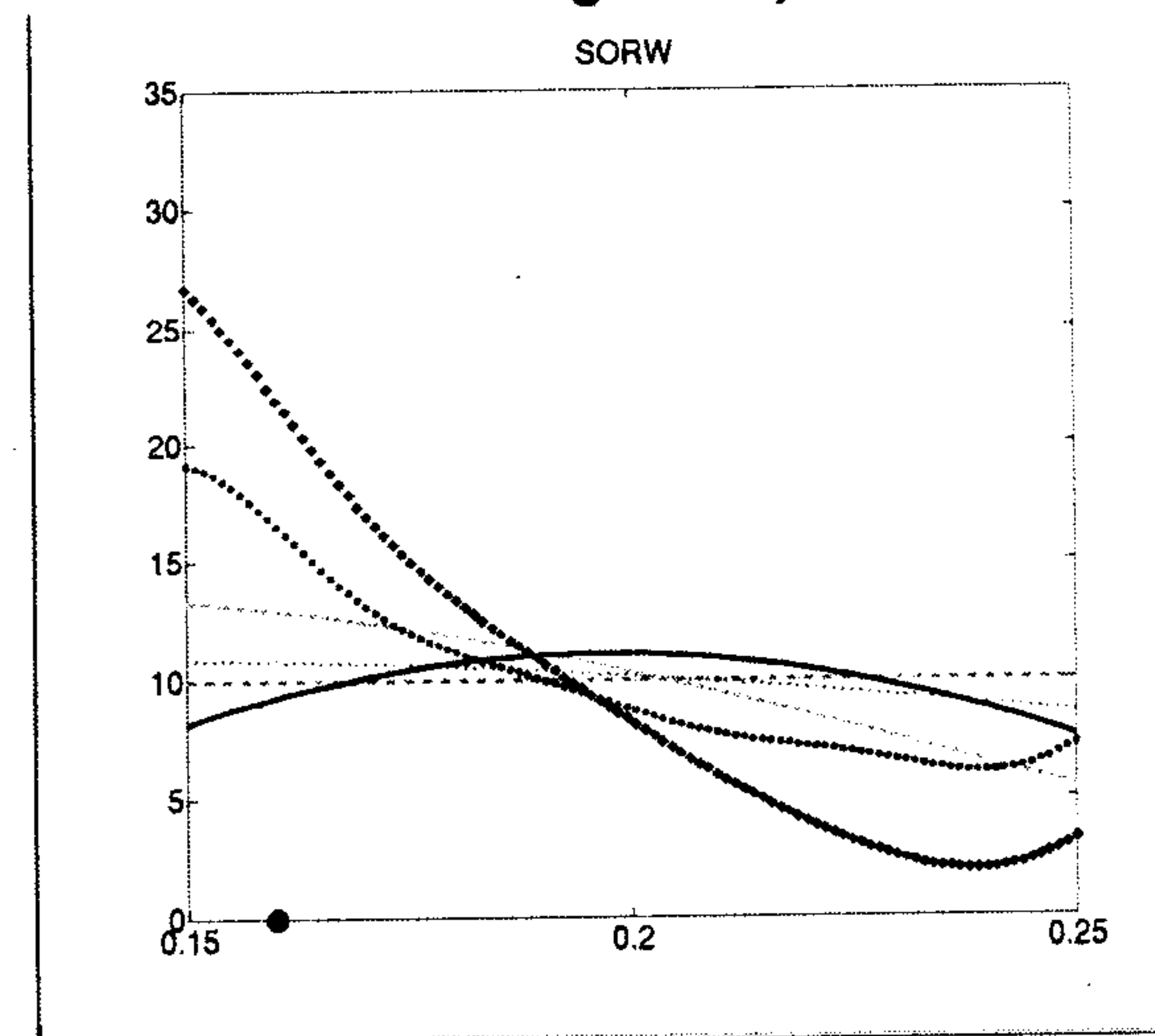


Figure 4d)

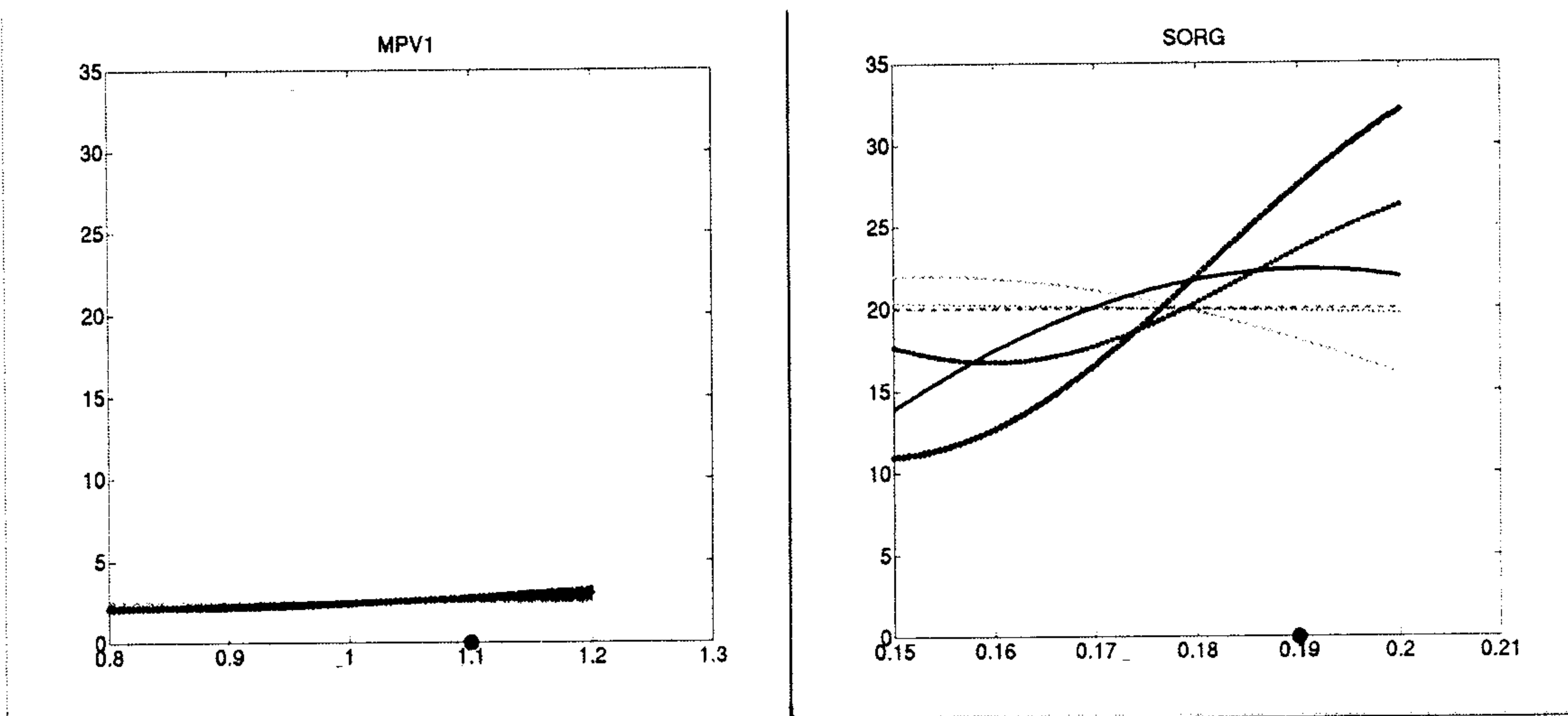


Figure 4e)

Figure 4f)

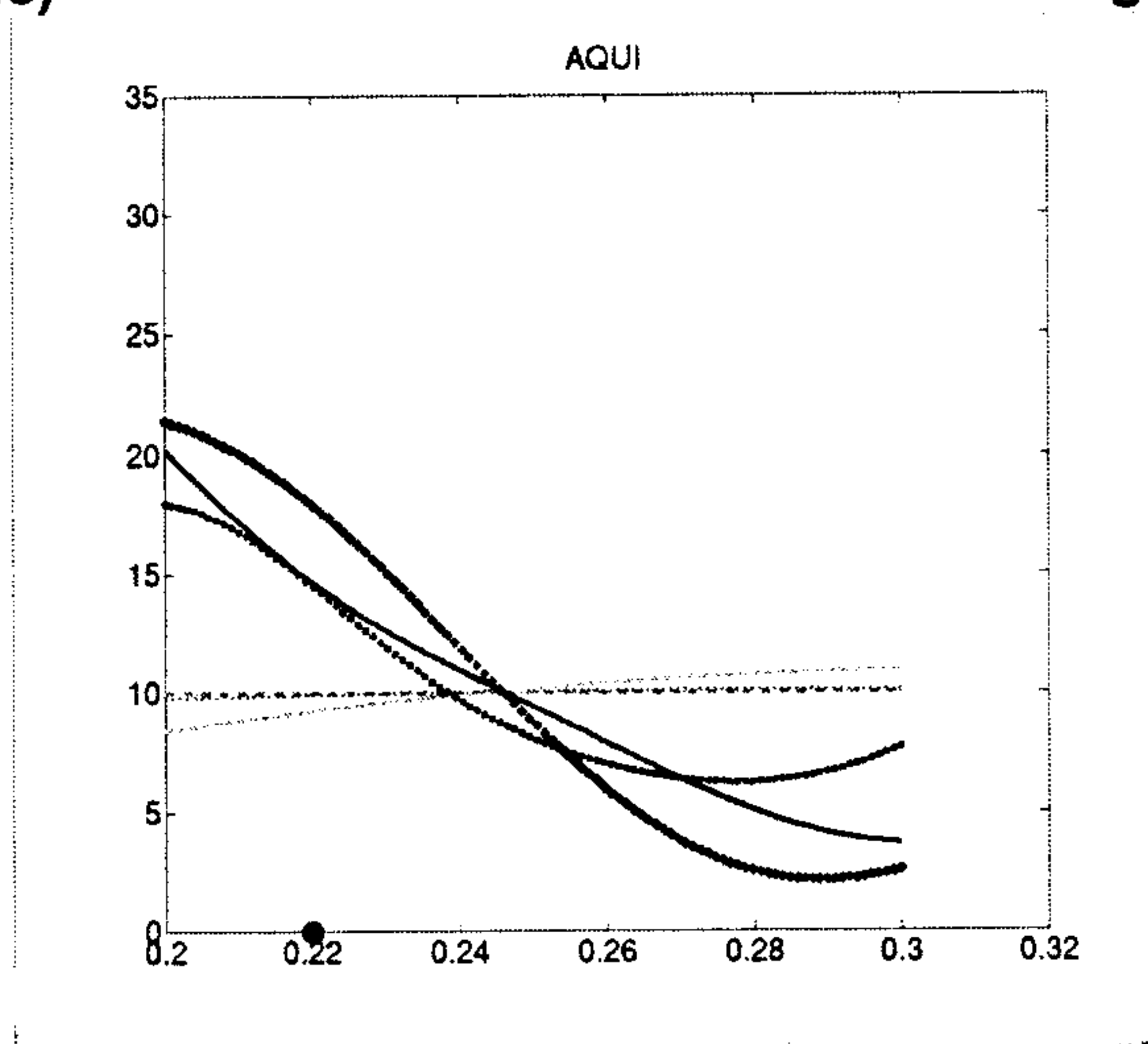


Figure 4g)

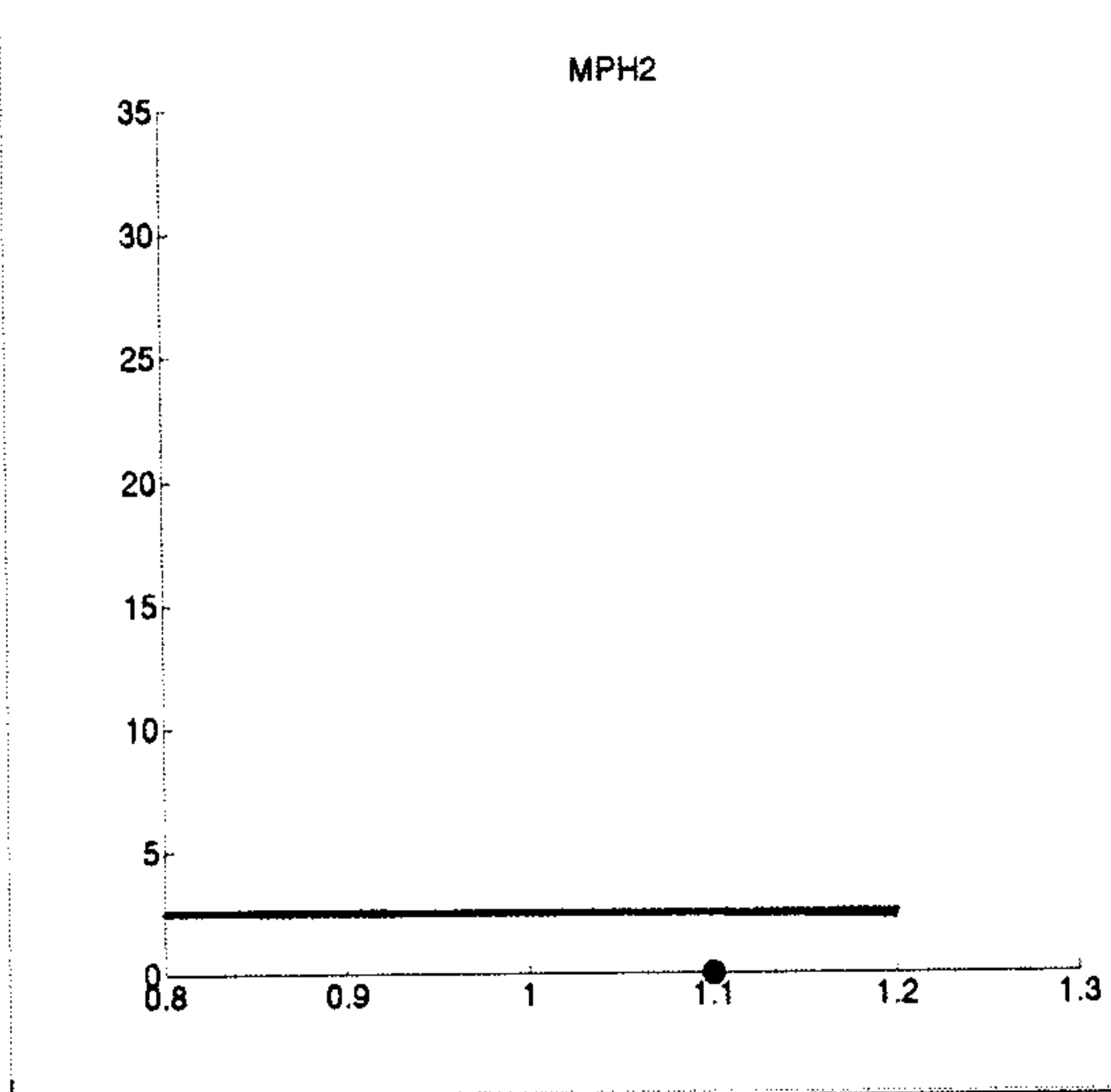


Figure 5a)

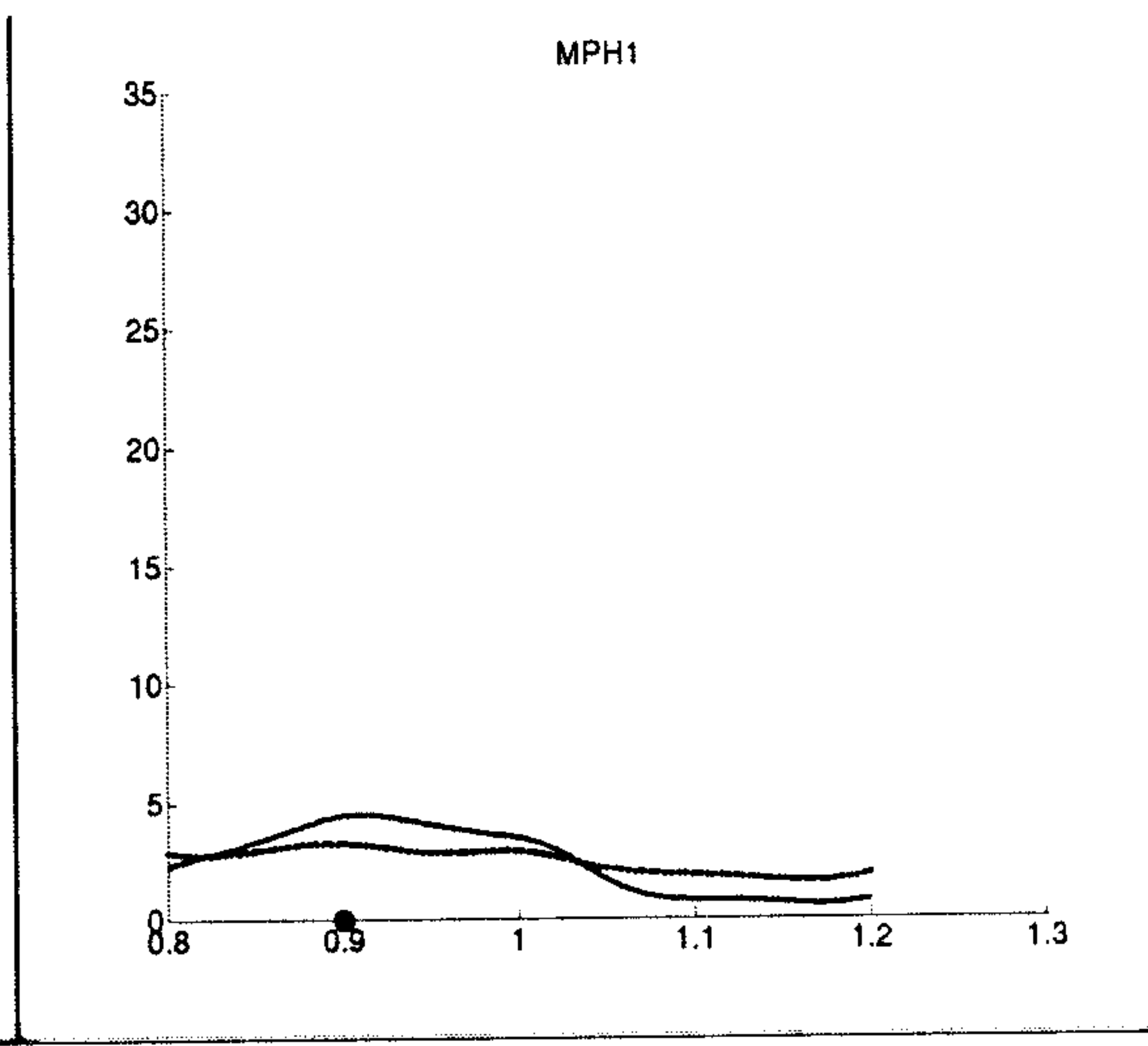


Figure 5b)

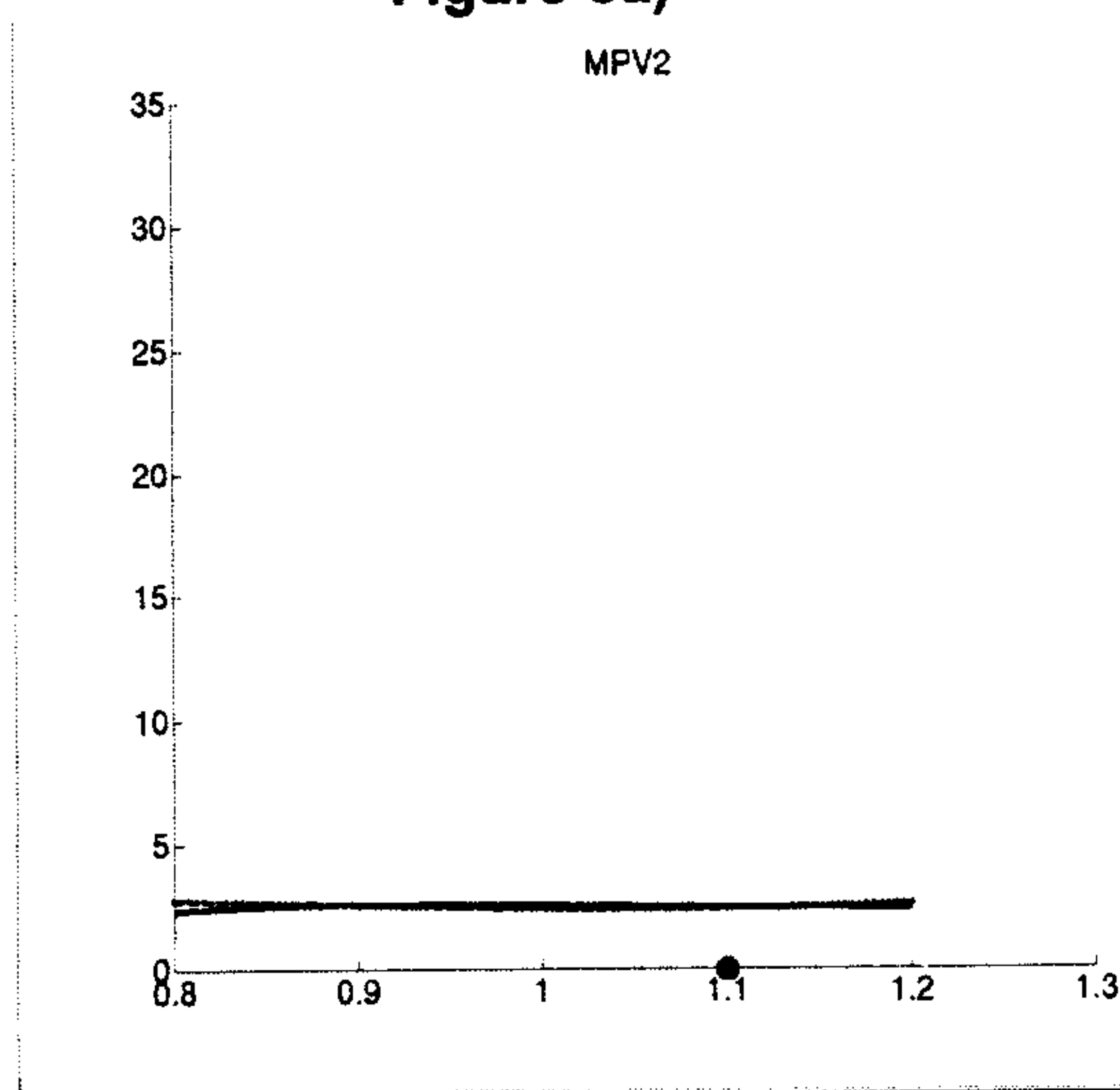


Figure 5c)

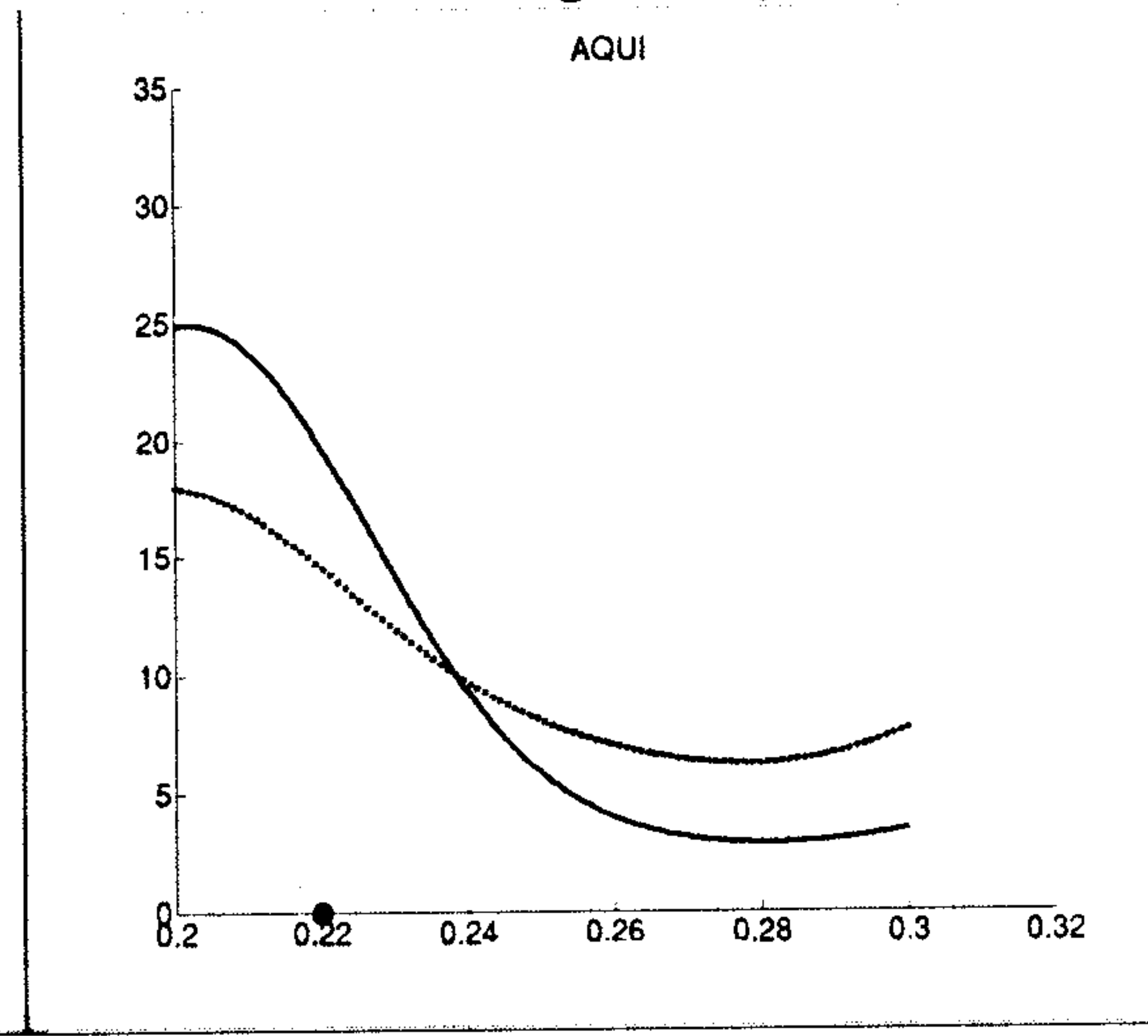


Figure 5d)

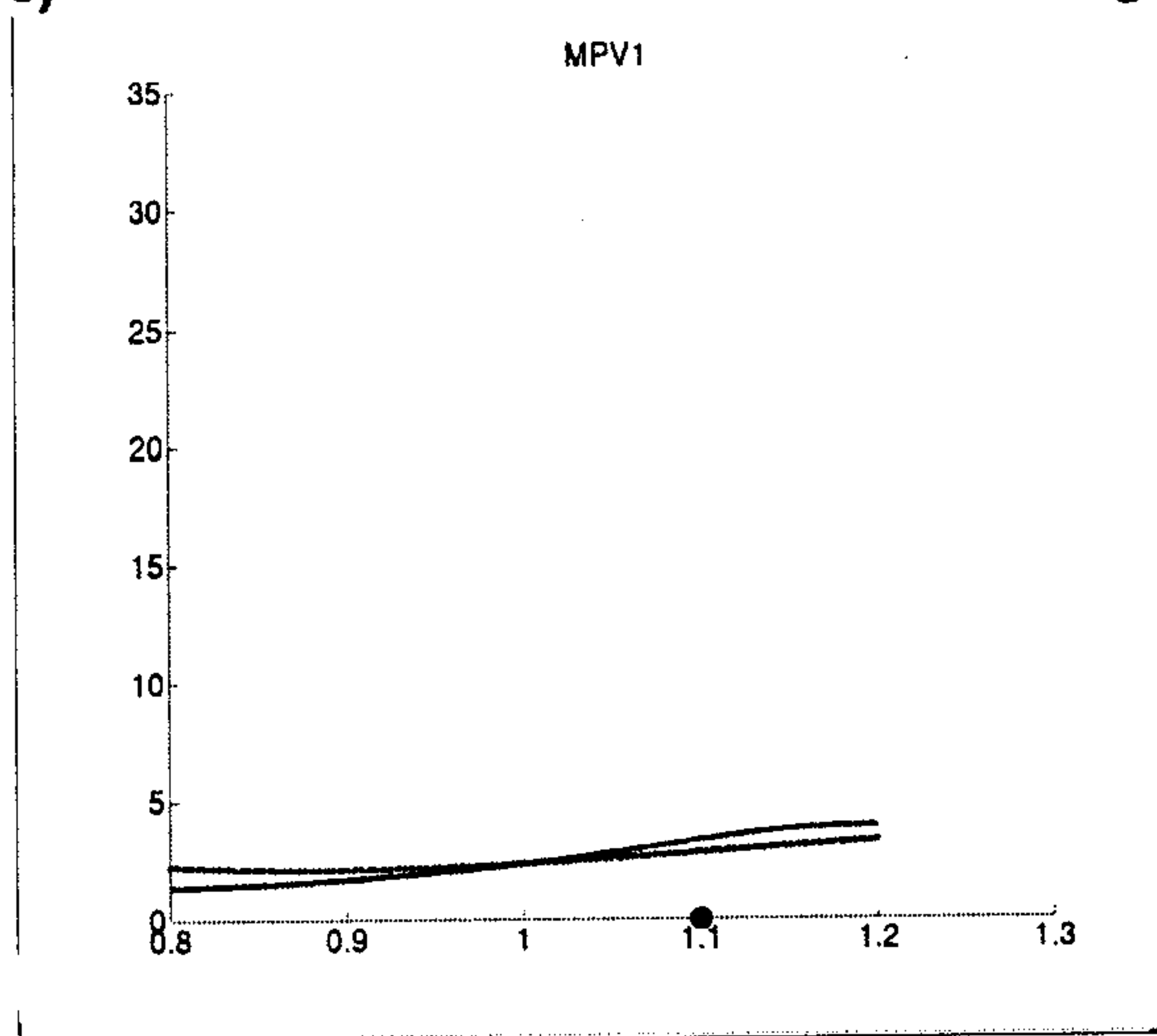


Figure 5e)

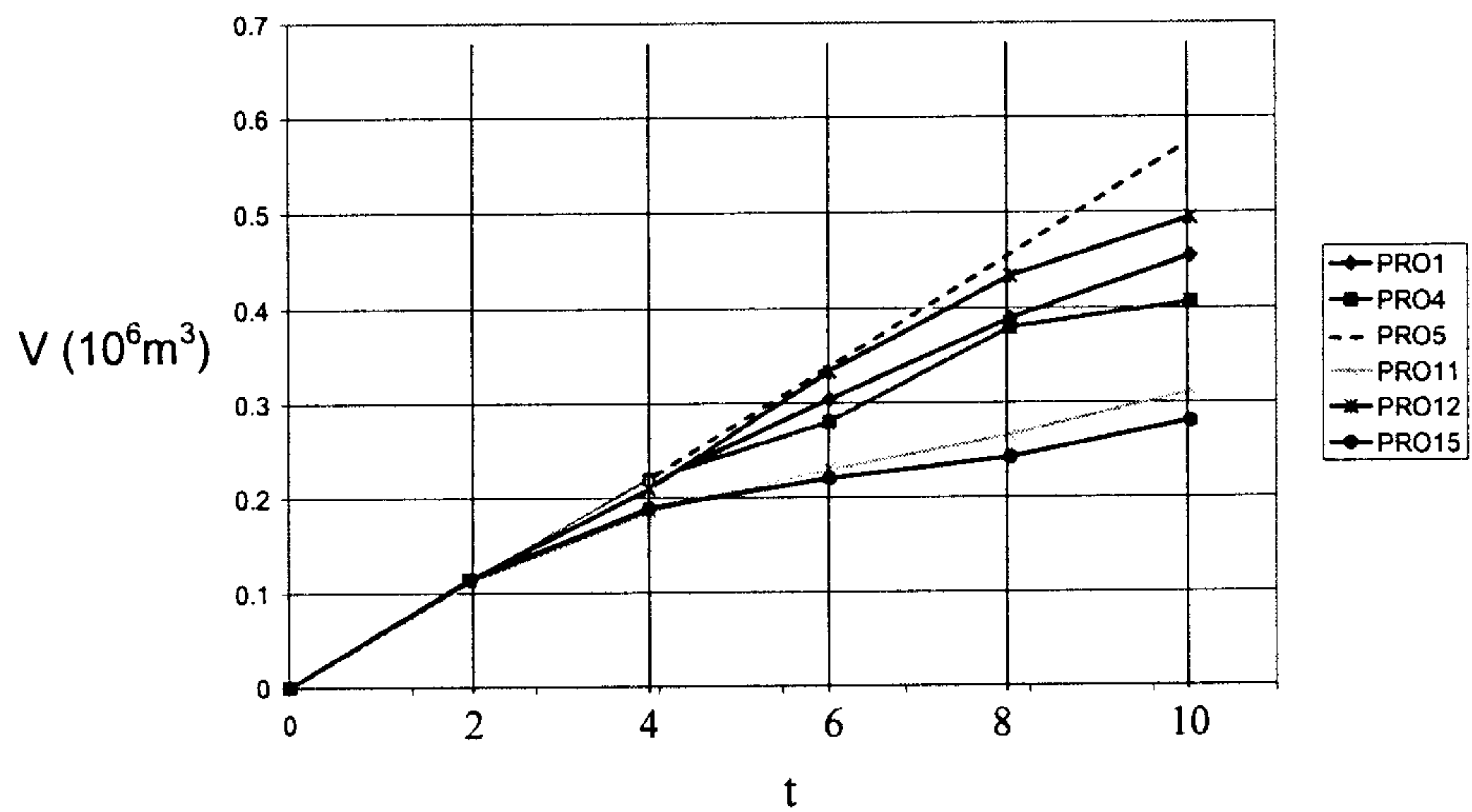


Figure 6a)

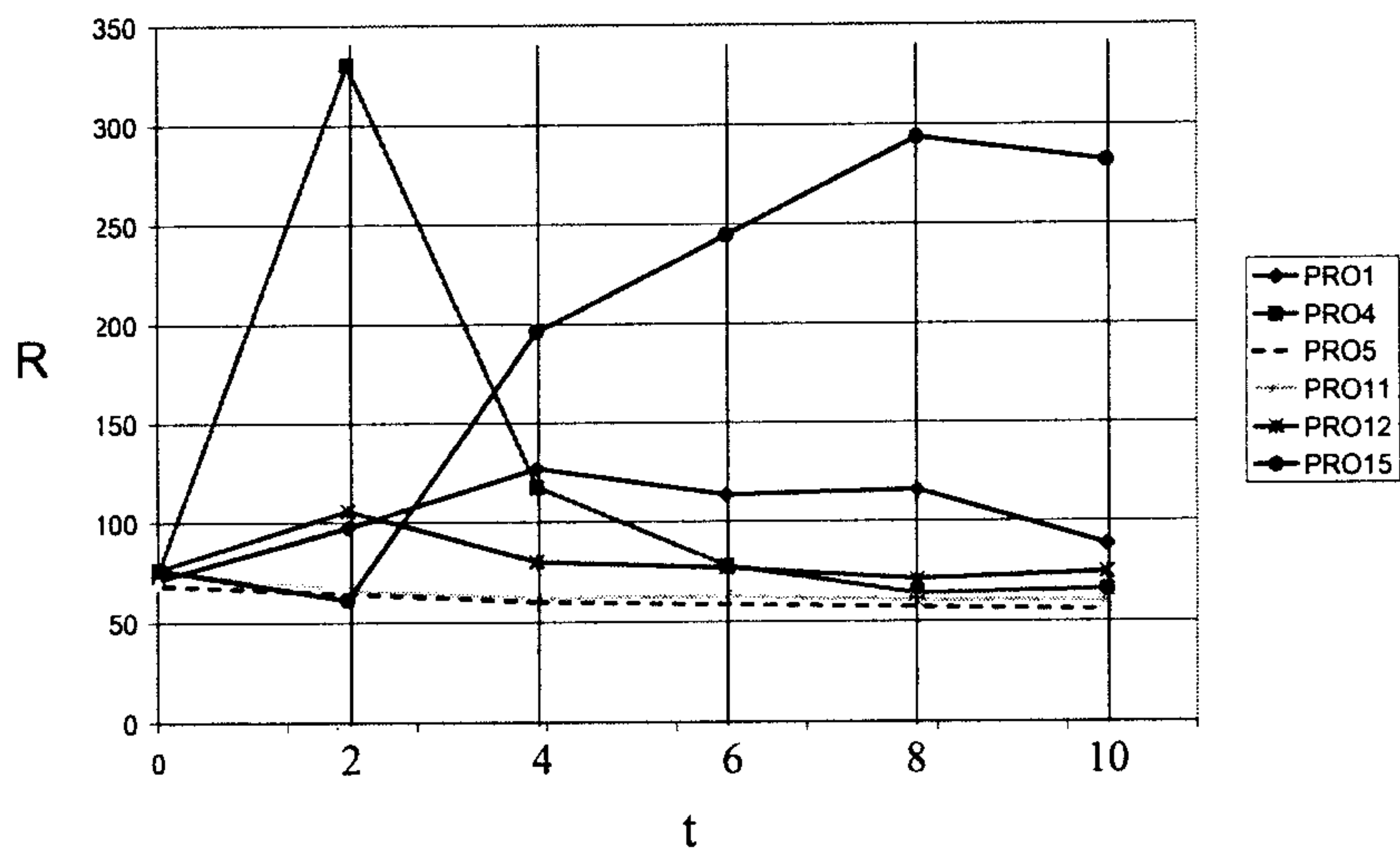


Figure 6b)

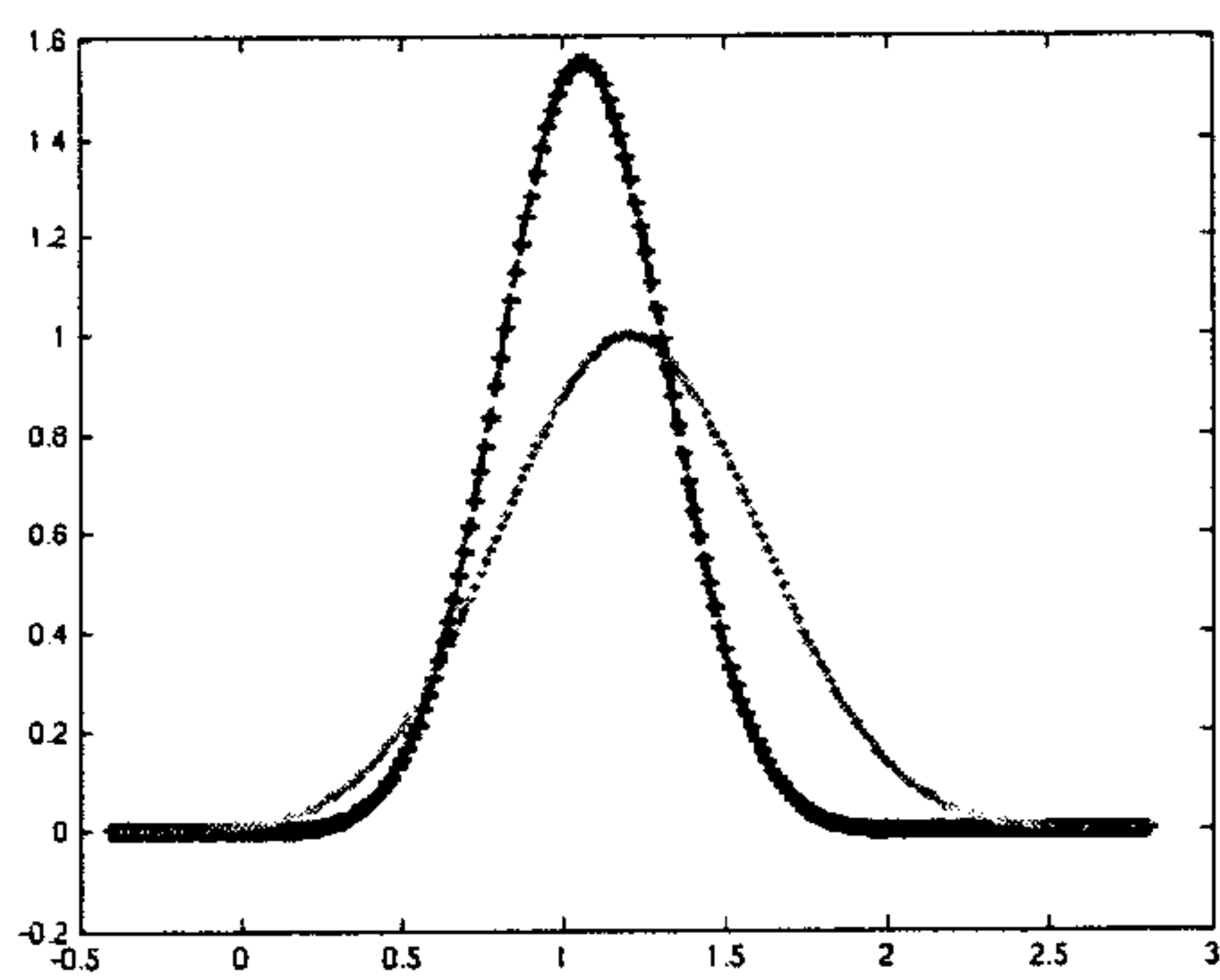


Figure 7a)

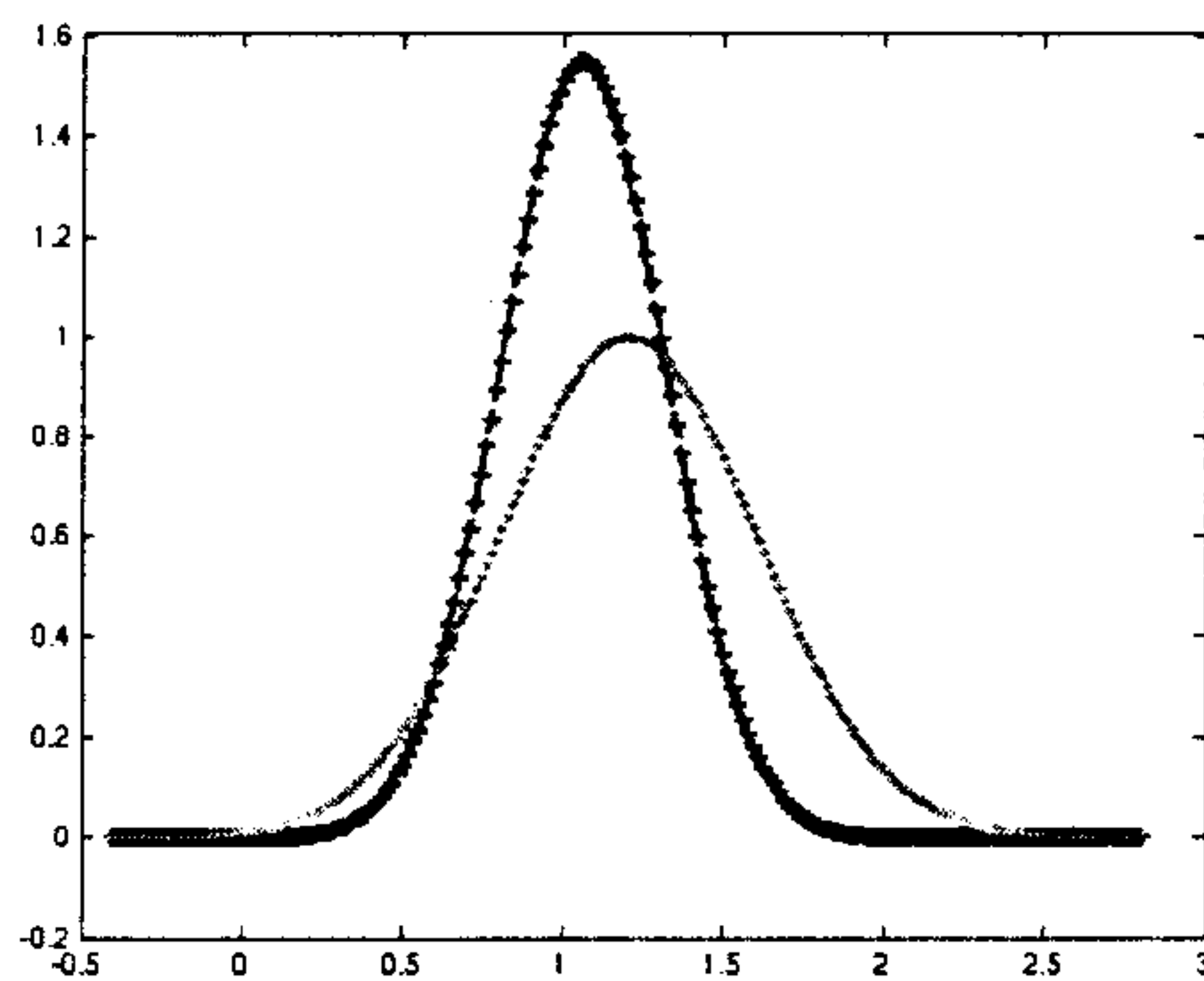


Figure 7b)

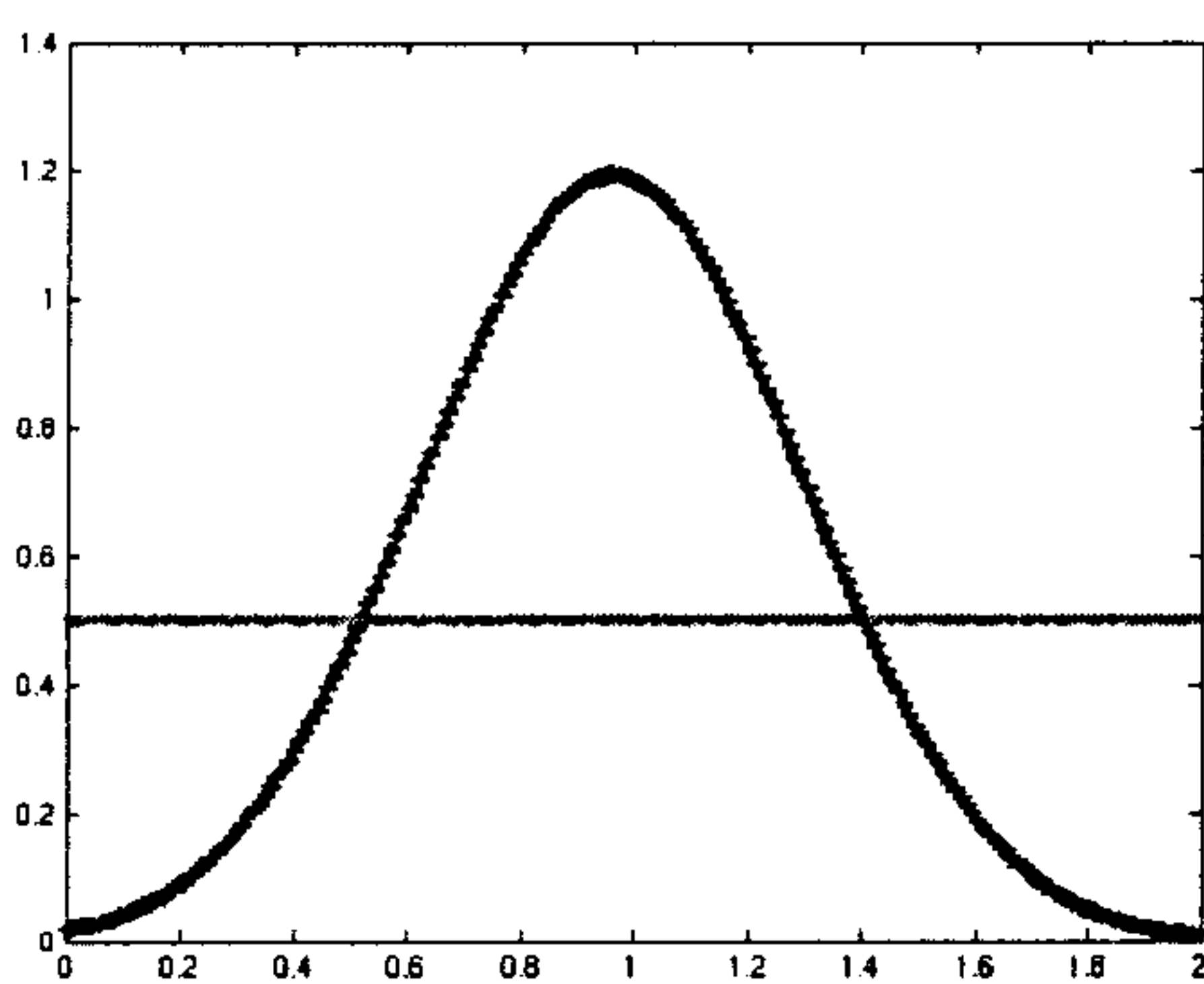


Figure 8a)

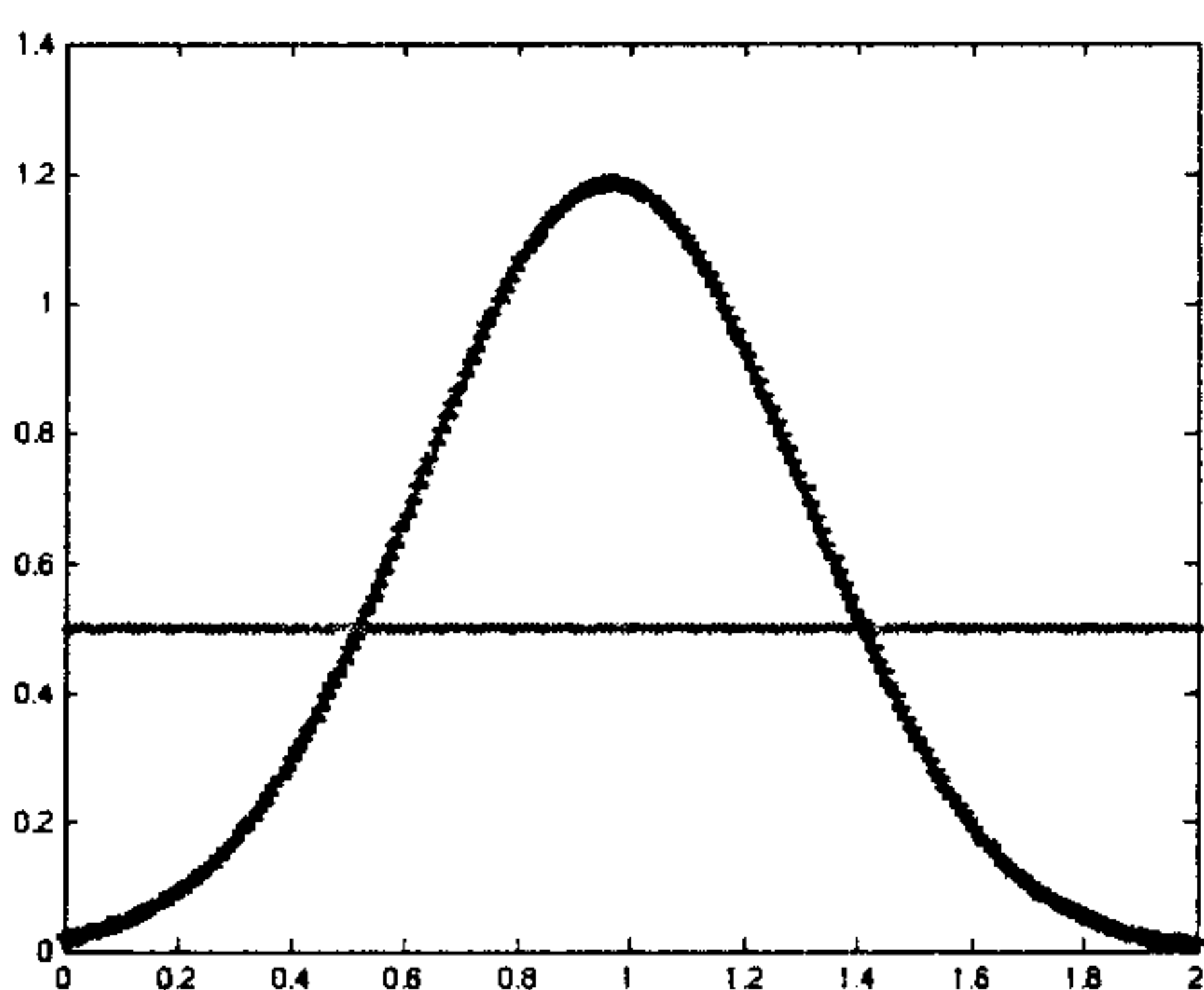


Figure 8b)

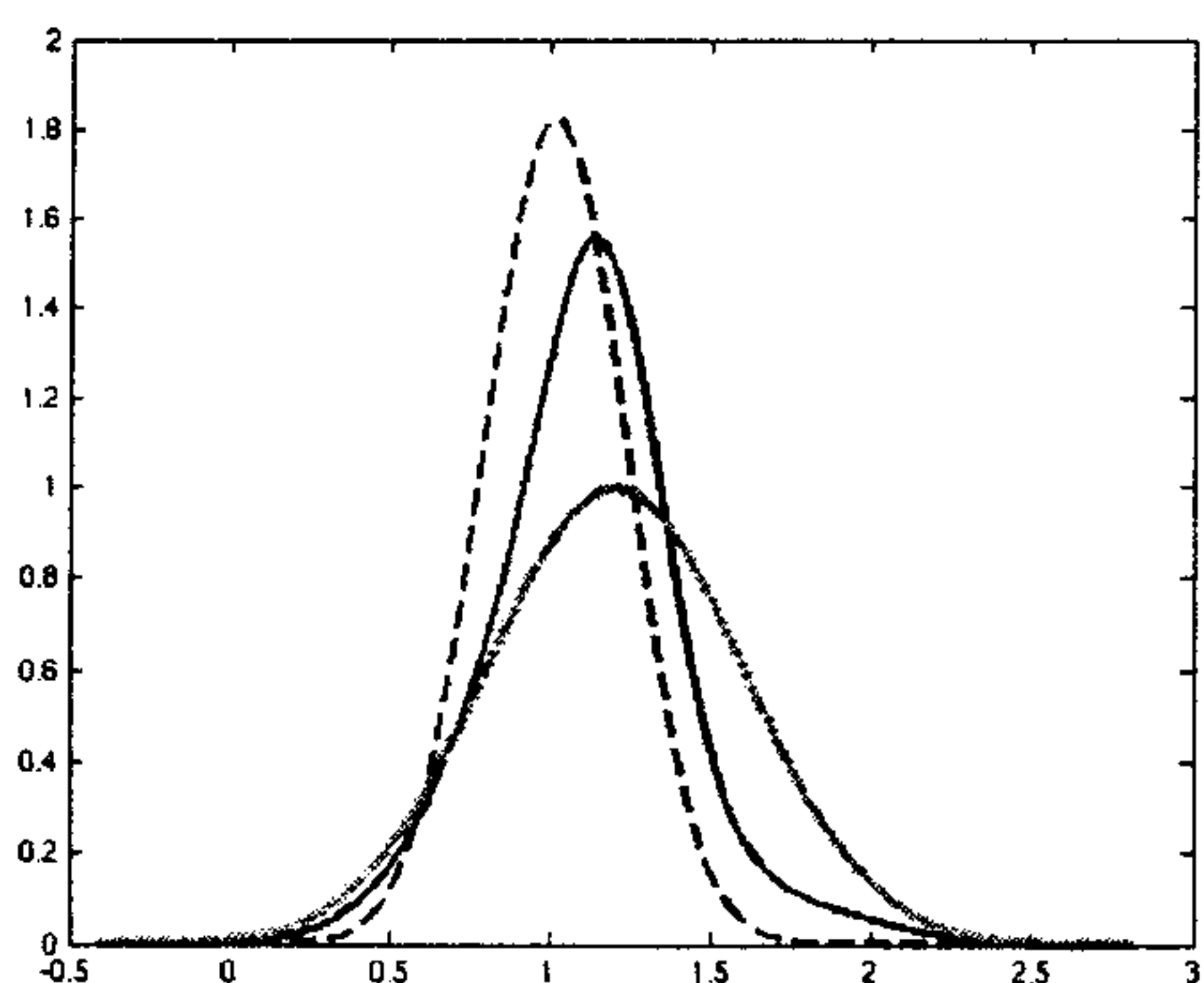


Figure 9a)

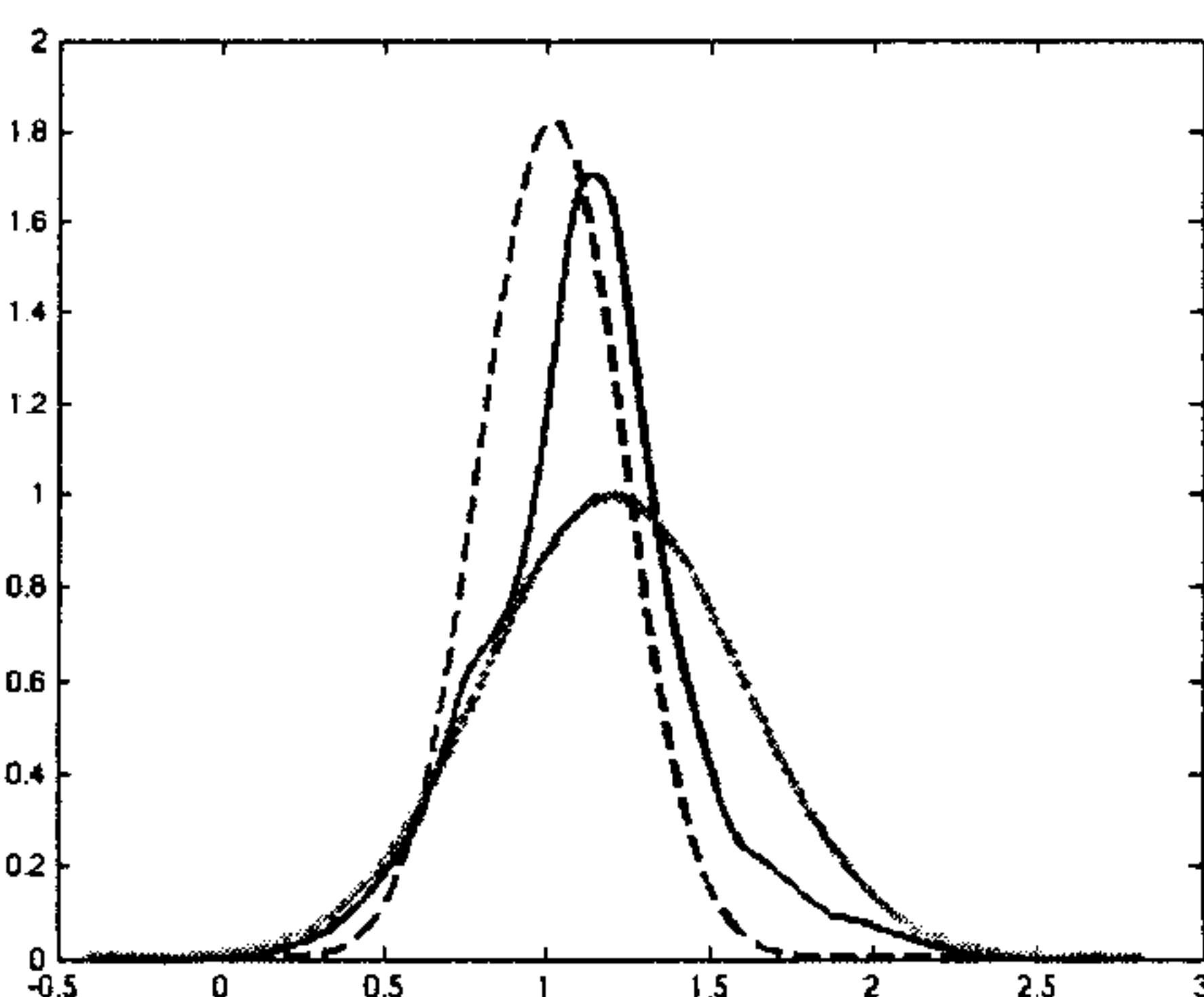


Figure 9b)

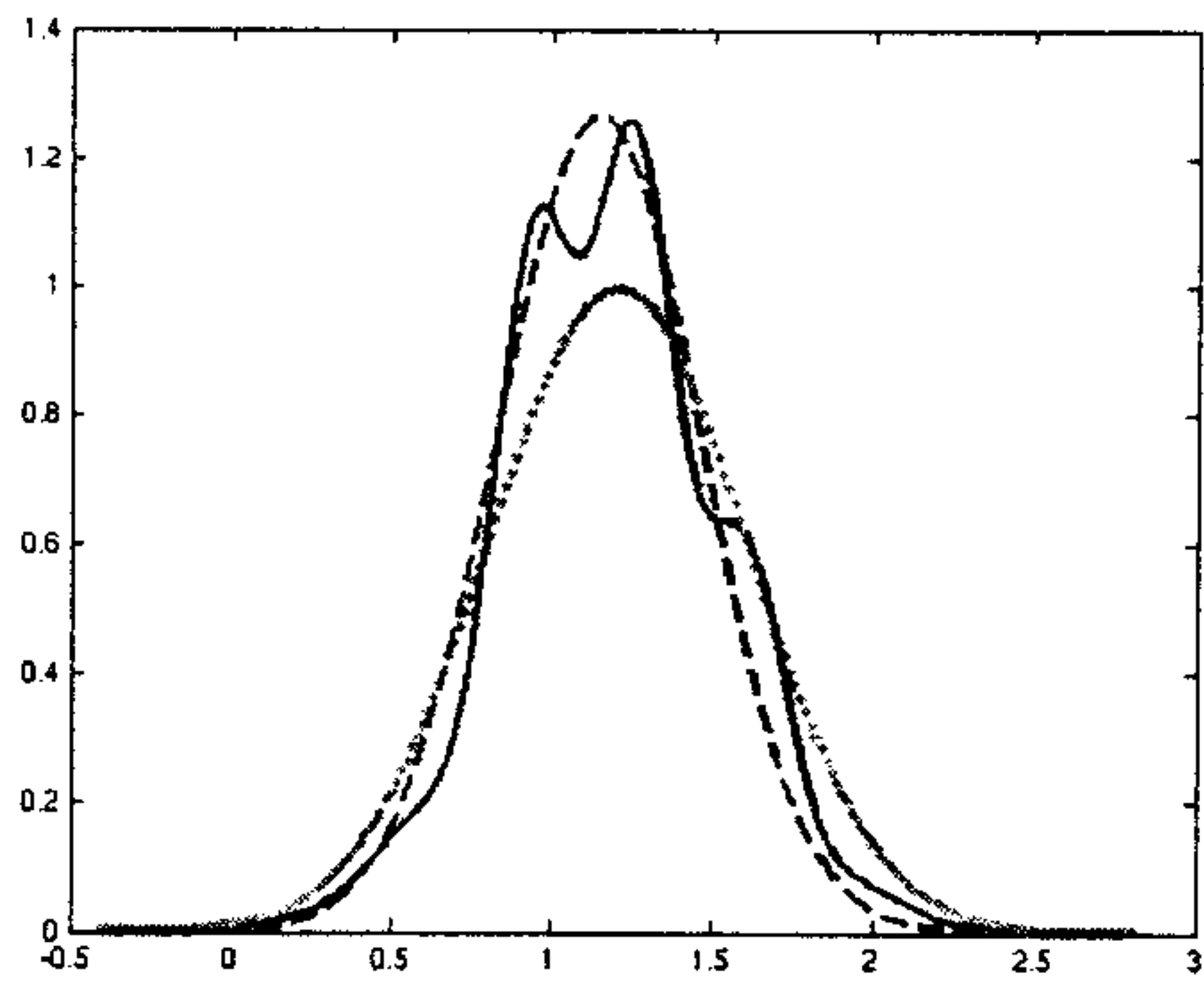


Figure 9c)

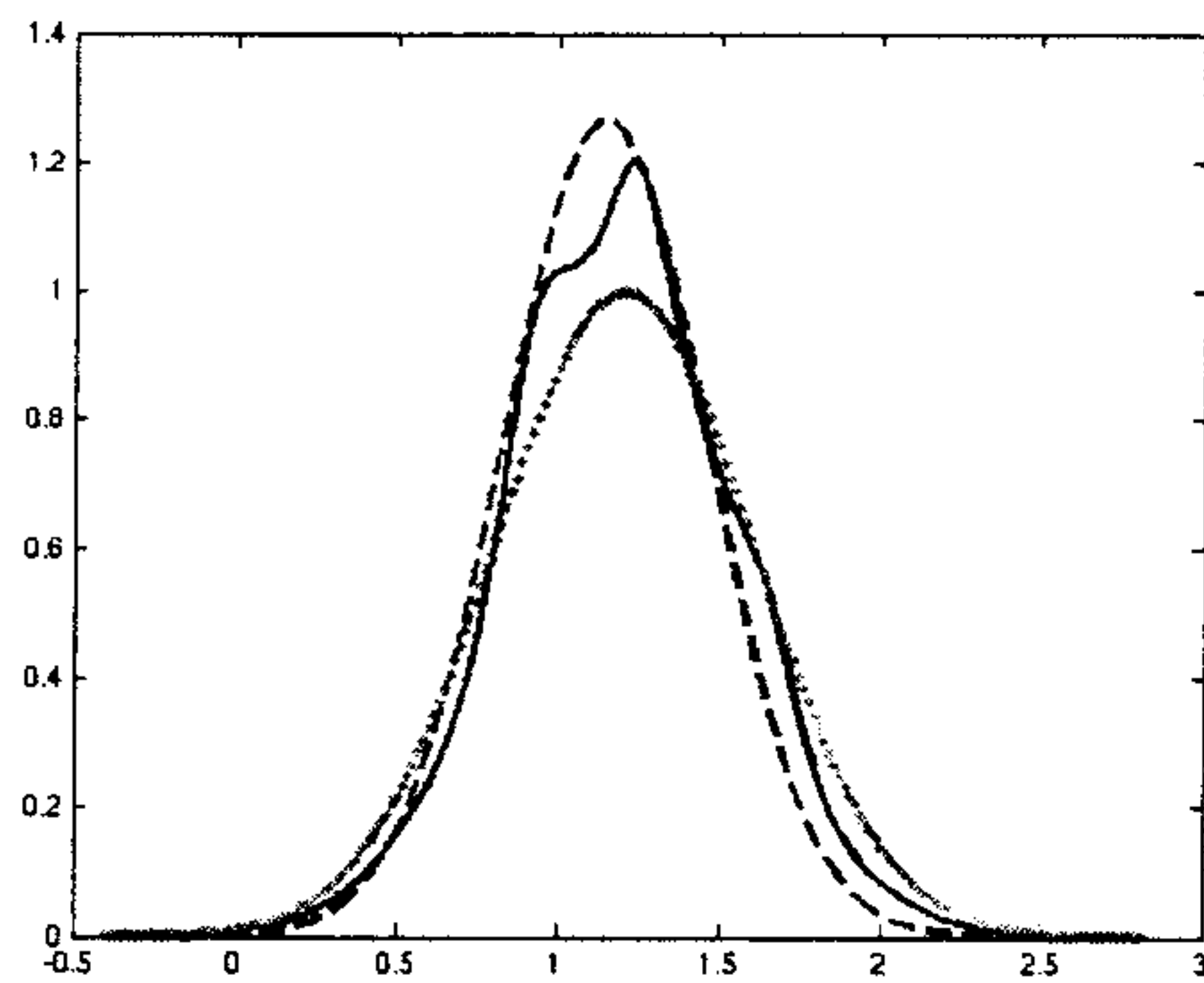


Figure 9d)

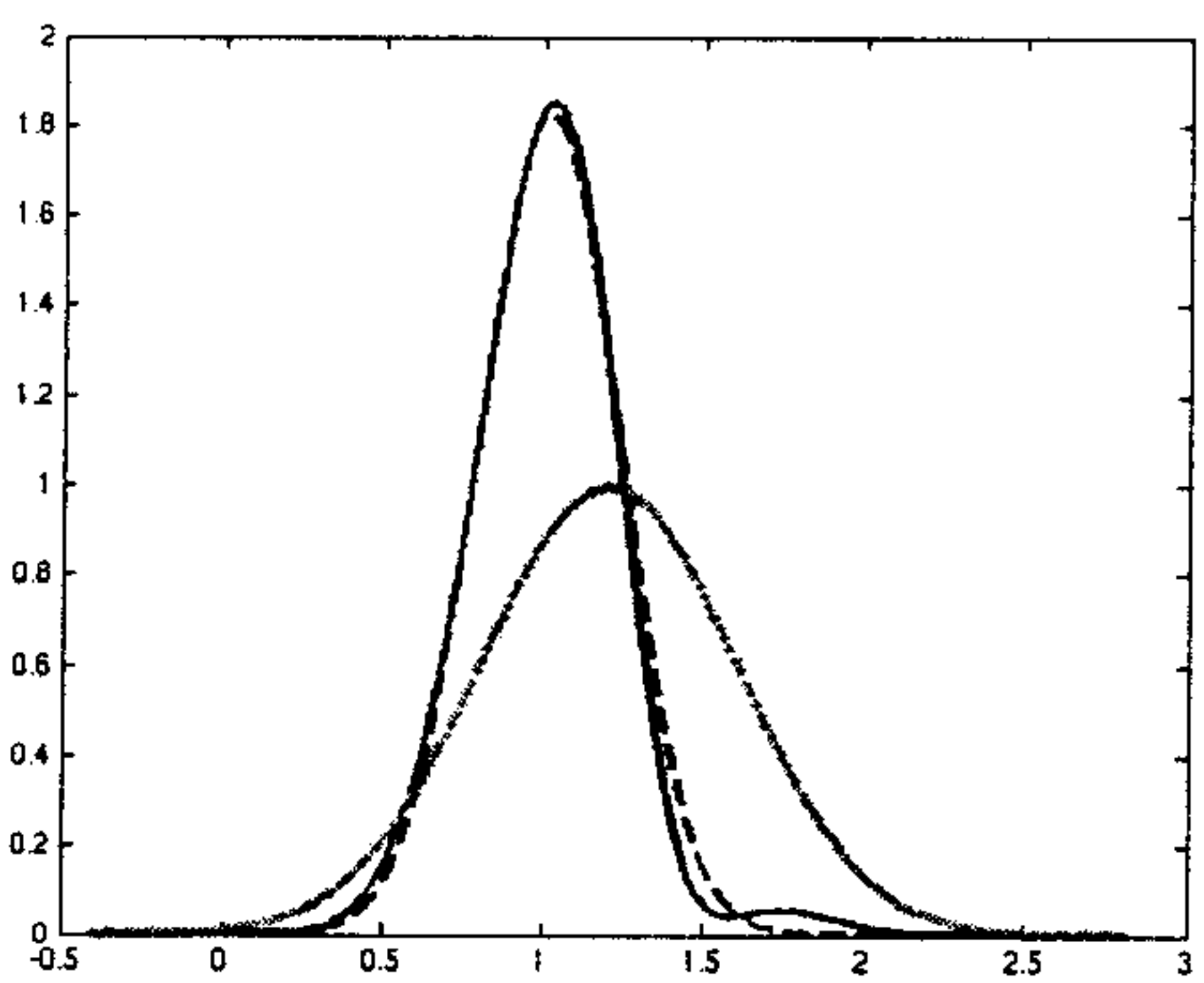


Figure 10a)

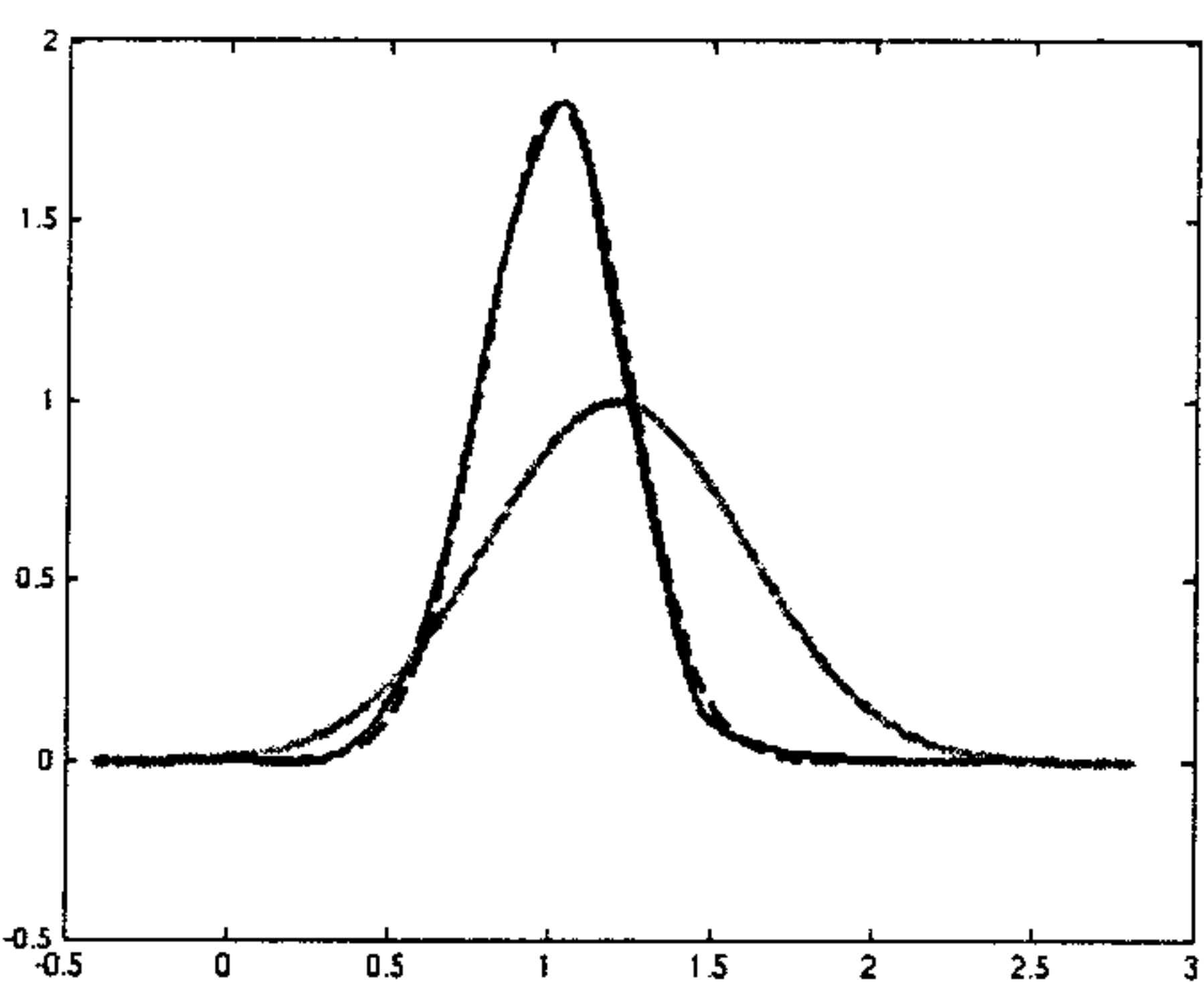


Figure 10b)

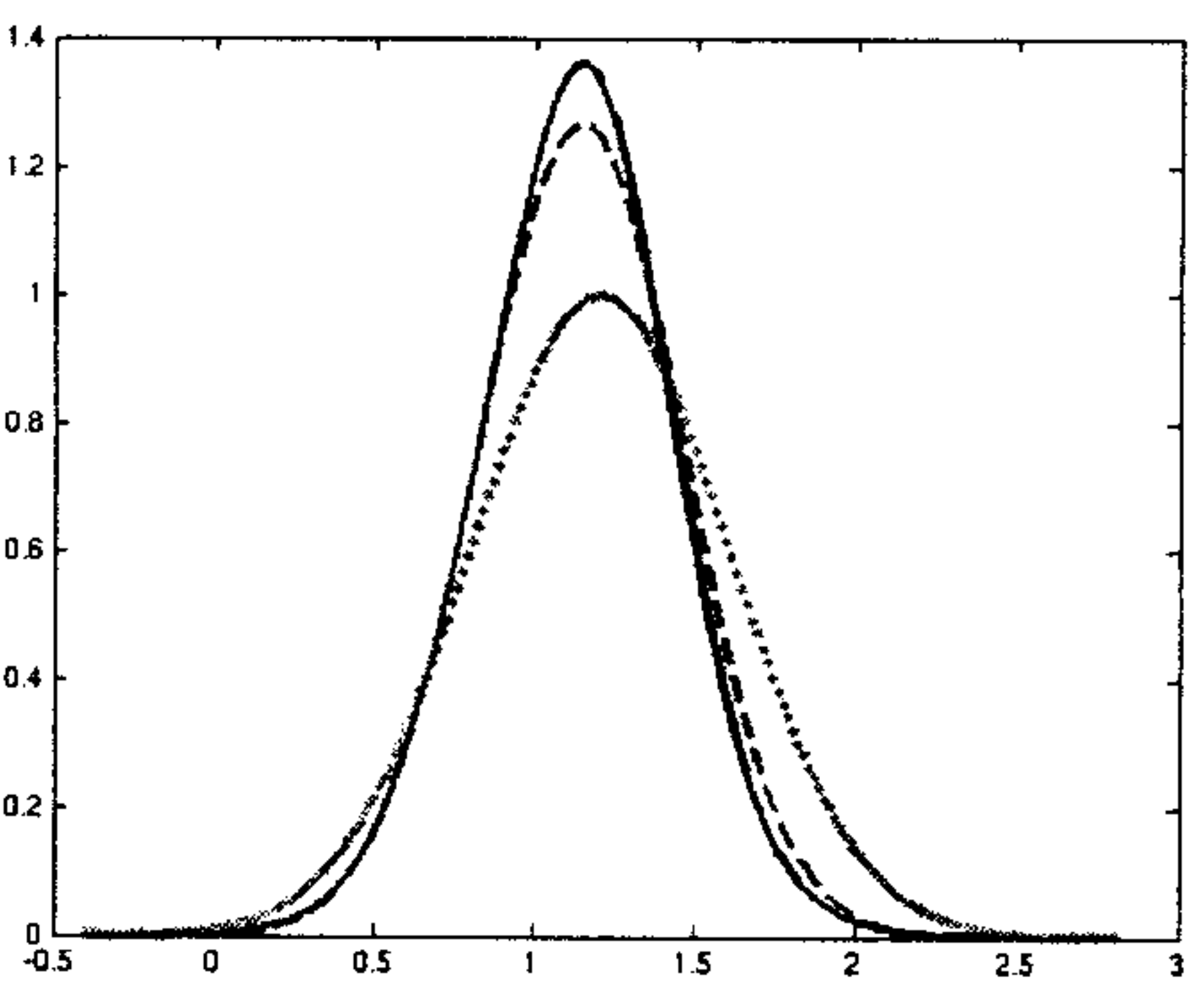


Figure 10c)

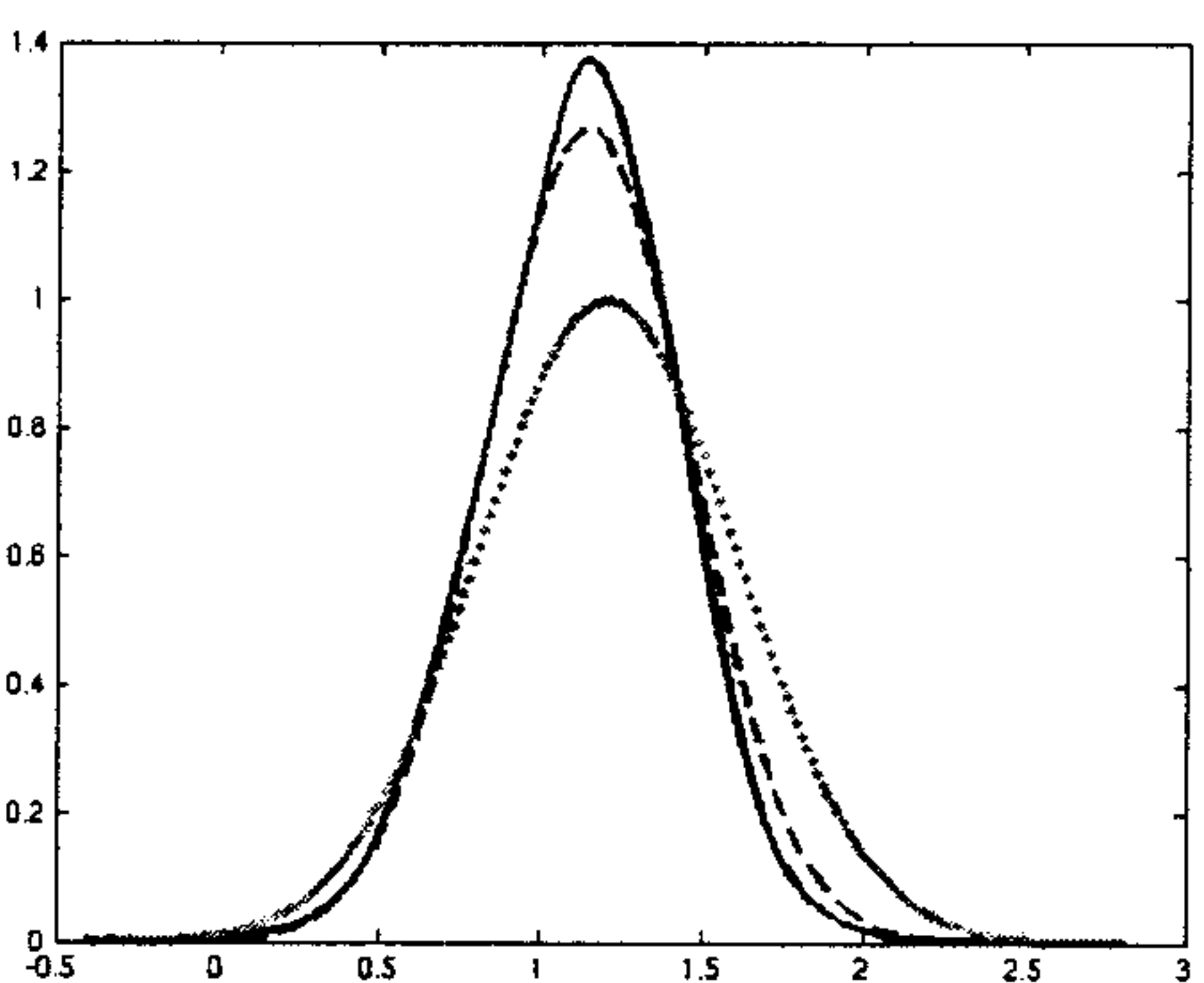


Figure 10d)

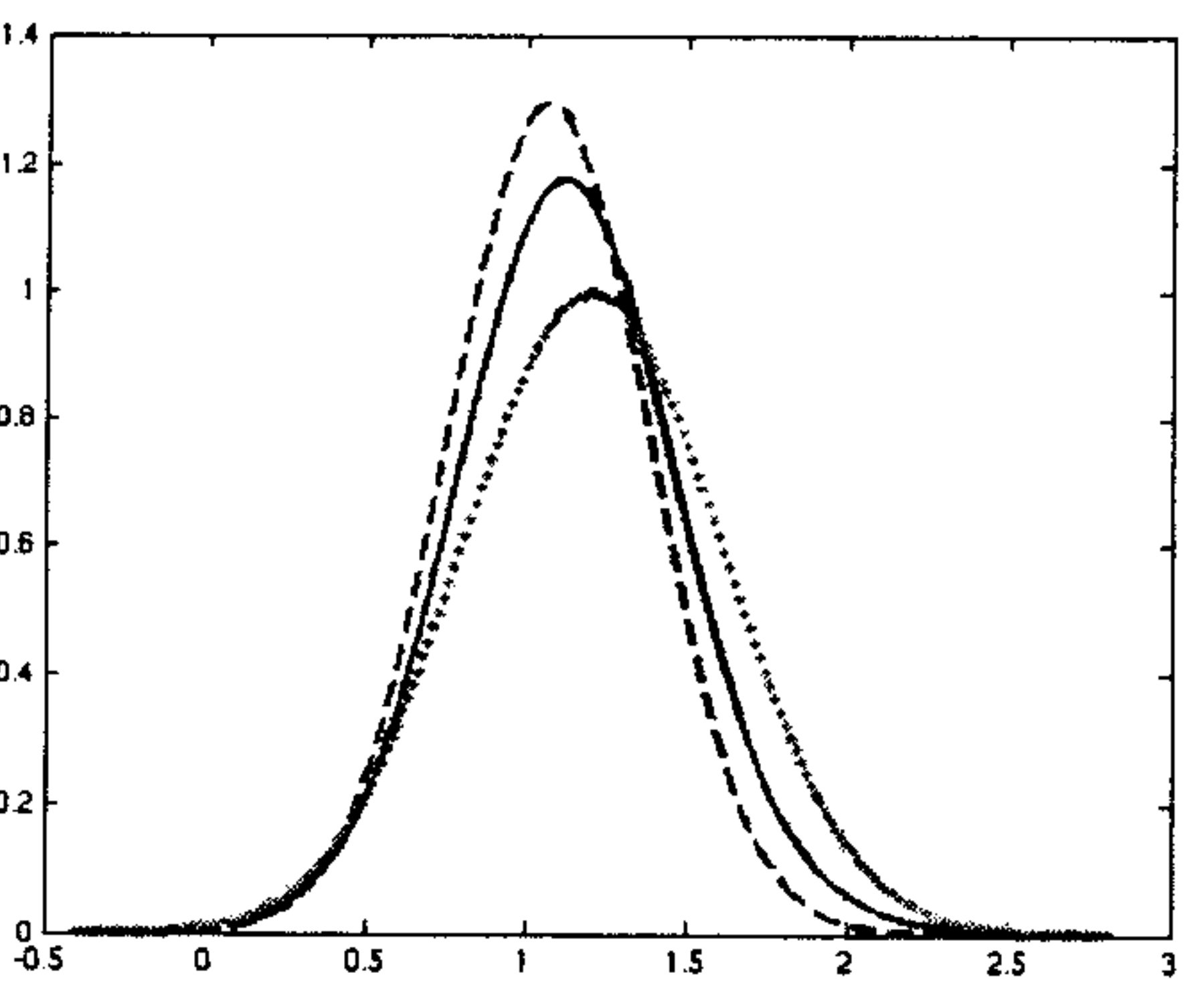


Figure 11a)

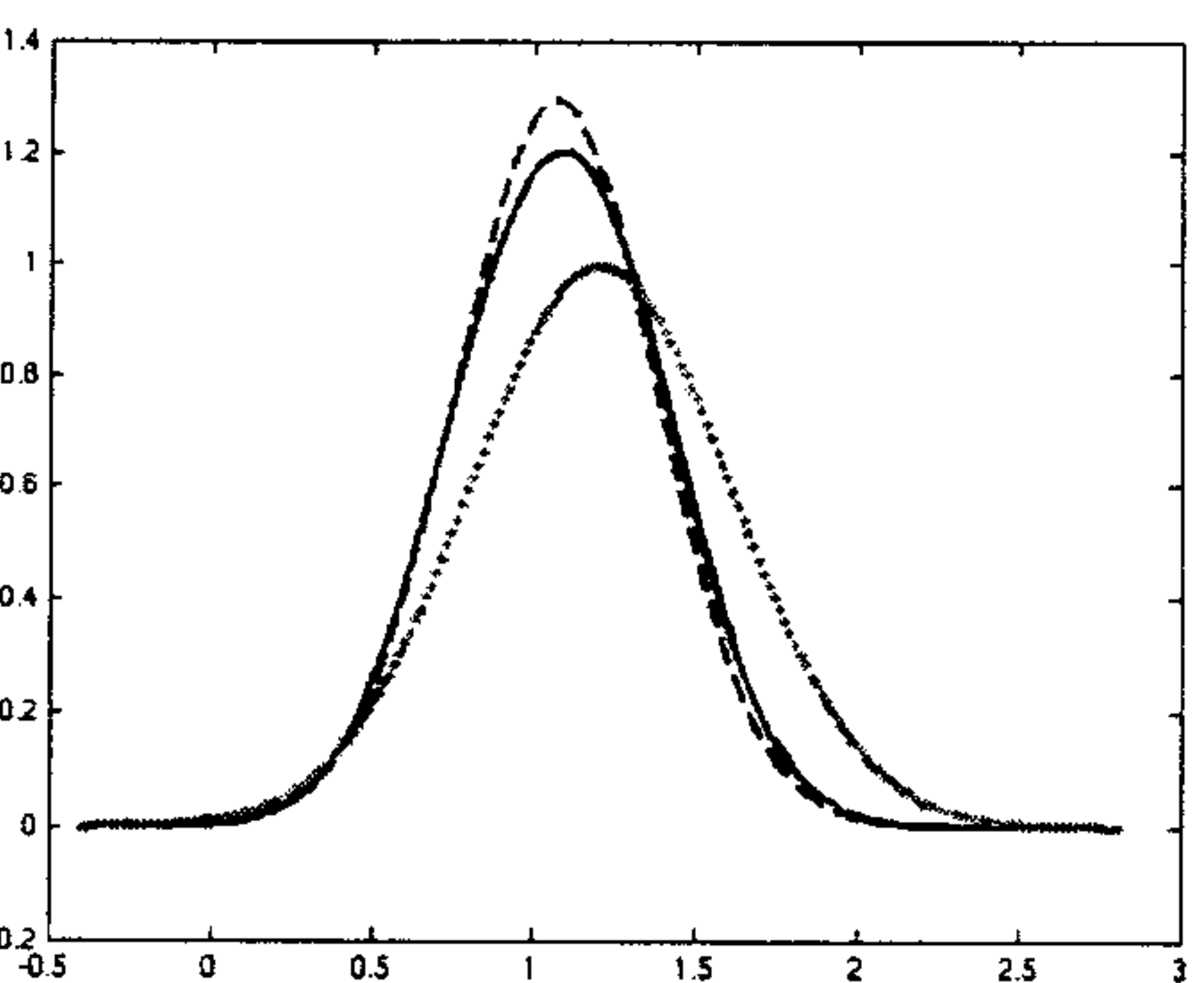


Figure 11b)

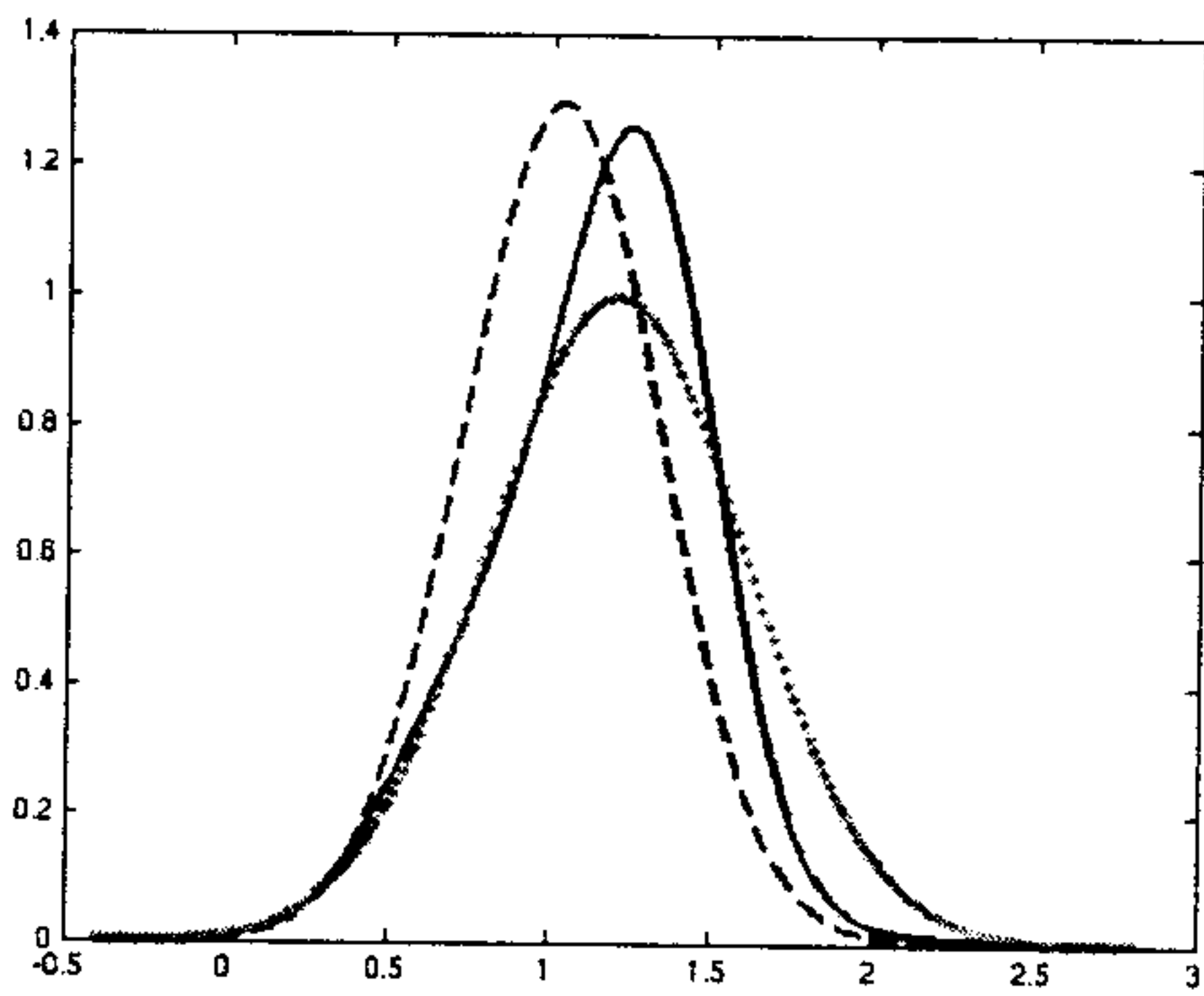


Figure 11c)

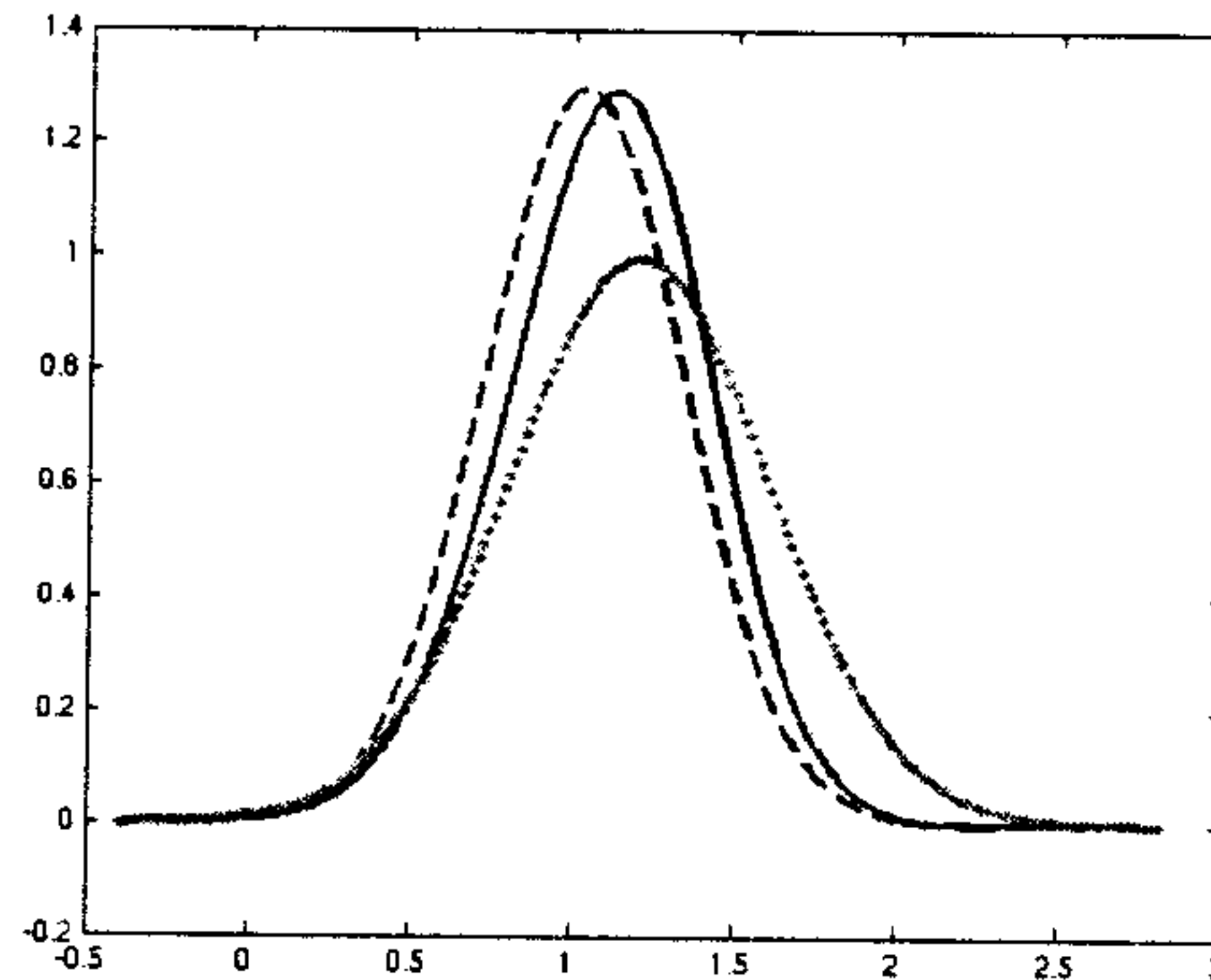


Figure 11d)

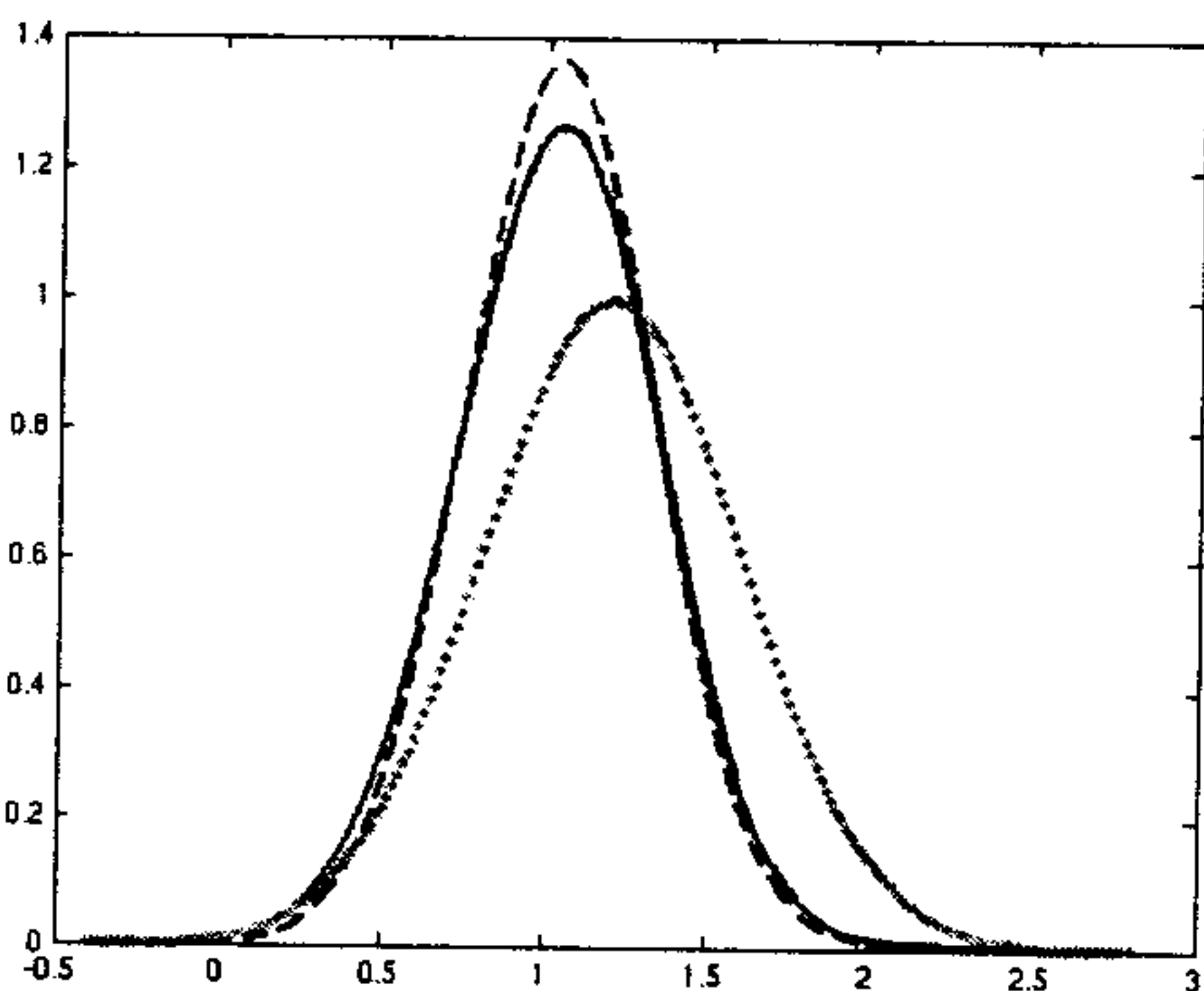


Figure 11e)

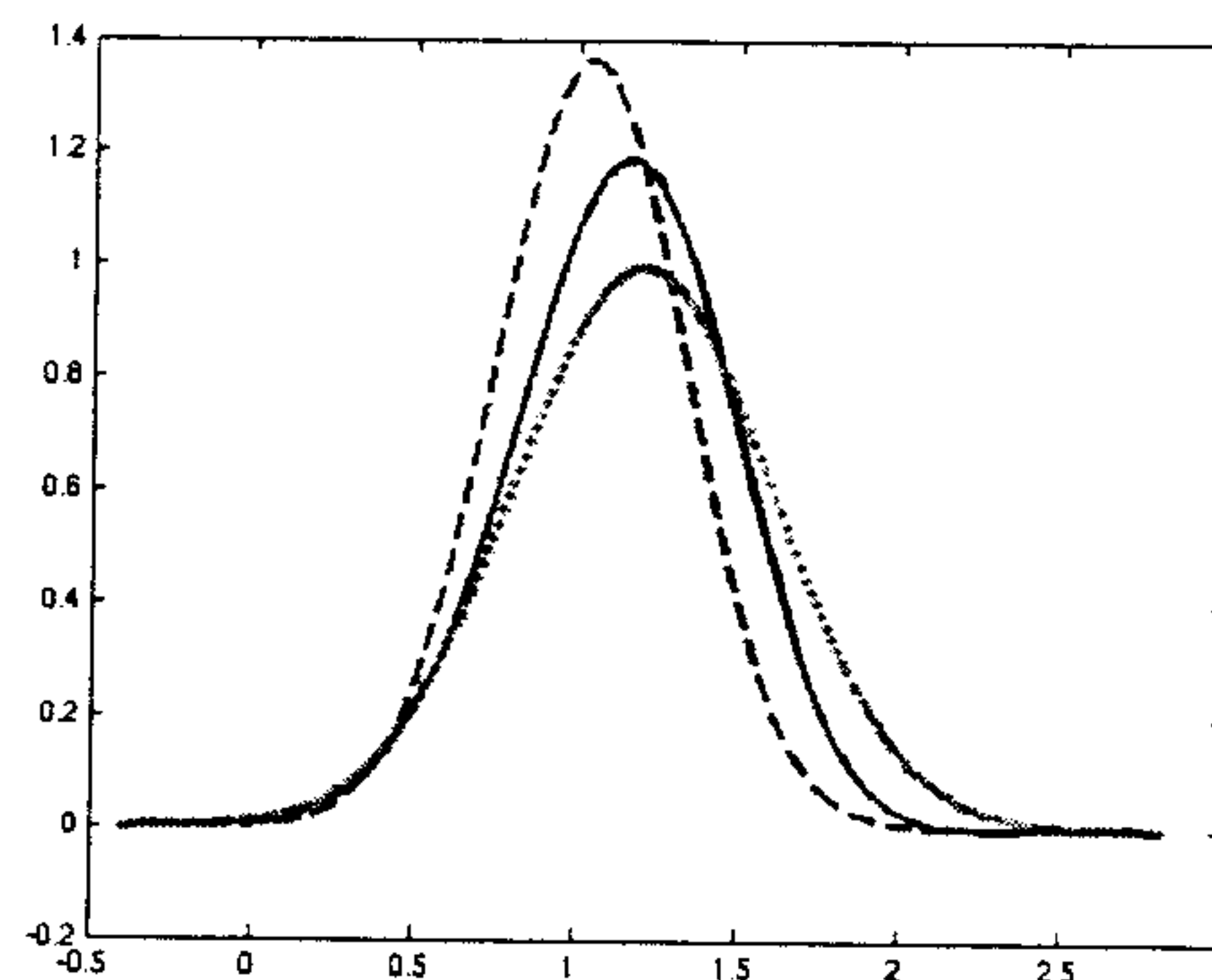


Figure 11f)

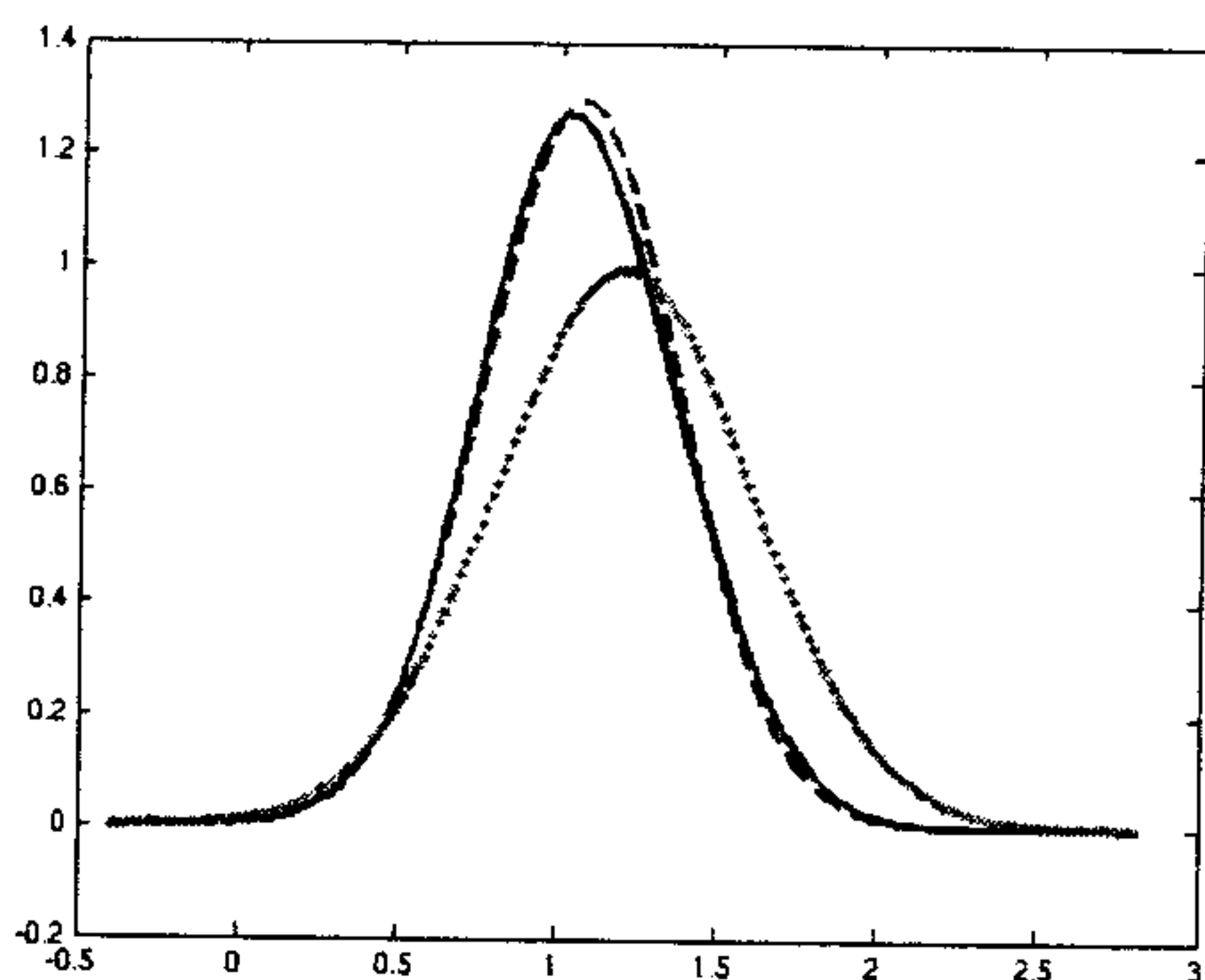


Figure 12a)

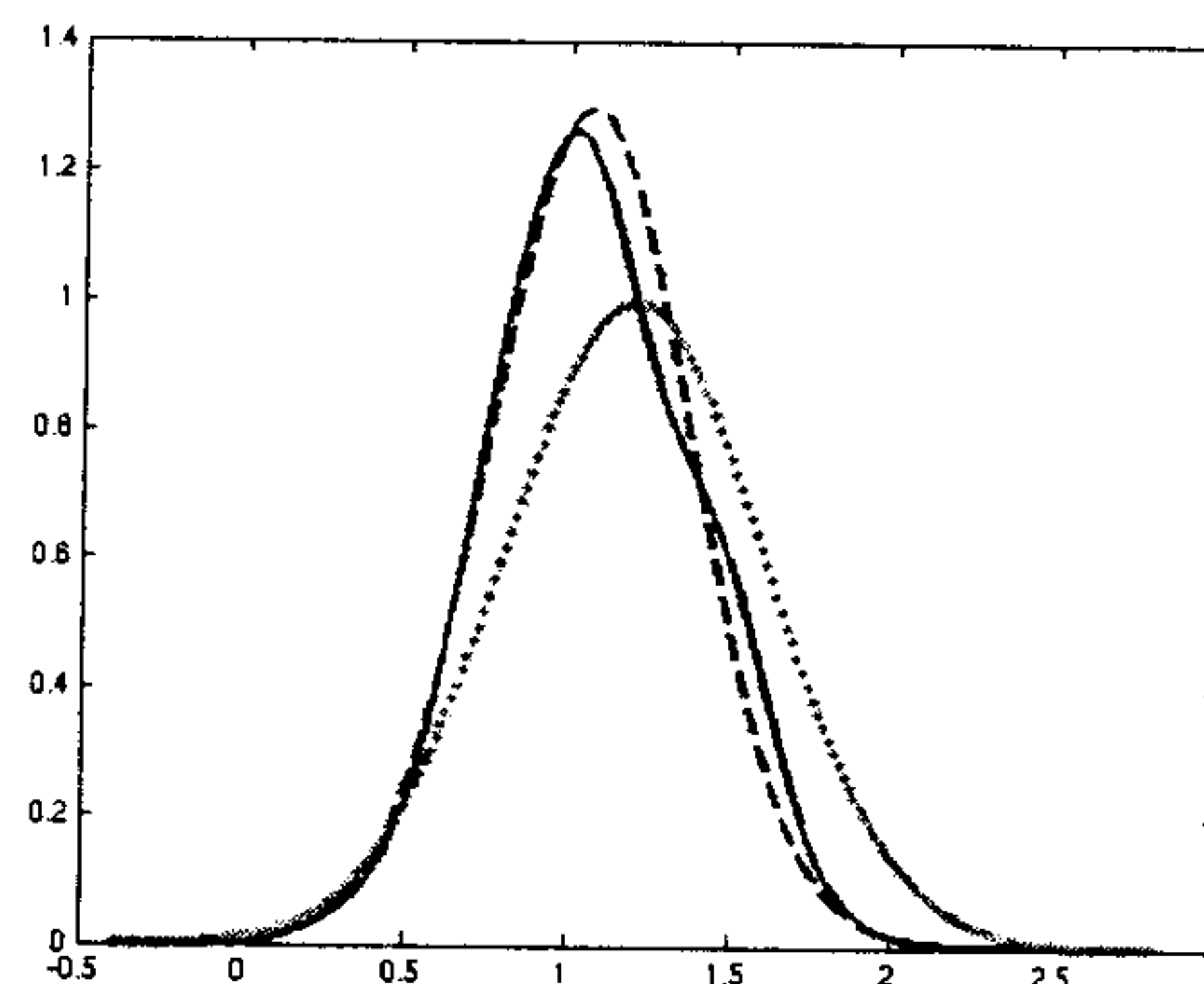


Figure 12b)

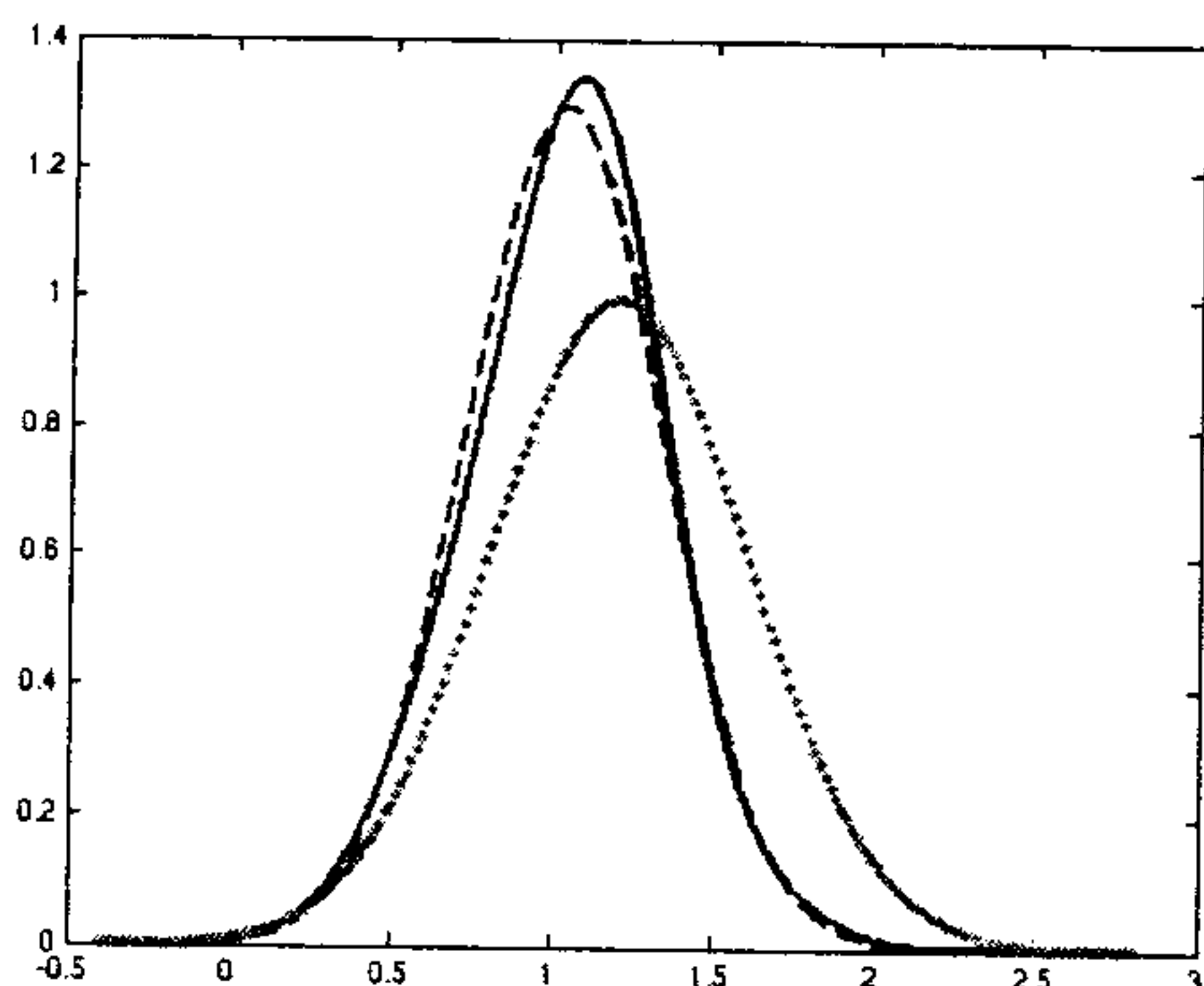


Figure 12c)

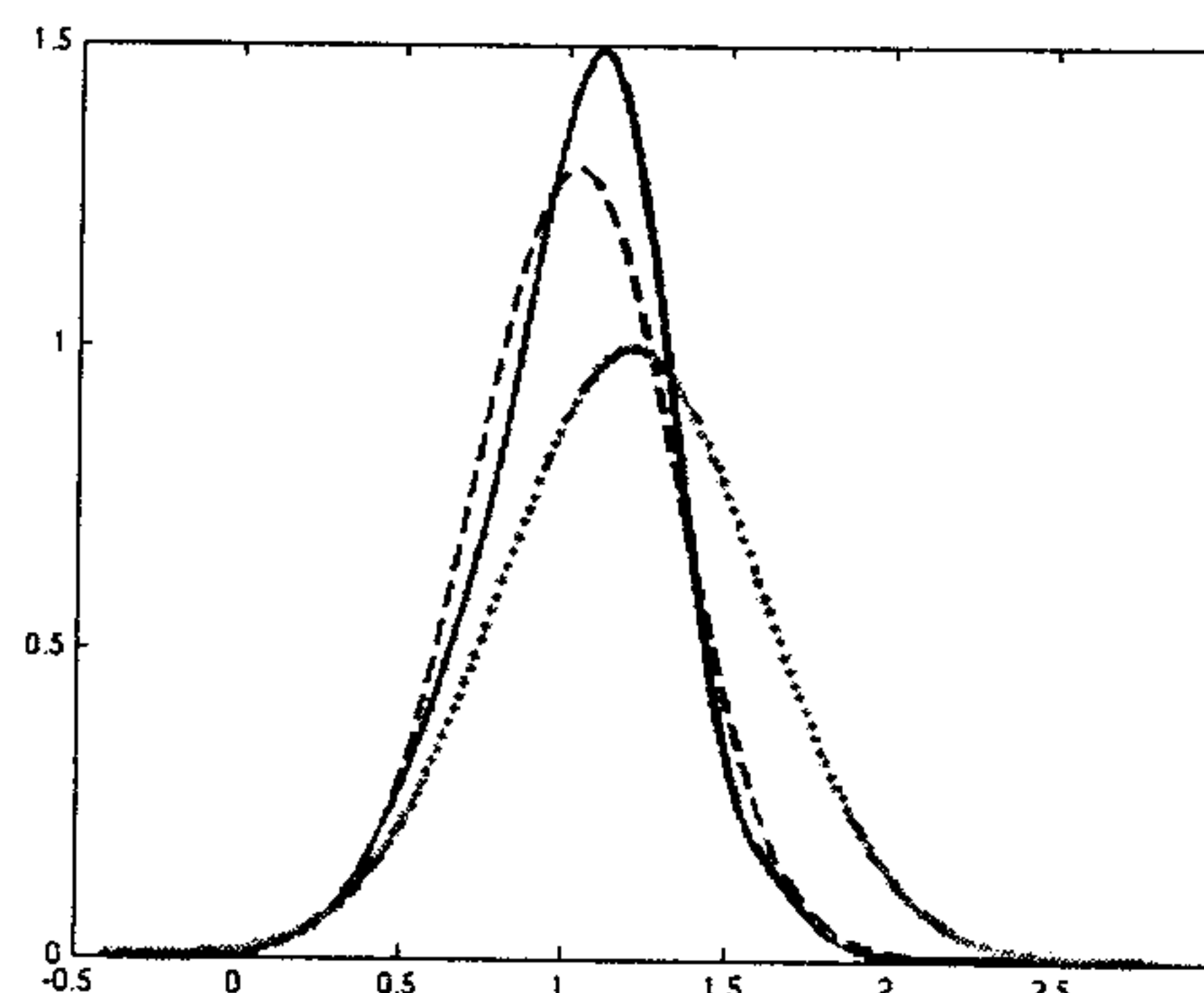


Figure 12d)

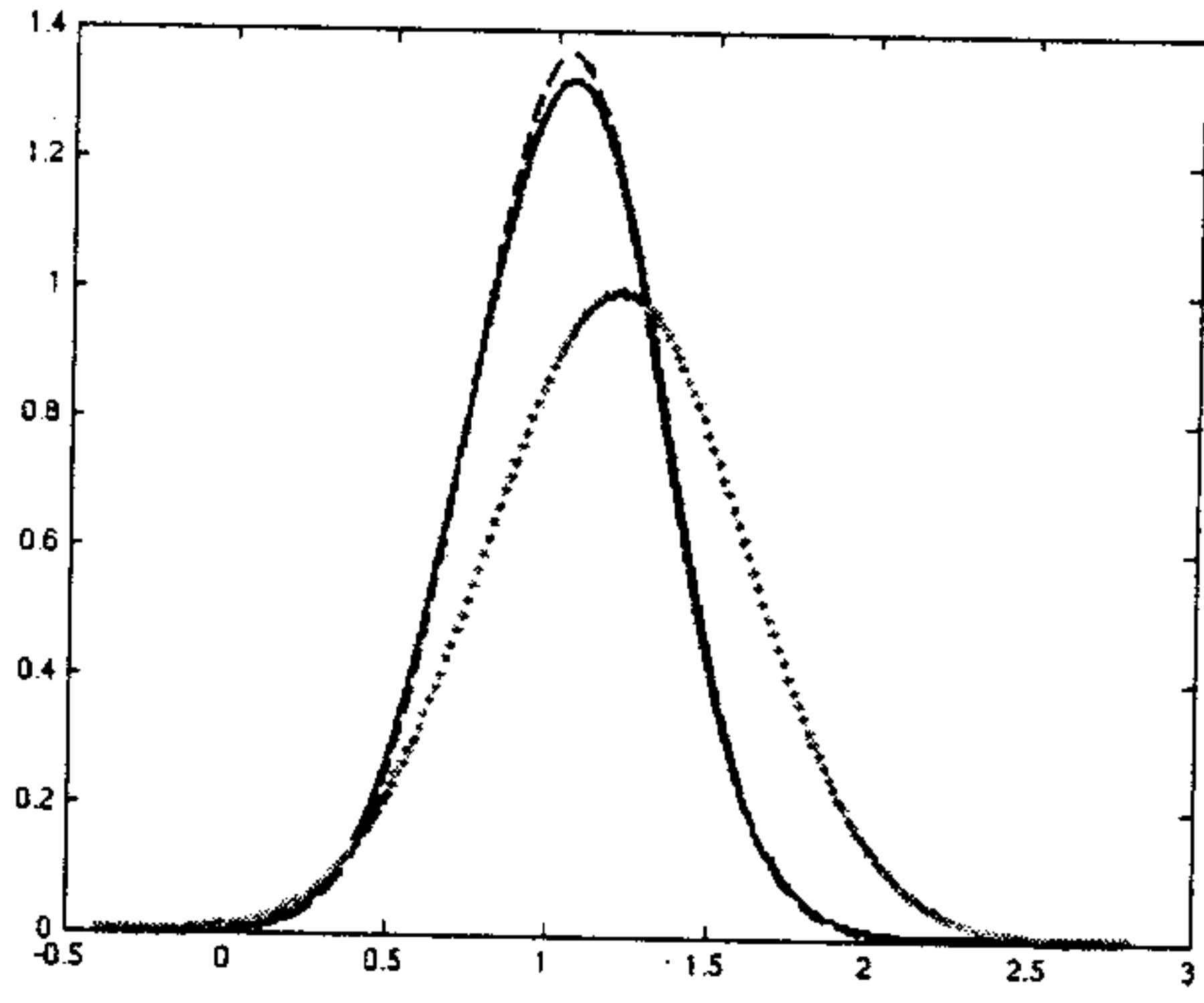


Figure 12e)

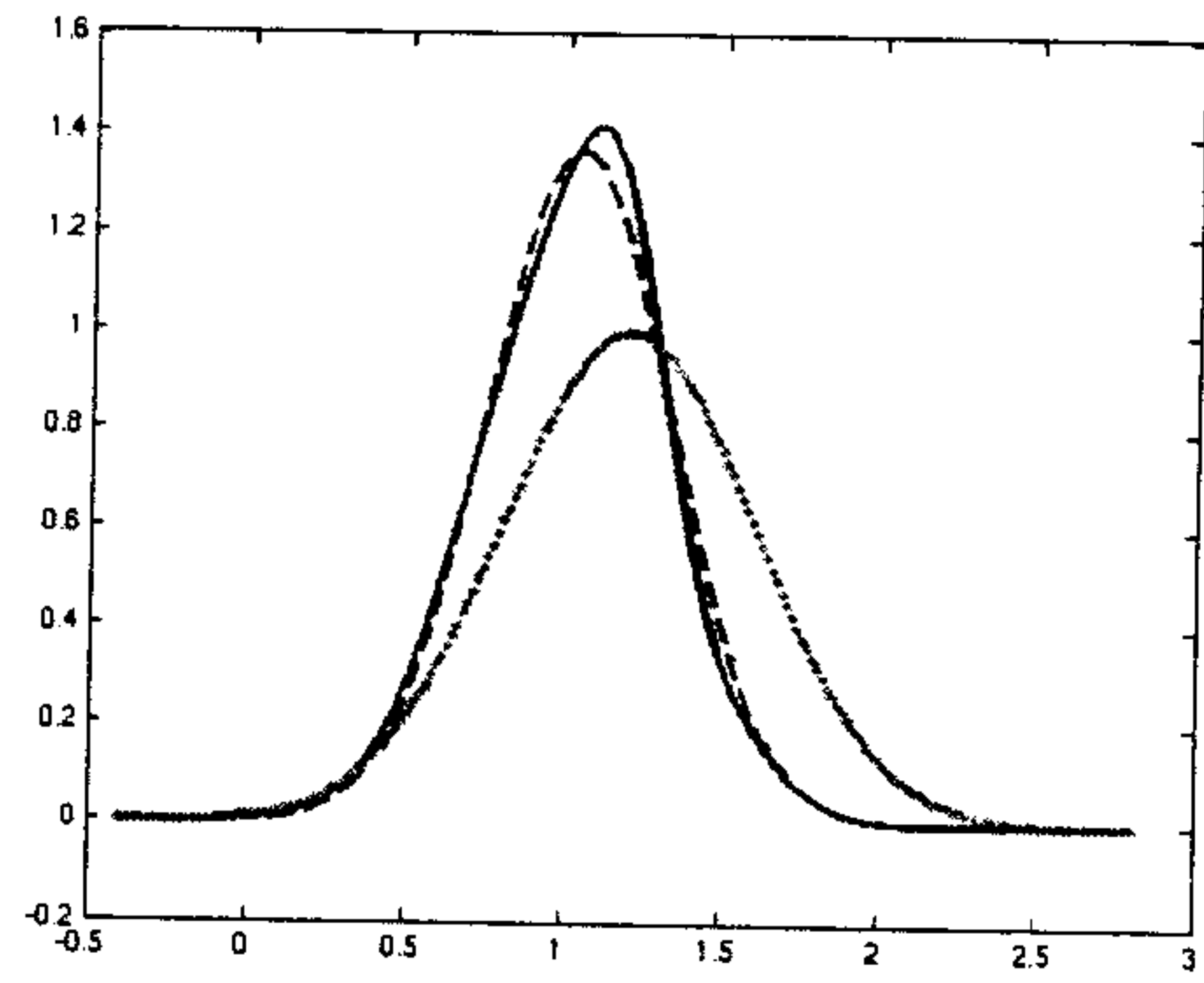


Figure 12f)

