$$\mathbf{k} \triangleq ({}^1\mathbf{k}_0^T, {}^1\mathbf{k}_0^T)^T$$

$\boldsymbol{\tau}_{th}$ is the generalized thruster force

$\boldsymbol{\delta}$ disturbance vector

As discussed in the previous sections, the generalized velocity vector $\boldsymbol{\nu}$ defined in equation (3.15) has to be replaced with the relative velocity vector $\boldsymbol{\nu}_{rel} = ((\mathbf{v}_{u/o} - \mathbf{v}_{fluid/0})^T, \boldsymbol{\omega}_{1/0}^T)^T$ in order to take into account eventual irrotational uniform constant underwater currents. Currents with time changing velocity give rise to an additional hydrodynamic force sometimes called horizontal buoyancy (see section 3.2.4) that can be modelled and simulated only with the knowledge of the fluids inertial acceleration which is usually unaccessible. The velocity gradient of non uniform currents may cause a pressure gradient on the vehicles hull that induces another hydrodynamic load. This latter phenomenon is generally unmodeled as it would require a complete knowledge of the current velocity filed.

3.4 Underwater Manipulator Model

The model of an underwater manipulator can be deduced on the basis of the standard model of a land industrial manipulator and the hydrodynamic forces acting on an underwater rigid body described in the previous sections. Both Lagrange and Newton-Euler methods have been adopted in the literature for the synthesis of an underwater arm model. Schjølberg et al.[22] have derived a lumped parameter dynamic model of an underwater manipulator-vehicle system by an iterative Newton-Euler method. The proposed model is an extension of the classical land manipulator model as outlined by Spong et al.[47] to the underwater environment. The resulting dynamic equation has the same structure of land manipulators, i.e.

$$M(\mathbf{q})\ddot{\mathbf{q}} + C(\mathbf{q}, \dot{\mathbf{q}})\dot{\mathbf{q}} + D(\mathbf{q}, \dot{\mathbf{q}})\dot{\mathbf{q}} + \mathbf{g}(\mathbf{q}) = \boldsymbol{\tau}_m \tag{3.53}$$

being $\mathbf{q}$ the generalized link coordinate vector, M the sum of the standard inertia matrix with the added mass one, C the centripetal-Coriolis matrix including the added mass terms responsible of the hydrodynamic coupling (D'Alembert paradox) discussed previously for a single rigid body, D the hydrodynamic lift and drag generalized forces, $\mathbf{g}$ the weight and buoyancy generalized forces and $\boldsymbol{\tau}_m$ the applied joint generalized forces. The same model structure has been derived with the use of Kane's equations by Tarn et al. both for a single n-axis manipulator-vehicle system [48] and for multiple manipulators-vehicle system [24]. A manipulator-vehicle dynamic model may be use either for simulation purposes, or for control system design.

As far as simulation is concerned McMillan et al.[23] have described an efficient simulator based on an articulated-body algorithm taking into account the major hydrodynamic forces on a manipulator-vehicle system. They show that the computational requirement for a mobile six degrees of freedom underwater manipulator is about double then for a land system, although the amount of computation still grows linearly with the degrees of freedom. In their work drag is modeled as a distributed effect on each link which is approximated with a cylinder. The same kind of drag model is adopted in [49] where a fixed base partially (or totally) immersed manipulator is considered.

Before taking into account the control system design of an underwater-vehicle system it is necessary to understand the actual relevance of hydrodynamic effects on an underwater industrial manipulator and on the overall arm-vehicle system. To this extent most interesting is the experience of the Deep Submergence Laboratory of the Woods Hole Oceanographic Institution as reported in [50]: as the major hydrodynamic effect on a fixed base manipulator is damping, the authors conclude that on their system *"while performing routine tasks hydrodynamic effects have no significant effect on manipulator control"*. On the other hand for a mobile base system the effects of fast manipulator motion on the vehicle have experimentally shown to be relevant. Intuitively this is reasonable as it corresponds to a "swimming-like" coupling effect between the manipulator and the vehicle. This phenomenon has been extensively analyzed by McLain et al.[51] [52] [53] [25] both theoretically and experimentally. Their approach basically consists of two steps: first a model arm, developed on the basis of Sarpkaya's study of the added mass and drag coefficients on a cylinder [54], is experimentally identified. Then the dynamic model of the arm is used to compensate the arm-vehicle coupling effect by a model-based feedforward signal. Experimental tests carried out with a single link arm on an unmanned underwater vehicle [53] show good improvements in the control performance with only a small (5%) increase in vehicle thrust. Nevertheless the implementation of such approach on a real system would require an accurate and complete identification of the underwater arm model and a much higher computational burden with all the drawbacks that this implies. In the words of Sayers et al. [50] "*when operating at normal speeds in the real environment the principal benefit resulting from a full model for combined manipulator/vehicle motion is not the ability to compensate for dynamic effects on-line, but rather the assistance it provides in planning alternative motions off-line*". As a matter of fact to avoid a complete off-line identification procedure an adaptive approach may be considered. Simulation results of an adaptive scheme for underwater manipulator-vehicle control have been reported by Mahesh et al.[21] and adaptive algorithms for underwater manipulators have been analyzed also by Ramadorai et al. [12]. Assuming a full knowledge of the system parameters a feedback linearization control of the vehicle manipulator system may be implemented as suggested Schjølberg et al.[55]. Finally robust control approaches to the problem have been taken into account in [27].

Chapter 4
Identification

In this chapter the topic of underwater robotic system identification will be addressed. Experimental results regarding open frame ROVs will be outlined.

4.1 Estimation approach

In order to describe the adopted estimation approach some notation will be introduced. Suppose that a model of some phenomena is given in the form

$$\mathbf{y_0}(t) = H(\mathbf{x}(t), t, \boldsymbol{\theta}) \tag{4.1}$$

being t the time, $\mathbf{y_0}(t) \in \Re^{n\times 1}$ a measurable quantity (either deterministic or stochastic), $\mathbf{x}(t) \in \Re^{m\times 1}$ a time dependent variable and $\boldsymbol{\theta} \in \Re^{p\times 1}$ a vector of parameters: the problem of parameter estimation consists in calculating some "estimate" $\hat{\boldsymbol{\theta}}$ (estimates will be denoted by a hat) of the parameter vector $\boldsymbol{\theta}$ given noisy measures $\mathbf{y}(t) = \mathbf{y_0}(t) + \boldsymbol{\varepsilon}$ (being $\boldsymbol{\varepsilon}$ the measurement noise). Notice that due to the unavoidable measurement noise the measures $\mathbf{y}(t)$ of $\mathbf{y_0}(t)$ are random variables being H either a stochastic or a deterministic model. The approach to the estimation problem is different according to the nature of the parameter vector $\boldsymbol{\theta}$: if it is a vector of unknown constants a so called *non-Bayesian* approach should be followed, otherwise a *Bayesian* one. Within the Bayesian approach the parameter vector is a random variable having a probability density function (pdf) $p(\boldsymbol{\theta})$ and one may argue that a good estimate of it could be the mode of the pdf of $\boldsymbol{\theta}$ conditioned to the measurements $\mathbf{y}$, i.e.

$$\hat{\boldsymbol{\theta}} \triangleq \arg\max_{\boldsymbol{\theta}} p(\boldsymbol{\theta}|\mathbf{y}) = \arg\max_{\boldsymbol{\theta}} \frac{p(\mathbf{y}|\boldsymbol{\theta})p(\boldsymbol{\theta})}{p(\mathbf{y})} = \arg\max_{\boldsymbol{\theta}} p(\mathbf{y}|\boldsymbol{\theta})p(\boldsymbol{\theta})$$

Within the non-Bayesian approach the parameter vector is an unknown deterministic quantity which has no probability density function (or better, its pdf is a Dirac function centered on the unknown value $\boldsymbol{\theta}$). A typical non-Bayesian estimator is the maximum likelihood estimator, i.e. the parameter estimate $\hat{\boldsymbol{\theta}}_{MLE}$ is the mode of the measurement probability density function given the deterministic parameter vector,

$$\hat{\boldsymbol{\theta}}_{MLE} \triangleq \arg\max_{\boldsymbol{\theta}} p(\mathbf{y}|\boldsymbol{\theta}) \tag{4.2}$$

The pdf of the measurements $\mathbf{y}$ given $\boldsymbol{\theta}$ is called *likelihood function* $\Lambda_{\mathbf{y}}(\boldsymbol{\theta}) \triangleq p(\mathbf{y}|\boldsymbol{\theta})$ and the estimator described by the above equation is called *maximum likelihood estimator* (MLE). The underwater robotic models developed in the previous chapters are deterministic models that contain only deterministic parameters as masses, drag coeffi-

cients and geometrical quantities that should be regarded as unknown constants rather then random variables. This suggests to work within the non-Bayesian framework and in particular estimation should be performed with a maximum likelihood technique. Notice that according to the above statements regarding the stochastic nature of $\mathbf{y}(t)$ due to the measurement noise $\boldsymbol{\varepsilon}$, in both the Bayesian and non-Bayesian frameworks, the estimate $\hat{\boldsymbol{\theta}}$ of the parameter vector $\boldsymbol{\theta}$ is a random variable either if $\boldsymbol{\theta}$ is a deterministic quantity or a stochastic one. Indeed important informations on the model structure can be deduced analyzing the covariance of $\hat{\boldsymbol{\theta}}$. As a matter of fact the underwater robotic models to identify are linear in the unknown parameters, thus least squares estimation can be applied.

4.1.1 Least Squares Technique

Within this paragraph the least squares technique for estimation of models linear in their parameters will be reviewed in order to introduce the adopted notation and to outline the criteria that have been used for input and model selection. As all the following results are standard in the identification literature many of them will be reported only for reference and without proof.

If the equation (4.1) model happens to be linear in $\boldsymbol{\theta}$,

$$\mathbf{y}_0(t) = H(\mathbf{x}(t), t)\,\boldsymbol{\theta} \Rightarrow \mathbf{y}(t) = H(\mathbf{x}(t), t)\,\boldsymbol{\theta} + \boldsymbol{\varepsilon}$$

one can analytically calculate the *least squares estimate* (LS) $\hat{\boldsymbol{\theta}}_{LS}$ defined as

$$\hat{\boldsymbol{\theta}}_{LS} \triangleq \arg\min_{\boldsymbol{\theta}} J_{LS} = \arg\min_{\boldsymbol{\theta}} \| \mathbf{y}(t) - H(\mathbf{x}(t), t)\,\boldsymbol{\theta} \|^2 \tag{4.3}$$

being J_{LS} the least squares cost function

$$J_{LS} \triangleq (\mathbf{y}(t) - H(\mathbf{x}(t), t)\,\boldsymbol{\theta})^T (\mathbf{y}(t) - H(\mathbf{x}(t), t)\,\boldsymbol{\theta}) \tag{4.4}$$

equivalent to the squared norm of the measuring error vector $\|\boldsymbol{\varepsilon}\|^2$ for a deterministic model H. By direct calculation it follows that

$$\hat{\boldsymbol{\theta}}_{LS} = (H^T H)^{-1} H^T \mathbf{y} \tag{4.5}$$

showing that the existence of this estimator relies on the existence of the inverse of $H^T H$ (*observability condition*). If the measurement vector $\mathbf{y}$ of a given process is a stochastic variable having mean $E_{\mathbf{y}}[\mathbf{y}] = H\boldsymbol{\theta}$ and covariance $E_{\mathbf{y}}[(\mathbf{y} - H\boldsymbol{\theta})(\mathbf{y} - H\boldsymbol{\theta})^T] = \sigma^2 I$[1] being I the identity matrix, then the least squares estimate (4.5) has the following properties [56]:

(1) it is linear in $\mathbf{y}$

[1] if H is a deterministic model this is equivalent to the statement that the measurement noise has zero mean and covariance $I\sigma^2$

(2) it is *unbiased*, i.e. $E_{\mathbf{y}}[\hat{\boldsymbol{\theta}}_{LS}] = \boldsymbol{\theta}$

(3) $cov(\hat{\boldsymbol{\theta}}_{LS}) \triangleq E_{\mathbf{y}}[(\hat{\boldsymbol{\theta}}_{LS} - \boldsymbol{\theta})(\hat{\boldsymbol{\theta}}_{LS} - \boldsymbol{\theta})^T] = \sigma^2(H^TH)^{-1}$

(4) $cov(\hat{\boldsymbol{\theta}}_{LS}) \leq cov(\hat{\boldsymbol{\theta}}_{ULE})$ being $\hat{\boldsymbol{\theta}}_{ULE}$ the estimate given by *any other* unbiased linear estimator.

This last property is equivalent to the statement that in the given hypothesis the least squares estimator is a so called *best linear unbiased estimator* (BLUE). Moreover in the above equations the function $E_{\mathbf{y}}$ is the expectation operator defined as

$$E_{\mathbf{y}}[f(\mathbf{y})] = \int f(\mathbf{y})\, p_{\mathbf{y}}(\mathbf{y})\, d\mathbf{y}$$

Property (1) is immediate, properties (2) and (3) can be proven by direct calculation

$$E_{\mathbf{y}}(\hat{\boldsymbol{\theta}}_{LS}) = E_{\mathbf{y}}[(H^TH)^{-1}H^T\mathbf{y}] = (H^TH)^{-1}H^TE_{\mathbf{y}}[\mathbf{y}] = (H^TH)^{-1}H^TH\boldsymbol{\theta} = \boldsymbol{\theta}$$

and

$$\hat{\boldsymbol{\theta}}_{LS} - \boldsymbol{\theta} = (H^TH)^{-1}H^T\mathbf{y} - \boldsymbol{\theta} = (H^TH)^{-1}H^T(\mathbf{y} - H\boldsymbol{\theta}) \Rightarrow$$

$$cov(\hat{\boldsymbol{\theta}}_{LS}) \triangleq E_{\mathbf{y}}[(\hat{\boldsymbol{\theta}}_{LS} - \boldsymbol{\theta})(\hat{\boldsymbol{\theta}}_{LS} - \boldsymbol{\theta})^T] =$$

$$= (H^TH)^{-1}H^TE_{\mathbf{y}}[(\mathbf{y} - H\boldsymbol{\theta})(\mathbf{y} - H\boldsymbol{\theta})^T]H(H^TH)^{-1} =$$

$$= (H^TH)^{-1}H^T\,\sigma^2 I\,H(H^TH)^{-1} = \sigma^2(H^TH)^{-1} \Rightarrow$$

$$cov(\hat{\boldsymbol{\theta}}_{LS}) = \sigma^2(H^TH)^{-1} \tag{4.6}$$

Property (4) can be proven as follows [56]: by definition a linear unbiased estimator C is such that $C\mathbf{y} = \hat{\boldsymbol{\theta}}_{LUE}$ and

$$E_{\mathbf{y}}[\hat{\boldsymbol{\theta}}_{LUE}] = E_{\mathbf{y}}[C\mathbf{y}] = CE_{\mathbf{y}}[\mathbf{y}] = CH\boldsymbol{\theta} = \boldsymbol{\theta}$$

for any $\boldsymbol{\theta}$ so that

$$CH = I \tag{4.7}$$

By direct calculation the covariance of $\hat{\boldsymbol{\theta}}_{ULE}$ is

$$\begin{aligned} cov(\hat{\boldsymbol{\theta}}_{LUE}) &\triangleq E_{\mathbf{y}}[(\hat{\boldsymbol{\theta}}_{LUE} - \boldsymbol{\theta})(\hat{\boldsymbol{\theta}}_{LUE} - \boldsymbol{\theta})^T] = E_{\mathbf{y}}[(C\mathbf{y} - \boldsymbol{\theta})(C\mathbf{y} - \boldsymbol{\theta})^T] = \\ &= E_{\mathbf{y}}[(C\mathbf{y} - CH\boldsymbol{\theta} + CH\boldsymbol{\theta} - \boldsymbol{\theta})(C\mathbf{y} - CH\boldsymbol{\theta} + CH\boldsymbol{\theta} - \boldsymbol{\theta})^T] = \\ &= E_{\mathbf{y}}[(C\mathbf{y} - CH\boldsymbol{\theta})(C\mathbf{y} - CH\boldsymbol{\theta})^T] = CE_{\mathbf{y}}[(\mathbf{y} - H\boldsymbol{\theta})(\mathbf{y} - H\boldsymbol{\theta})^T]C^T = \\ &= CC^T\sigma^2 \end{aligned}$$

Next consider the positive semi-definite matrix DD^T being $D = C - (H^T H)^{-1} H^T$. By direct calculation the following holds:

$$\begin{aligned} DD^T &= (C - (H^T H)^{-1} H^T)(C - (H^T H)^{-1} H^T)^T = \\ &= CC^T - (H^T H)^{-1} H^T C^T - CH(H^T H)^{-1} + (H^T H)^{-1} = \\ &= CC^T - (H^T H)^{-1} \geq 0 \end{aligned}$$

which implies property (4).

With the same kind of calculations it can be shown that the BLUE estimator for a process having mean $E_{\mathbf{y}}[\mathbf{y}] = H\boldsymbol{\theta}$ and covariance $E_{\mathbf{y}}[(\mathbf{y} - H\boldsymbol{\theta})(\mathbf{y} - H\boldsymbol{\theta})^T] = \Sigma$ is given by

$$\hat{\boldsymbol{\theta}}_{WLS} = (H^T \Sigma^{-1} H)^{-1} H^T \Sigma^{-1} \mathbf{y} \tag{4.8}$$

having covariance

$$cov(\hat{\boldsymbol{\theta}}_{WLS}) \triangleq E_{\mathbf{y}}[(\hat{\boldsymbol{\theta}}_{WLS} - \boldsymbol{\theta})(\hat{\boldsymbol{\theta}}_{WLS} - \boldsymbol{\theta})^T] = (H^T \Sigma^{-1} H)^{-1} \tag{4.9}$$

This BLUE estimator may be regarded as a *Weighted Least Squares* estimator as it can be seen by direct calculation that $\hat{\boldsymbol{\theta}}_{WLS}$ given by equation (4.8) minimizes with respect to $\boldsymbol{\theta}$ the cost function J_{WLS} defined as

$$J_{WLS} \triangleq (\ \mathbf{y}(t) - H(\mathbf{x}(t), t)\ \boldsymbol{\theta}\ )^T\ \Sigma^{-1}\ (\ \mathbf{y}(t) - H(\mathbf{x}(t), t)\ \boldsymbol{\theta}\ ) \tag{4.10}$$

Equation (4.6) tells us a great deal about the precision and reliability of the estimates calculated by the least squares technique: if H is a deterministic model then σ is the standard deviation of the measurement noise and the covariance of the estimates will be proportional to it. Even more interesting is the dependance of $cov(\hat{\boldsymbol{\theta}}_{LS})$ on $(H^T H)^{-1}$. This matrix depends on the input signal $\mathbf{x}(t)$ that in the identification experiments is designed by the experimenter keeping into account all the required constraints. Indeed H is sometimes referred to as the *design matrix* in estimation theory. The existence of the inverse of $H^T H$ is the so called *observability condition* which of course depends on the input vector $\mathbf{x}(t)$. The relationship between the parameter estimate covariance and the conditioning of the regressor H can be understood considering the singular value decomposition (SVD) [57] of H

$$\begin{aligned} H &= USV^T \\ H &\in \Re^{n \times p}; U \in \Re^{n \times n}; V \in \Re^{p \times p} : U^T U = I,\ V^T V = I \\ S &= diag(s_1, s_2, \cdots, s_p) \end{aligned}$$

being $s_1, s_2, \cdots, s_p$ the singular values of H. Given this SVD decomposition, from equation (4.6) it follows that

$$\sigma^2_{\hat{\theta}_i} \triangleq (cov(\hat{\boldsymbol{\theta}}_{LS}))_i = (\sigma^2 (H^T H)^{-1})_i = \sigma^2 (S^T S)^{-1}_i = \frac{\sigma^2}{s_i^2}$$

which clearly shows the relationship between the parameter estimate variance, the measurement noise variance σ^2 and the regressor singular values. If $\mathbf{x}(t)$ guarantees $H^T H$ to be nonsingular it is called *persistently exciting* for the model. The degree of excitations provided to a system by an input can be measured by several indicators as the determinant of $H^T H$, or its trace or its condition number. The issue of designing persistently exciting inputs is the topic of a very wide literature a discussion of which goes beyond the scope of this work. For a detailed discussion of such topics refer to [56] [58].

4.1.2 Consistency and Efficiency

The concepts of consistency and efficiency are related to the asymptotic properties of an estimator as functions of the available "information". In particular an estimator, for both a random or deterministic parameter vector, is said to be *consistent* if the estimate converges to the true value in some stochastic sense [59], e.g. in the mean square sense if

$$\lim_{t\to\infty} E[(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})^T(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})] = 0$$

where the expectation is taken over $\mathbf{y}$ and $\mathbf{x}$. The concept of efficiency is instead related to the covariance of an estimator. In this regard the covariance of the estimate of either a random or deterministic parameter vector has to satisfy the *Cramer-Rao lower bound* stating that

$$\left(cov(\hat{\boldsymbol{\theta}}) - M^{-1}\right) \text{ is positive semi-definite}$$

or equivalently

$$cov(\hat{\boldsymbol{\theta}}) \geq M^{-1}$$

being M the *Fisher information matrix* defined for the deterministic parameter case as [59]

$$M \triangleq E_{\mathbf{y}}[(\nabla_{\boldsymbol{\theta}} \ln \Lambda_{\mathbf{y}}(\boldsymbol{\theta}))(\nabla_{\boldsymbol{\theta}} \ln \Lambda_{\mathbf{y}}(\boldsymbol{\theta}))^T]|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} = -E_{\mathbf{y}}[\nabla_{\boldsymbol{\theta}} \nabla_{\boldsymbol{\theta}}^T \ln \Lambda_{\mathbf{y}}(\boldsymbol{\theta})]|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}$$

being $\Lambda_{\mathbf{y}}(\boldsymbol{\theta}) \triangleq p(\mathbf{y}|\boldsymbol{\theta})$ the likelihood function, $\boldsymbol{\theta}_0$ the value of the unknown deterministic parameter vector, $\nabla_{\boldsymbol{\theta}}$ gradient operator $\nabla_{\boldsymbol{\theta}} \triangleq (\frac{\partial}{\partial\theta_1}, \frac{\partial}{\partial\theta_2}, \cdots, \frac{\partial}{\partial\theta p})^T$ and $\nabla_{\boldsymbol{\theta}} \nabla_{\boldsymbol{\theta}}^T$ the Hessian. The Fisher information matrix of a random parameter vector can be defined as above simply replacing the combined probability density function $p(\mathbf{y}, \boldsymbol{\theta})$ to $\Lambda_{\mathbf{y}}(\boldsymbol{\theta})$. An estimator is said to be *efficient* if its covariance matrix is equal to the inverse of the Fisher information matrix.

4.1.3 On the normal distribution case

If the process $\mathbf{y}$ is normally distributed with mean $H\boldsymbol{\theta}$ and covariance $\sigma^2 I$ the LS estimate is efficient and normally distributed with mean $\hat{\boldsymbol{\theta}}_{LS}$ given by equation (4.5) and covariance (4.6) [56]. The efficiency of the LS estimator in the normal case can be proven thanks to the fact that if $\mathbf{y}$ is normal the likelihood function and the Fisher infor-

mation matrix can be calculated explicitly. The fact that the estimate is itself normally distributed follows from the fact that linear functions of normal variables are normal themselves. This last property is very useful as if the parameter vector is known to be normally distributed with known mean and variance the standard Gaussian hypothesis testing technique [59] may be applied to the overfitting or model selection problem. Overfitting of the data by the model can be detected evaluating the variance of the parameter estimate. Roughly speaking, if the parameter variance is too large the parameter itself is said to be *statistically insignificant* and it might just as well be put to zero. More precisely given the two hypothesis h_0 and h_1:

$$\left.\begin{array}{l} h_0 : \theta_i = 0 \\ h_1 : \theta_i \neq 0 \end{array}\right\} \text{such that } p(\text{accept } h_1 |\ h_0 \text{ true}) = \alpha$$

being θ_i normally distributed with standard deviation σ_{θ_i} and α some arbitrary constant (usually 5%), the h_1 hypothesis is accepted if $\frac{|\theta_i|}{\sigma_{\theta_i}} > c(\alpha)$ being $c = 1.96$ if $\alpha = 5\%$[2]. In percentile notation it can be said that if the parameter relative percentile error is larger then 51.02%, i.e.

$$100\frac{\hat{\sigma}_{\theta_i}}{|\hat{\theta}|} > \frac{100}{1.96}\% = 51.02\ \% \tag{4.11}$$

there is a 95% confidence limit that the parameter itself is statistically insignificant and the hypothesis $h_0 : \theta_i = 0$ is better to be accepted. Moreover in the normal case the measurement variance σ^2 normalized sum of the squared residuals (4.4) $\frac{J_{LS}(\hat{\boldsymbol{\theta}}_{LS})}{\sigma^2} = (\mathbf{y} - H\hat{\boldsymbol{\theta}}_{LS})^T(\mathbf{y} - H\hat{\boldsymbol{\theta}}_{LS})/\sigma^2$ has a $\chi^2(\nu)$ distribution of mean ν and standard deviation $\sqrt{2\nu}$ being $\nu \triangleq dim(\mathbf{y}) - dim(\boldsymbol{\theta})$ the number of degrees of freedom of the fit. As a consequence the value of the normalized cost function $J_{LS}(\hat{\boldsymbol{\theta}}_{LS})/\sigma^2$ can be used to measure quantitatively the goodness of the fit: a rule of thumb for a moderately good fit is that $\chi^2 \approx \nu$. Actually the value of $J_{LS}(\hat{\boldsymbol{\theta}}_{LS})/\sigma^2$ is generally used to test for underfitting as if it is larger then some threshold c, fixed so that the area under the χ^2 distribution between c and infinity is more then $\alpha\%$ being usually $\alpha = 1$ or 5, the data is said to be underfitted by the model. Notice that in order to evaluate underfitting the measurement variance σ^2 must be known. On the contrary, if it is not known, then assuming that the fit is good, the value of $J_{LS}(\hat{\boldsymbol{\theta}}_{LS})$ may be used to estimate the measurement variance σ^2 as shown in the following section. In such circumstance one is not allowed to use $J_{LS}(\hat{\boldsymbol{\theta}}_{LS})$ to asses underfitting anymore.

As far as deterministic models are concerned, i.e. $\mathbf{y}$ has a joint Gaussian distribution with mean $H\boldsymbol{\theta}$ and known covariance Σ, the WLS estimate $\hat{\boldsymbol{\theta}}_{WLS}$ (equation (4.8)) is equivalent to the MLE estimate $\hat{\boldsymbol{\theta}}_{MLE}$ (equation (4.2)). This follows from the fact that

[2] If $\alpha = 5\%$ then c is calculated such that if $N(x,0,1)$ is the normal distribution of x having zero mean and unit variance, $\int_{-c}^{c} N(x,0,1)dx = 1 - \alpha = 0.95 \Rightarrow c = 1.96$

in the given hypothesis the likelihood function $\Lambda_{\mathbf{y}}(\boldsymbol{\theta})$ is

$$\Lambda_{\mathbf{y}}(\boldsymbol{\theta}) = p(\mathbf{y}|\boldsymbol{\theta}) = \frac{1}{\sqrt{(2\pi)^n \det \Sigma}} \exp\{-\frac{1}{2}(\mathbf{y} - H\boldsymbol{\theta})^T \Sigma^{-1} (\mathbf{y} - H\boldsymbol{\theta})\}$$

($n \triangleq \dim(\mathbf{y})$) that is maximized minimizing the exponential argument, i.e. J_{WLS} (4.10). The consistency of the LS, WLS and MLE estimators is a known fact a proof of which may be found in any text book on estimation theory. As far as the efficiency of the MLE estimator is concerned the following theorem is reported from [56]: if an efficient unbiased estimator exists, then it is also the maximum likelihood estimator.

4.1.4 Measurement variance estimation

In order to use equation (4.6) in practice the variance σ^2 of the stochastic process $\mathbf{y}$ must be known. As noticed in the previous footnote (1), if the model H is deterministic such variance is the measurement noise variance which is thus usually known. Nevertheless it is not infrequent that such variance is not known, e.g. if $\mathbf{y}$ is calculated through some other model of unknown reliability, and must be estimated as well. An unbiased estimate of σ^2 is provided by [56]:

$$\hat{\sigma}^2 = \frac{J_{LS}(\hat{\boldsymbol{\theta}}_{LS})}{\dim(\mathbf{y}) - \dim(\boldsymbol{\theta})} = \frac{(\mathbf{y} - H\hat{\boldsymbol{\theta}}_{LS})^T(\mathbf{y} - H\hat{\boldsymbol{\theta}}_{LS})}{\dim(\mathbf{y}) - \dim(\boldsymbol{\theta})} \tag{4.12}$$

In the normal case, i.e. if $\mathbf{y}$ is normally distributed with mean $H\boldsymbol{\theta}$ and variance σ^2, such estimator is optimal in the sense that is a *minimum variance unbiased estimator* (MVUE) for σ^2[56].

4.2 On board sensor based ROV identification

As discussed in the first chapter of this work, the navigation and control systems design of variable configuration ROVs are strictly related to the degree of knowledge of the vehicles dynamic model. As these models are subject to mission dependent changes an on board sensor based system identification approach is highly recommended in order to be able to identify the most important dynamic parameters by simple in water tests rather than complex, time consuming and expensive towing tank techniques. Moreover being the ROV models linear in their parameters a least squares technique will be adopted as the LS estimator has been shown to be either the maximum likelihood one in the Gaussian case, or the best linear unbiased estimator in the more general case. Indeed also other estimation techniques as Kalman filter based ones or estimation error minimization by simulated annealing algorithms have been tested by the author as accounted in [60] [61]. The experimental results regard the identification of a simplified model of the heave, surge, sway and yaw axis of the ROMEO ROV of CNR-IAN. ROMEO

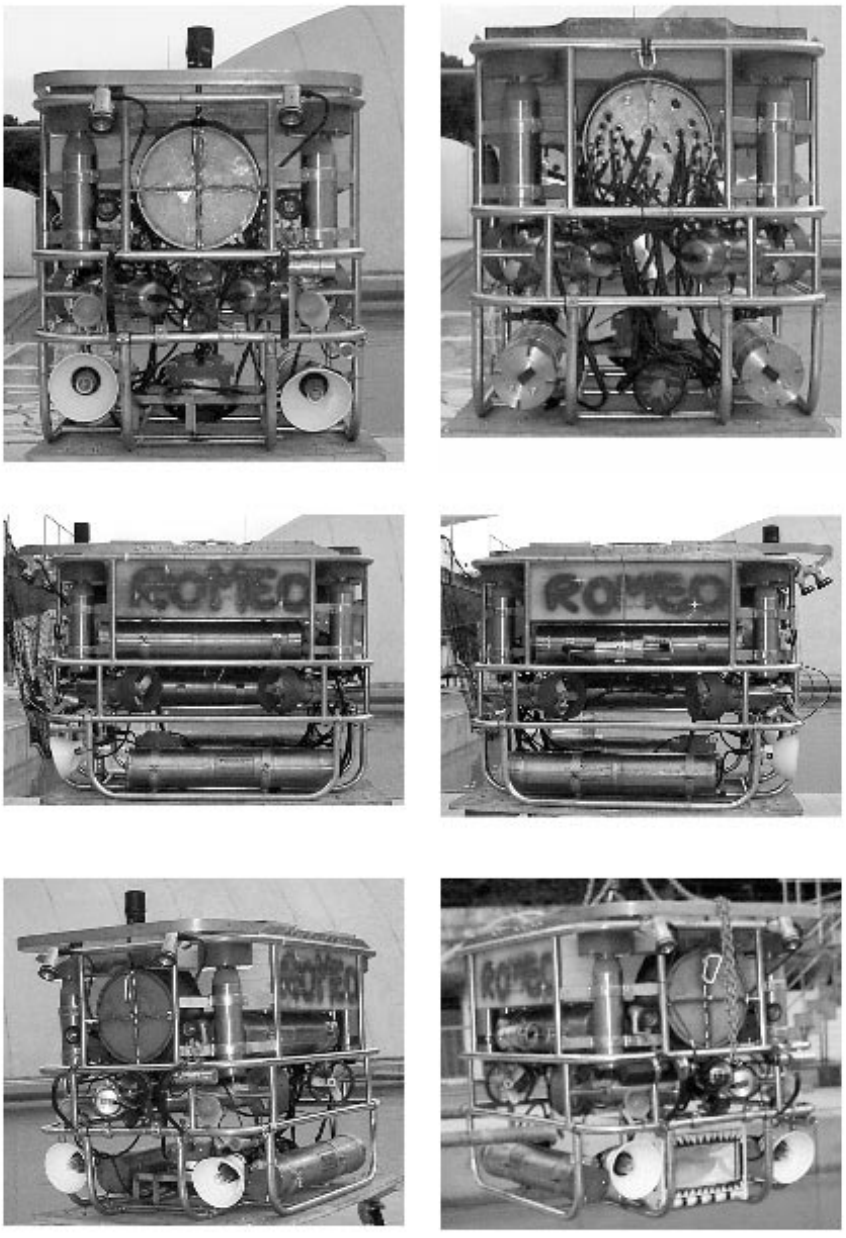
which is depicted in figure (4.1) is about $1m$ in height, $0.9m$ in width and $1.3m$ in length and its weight in air is of about $450Kg$. As shown in figure (4.1) the bottom of the vehicle carries a skid frame for payload and 2 cylindrical canisters for batteries while the upper part is made of a cylindrical canister for the electronics, 4 thrusters for propulsion in the horizontal plane, 4 for the vertical one, several instruments and sensors and, on the top, foam for buoyancy. The thrusters, canisters, and instruments are allocated so that the overall structure of the vehicle is symmetric with respect to both the xz and yz planes.

4.2.1 Model structure

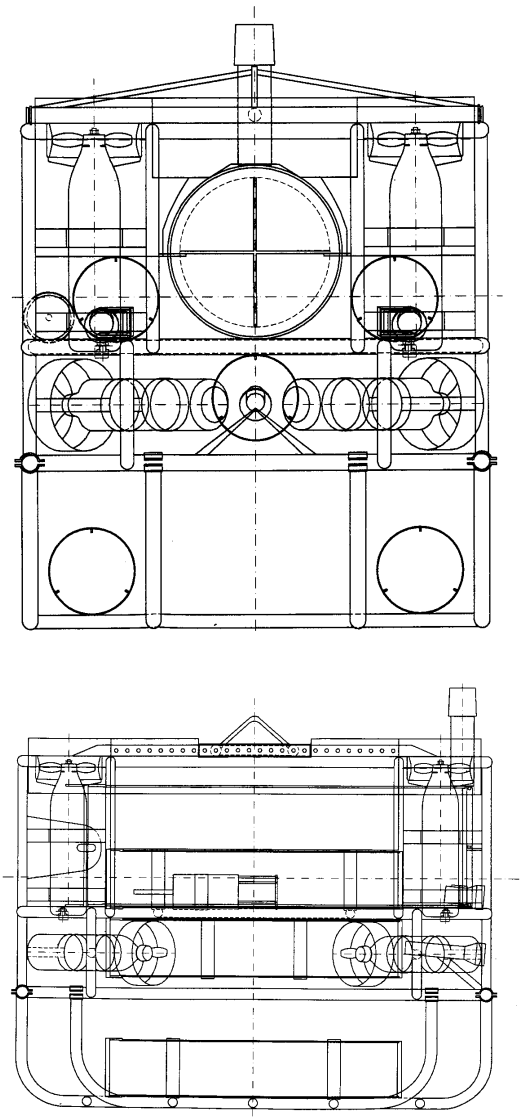
The experimental identification of a complete ROV model as the one given by equation (3.52) is not feasible with only standard on board sensors because it would require a complete state knowledge. Indeed it may be performed with more complex and expensive towing tank facilities as described by Goheen et al.[62] or Fryxell [63], but such approach is not indicated for systems having a variable and mission dependent configuration. Moreover in many standard manoeuvreing conditions, e.g. plane surge motion or vertical translation, and generally at low operating speeds, the coupling terms may be reasonably neglected without serious loss of information. As a consequence on board sensor based identification experiments usually refer to a simplified uncoupled model that can be deduced from equation (3.52) neglecting the off diagonal elements of the added mass matrix, the Coriolis and centripetal kinematic coupling terms and the drag ones. This approximation relies on the facts that *(i)* the off diagonal elements of the added mass matrix of a rigid body having three symmetry planes are identically null [34], *(ii)* the off diagonal elements of such positive definite matrix are much smaller than their diagonal counterparts [32], *(iii)* the hydrodynamic damping coupling is negligeable at low speeds. The resulting model structure for a single degree of freedom is:

$$m\dot{\xi} = -k_\xi \xi - k_{\xi|\xi|}\xi|\xi| + \tau_\xi + \varepsilon \tag{4.13}$$

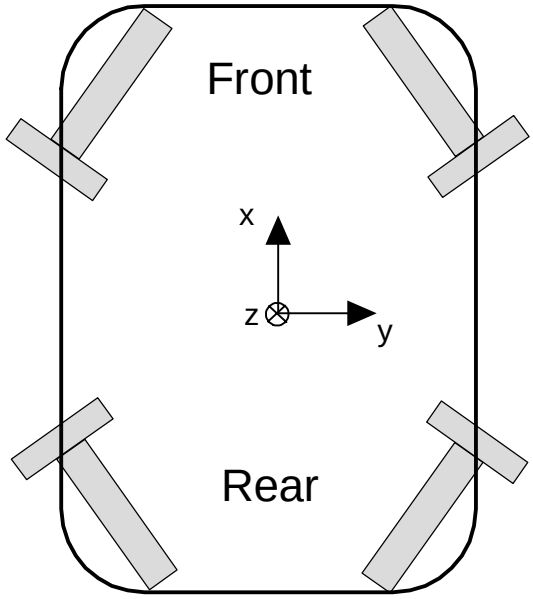
being m the inertia relative to the considered degree of freedom, ξ the $1D$ velocity (surge, sway, heave, yaw, pitch or roll rate) k_ξ and $k_{\xi|\xi|}$ the linear and quadratic drag coefficients, τ_ξ the applied force or torque and ε a disturbance modelling otherwise unmodeled phenomena as cable effects. This kind of uncoupled model structure is certainly one of the most common in the literature of underwater vehicles: as far as identification experiments are concerned it has been adopted, for example, to identify the yaw motion of the IFREMER VORTEX vehicle [64] or the surge motion of the NPS PHOENIX AUV [65]. Equation (4.13) relative to the heave, surge, sway, and yaw axis of the ROMEO ROV has been experimentally identified as described in the following sections. To estimate the parameter vector $\boldsymbol{\theta} = (m, k_\xi, k_{\xi|\xi|})^T$ from equation (4.13) the torque τ_ξ is assumed known and a linear regressor is considered. The knowledge of τ_ξ is actually related to the fact that the relation between applied thruster voltage and torque has been a priori identified for each single thruster in a thrust tunnel as de-



4.1.ROMEO: the bottom right pictures shows a different payload configuration, a palnkton sampling equipment



4.2.The ROMEO ROV



4.3.Sketched top view of ROMEOs horizontal thruster configuration.

scribed in the following section. A potentially serious drawback of such methods could be related to the fact that the identified thruster model does not take into account the propeller-propeller or propeller-hull interactions that occur on the vehicle in the operating conditions. These phenomena are due to the fact that on the great majority of ROVs more thrusters (8 on ROMEO) are present on the vehicle and may thus interfere between themselves or with the hull. As far as the ROMEO ROV is concerned this can be more easily understood with reference to figure (4.3) in which the horizontal thruster disposition is schematically depicted. If all four thrusters operate at the same time, it is reasonable to expect that the front and rear ones on the same side of the vehicle interfere with each other. It is also expected that the front ones will experience a propeller hull interaction pushing backwards as the rear ones pushing forward. As a consequence the efficiency of the thrusters is expected to be less then the one measured in the thrust tunnel. As will be described in the following sections, to model this phenomena an efficiency parameter has been introduced. This technique has been shown to be effective for the modeling of both the propeller-propeller interactions and the propeller hull ones. The experimental results reported in the following show that the propeller-hull and propeller-propeller interactions have indeed a most relevant effect on ROV dynamics. Nevertheless this topic has not been systematically addressed by the underwater robotics scientific community: to the knowledge of the author the only relevant reference to this phenomenon in ROV systems is due Goheen and Jefferys [66] who describe a *thruster installation coefficient*. In their words [66] the installation coefficient takes "*into account the differences in force that the thruster provides when it is operating in the proximity of the ROV, as opposed to when it is tested in open water*".

At last notice that in order to estimate $\boldsymbol{\theta} = (m, k_\xi, k_{\xi|\xi|})^T$ and the eventual efficiency parameters a two step procedure has been implemented: first a the drag and efficiency coefficients are estimated by constant velocity experiments, then with the aid of their knowledge a sinusoidal torque input is designed in order to identify the inertia m.

4.2.2 Thruster model identification

The modelling and control of underwater vehicle thruster systems has received a wide attention in the literature of the last years [45] [46] [44]. As shown by Yoerger et. al.[44], within the theory of ideal fluids a lumped parameter thruster model is given by (section (3.3.1))

$$\begin{aligned} \tau &= C_t n|n| \\ \dot{n} &= \beta T - \alpha n|n| \end{aligned}$$

being τ the output thruster force, C_t, β and α constant parameters, n the propeller revolution rate and T the input torque. Generally the servo velocity loop of the velocity controlled thruster system has a negligible time constant with respect to the overall vehicles' one [19], and thus the thruster dynamics can be neglected. Indeed most often [32]

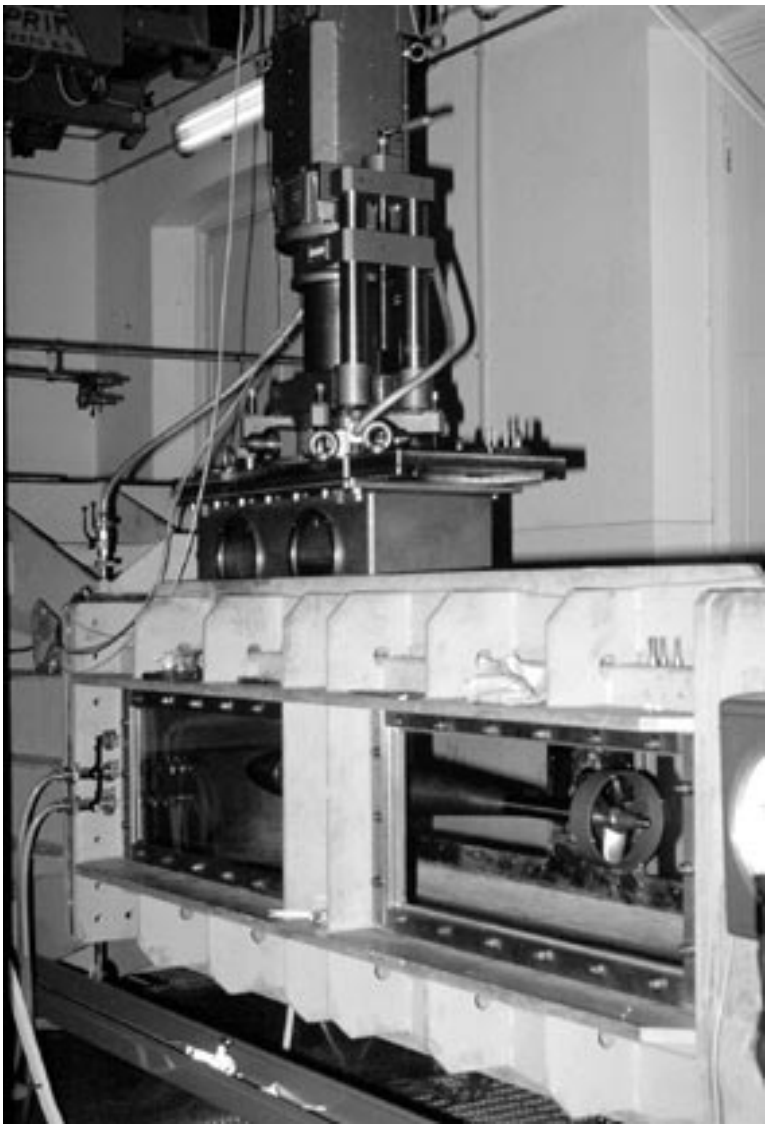
the applied thrust is modeled as $\tau = C_t|n|n - C_s|n|v$ being v the velocity of the fluid through the thruster (*velocity of advance*) and $-C_s|n|v$ a saturation term. In virtue of the creeping motion of UUVs, this last saturation term can be neglected in many standard operational conditions as widely accepted in the literature [32] [14] [44] [64]. Moreover in steady state conditions the neglected thrust drag term $-C_s|n|v$ will be somehow taken into account by the drag forces considered in the equation of motion (4.13) of the vehicle. Thus, neglecting the motor dynamics, the thruster force may be modelled as

$$\tau = CV|V| \tag{4.14}$$

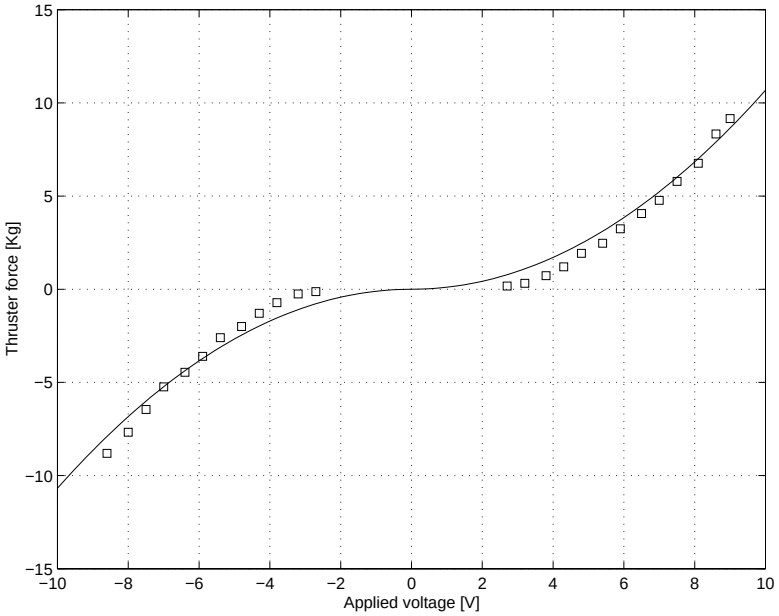
being C an unknown constant and V the applied control voltage. In many marine applications two different C parameters are requested for the positive and negative propeller revolution rates, as the thrusters do not behave symmetrically in the two directions, but most frequently UUV propellers, as ROMEOs ones, are designed to exhibit a symmetrical behaviour in the two directions. Equation (4.14) has been identified [67] for each of the eight ROMEO thrusters putting the whole thruster (motor and propeller) in a thrust tunnel (figure (4.4)) and measuring the force τ as a function of the input voltage V. Typical results of this measuring and identification method are shown in figure (4.5). It should be noticed that having neglected the velocity of advance, the proposed model is expected to be more accurate far from the propeller revolution rate inversion points. In particular high frequency sign changes of n, that may occur during hovering manoeuvres or would occur with pseudo random binary inputs typical of identification experiments, produce unmodeled turbulence next to the thrusters making the output thrust computed by the standard model less accurate.

4.2.3 Off line velocity estimation

As stated above, the proposed identification scheme consists in two steps: first the drag coefficients are estimated by constant velocity tests, and then their values are adopted to design a suboptimal inertia identification experiment with a sinusoidal input. Both steps are based on position measurements only so that a major issue is velocity estimation. As far as the drag experiments are concerned a simple least squares fitting of the position data is enough, but for the inertia identification tests a different filtering technique is required. More generally the problem of computing the numerical derivative of a signal given noisy samples is posed. Among the many possible signal processing techniques to face this problem attention is focused on the use of the Savitzky-Golay filters [68]. These are low pass filters designed in the time domain rather then in frequency domain. Within a moving window containing n_l points on the left and n_r on the right of the i^{th} data sample, the $n_l + n_r + 1$ points are least squares fitted with a polynomial of degree m and the filtered value of the i^{th} data sample is assumed to be the value of the polynomial in i. The derivative of the given signal in i is thus assumed to be the derivative of the polynomial in i. Notice that the fitted polynomial is adopted in the i^{th} point only, as when the moving window is shifted of one point the whole procedure is



4.4.Cavitation tunnel tests: preliminary propeller test. The thruster identification has been carried out putting the whole thruster in the tunnel.



4.5.Cavitation tunnel identification tests.

repeated. In order to design a Savitzky-Golay filter the windows left and right lengths n_l, n_r and the polynomial degree m must be chosen. If the data is processed off line n_r can be chosen to be non null so that the filter is non-causal: as far as the polynomial degree m is concerned it can be chosen adaptively as proposed in [69], but generally [68] m is fixed to 2 or 4. For most applications the moving window can be chosen to be symmetrical ($n_l = n_r$): guidelines for the choice of n_l and n_r may be found in [68]. Savitzky-Golay filters, that are most common among the chemists for noisy spectrometric data analysis, are among the most "natural" tools for derivative estimation. A detailed analysis of their properties goes beyond the scopes of this work, yet to have a qualitative understanding of their performance an example based on the equations of our interest is reported: consider the linear system

$$\begin{aligned} m\ddot{x} &= f - k_l\dot{x} \\ \text{being} \quad &: \quad f = f_0 + \Delta_f \sin(\omega t) \end{aligned}$$

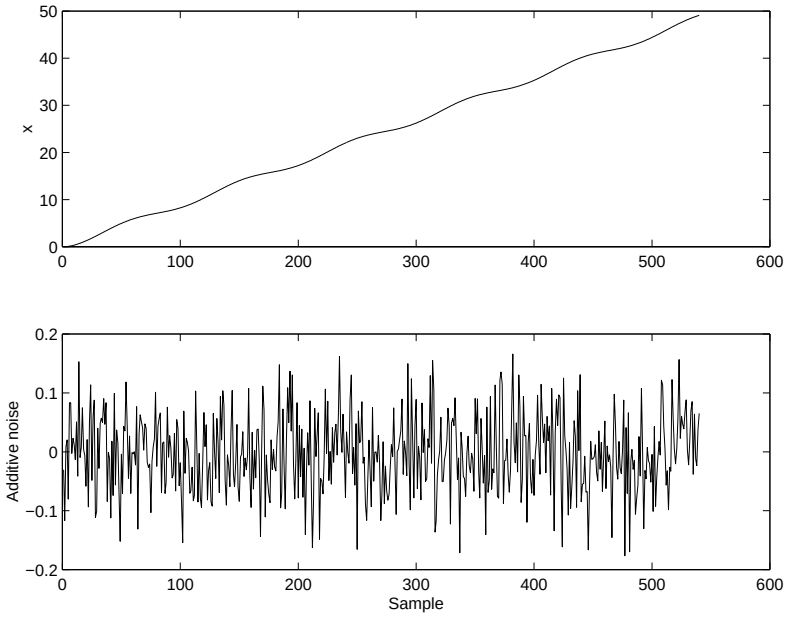
By direct calculation it follows that

$$\begin{aligned} \dot{x} &= x_0 e^{-t/\tau} + \frac{f_0}{k_l}(1 - e^{-t/\tau}) + \frac{\Delta_f}{k_l}\frac{\sin(\omega t) - \omega\tau(\cos(\omega t) - 1)}{(1 + \omega^2\tau^2)} \\ x &= \tau(x_0 - \frac{f_0}{k_l})(1 - e^{-t/\tau}) + \frac{f_0}{k_l}t + \frac{\Delta_f}{k_l}\frac{\omega\tau t - \tau\sin(\omega t) - \frac{1}{\omega}\cos(\omega t) + \frac{1}{\omega}}{(1 + \omega^2\tau^2)} \\ \text{being} \quad &: \quad \tau = m/k_l \end{aligned}$$

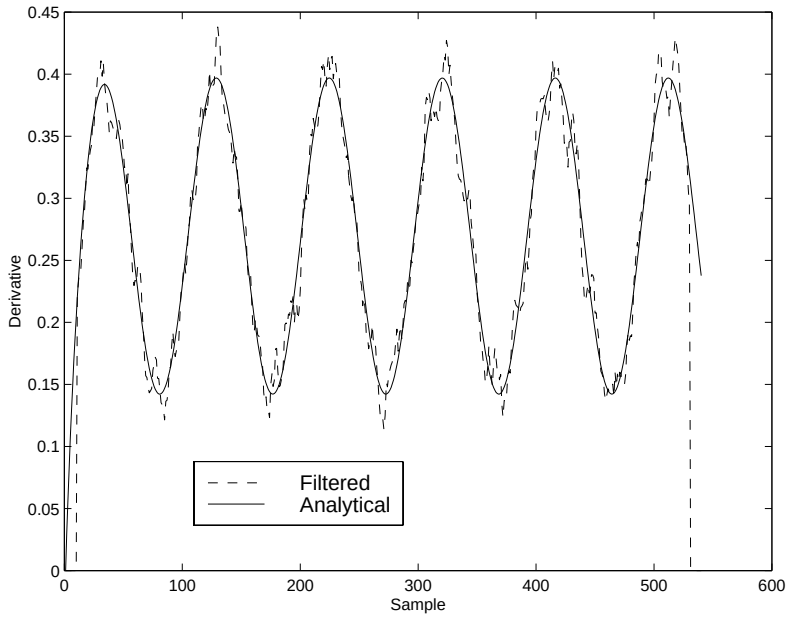
Assuming $f_0 = 35$, $\Delta_f = 25$, $\omega = 0.1963$, $m = 500$, $k_l = 170$, $x_0 = 0$, the position x on a $180s$ test evolves as shown in the top plot of figure (4.6) having adopted a $3Hz$ sampling rate, i.e. 540 data samples. Adding to this signal the zero mean normal noise having 0.07 standard deviation shown in the bottom plot of figure (4.6) and filtering the so computed noise corrupted position signal with a 4^{th} order Savitzky-Golay filter having $n_l = n_r = 10$ yields the result displayed in figure (4.7).

4.2.4 Heave model identification

As far as the heave axis is concerned, off-line identification has been performed to estimate linear and quadratic drag coefficients and buoyancy force [70]. The data for the identification experiments consists in depth and thrust measurements collected during up and down motions performed under the Antarctica ice-canopy during the XIII Italian Antarctica Expedition (1997-1998). Depth was measured directly by a depth-meter with a $10Hz$ sampling rate, while thrust was estimated by the thrust-tunnel identified model described in the previous subsection. Five different experiments, in the sequel labelled with numbers 1 to 5, have been performed with inputs of the kind shown in figure (4.8) each with a different vehicle weight. During experiment 1 and 2 the vehicle was positive, during experiment 3 it was roughly neutral and in the last two it was



4.6.Above: analytical position signal. Below: additive, zero mean, normal noise having standard deviation 0.07.



4.7.Estimated and real velocity $\dot{x}$.