On the use of kriging for enhanced data reconstruction in a separated transitional flat-plate boundary layer

Hasan Gunes[1,a)] and Ulrich Rist[2,b)]
[1]*Department of Mechanical Engineering, Istanbul Technical University, Gumussuyu, 34437 Istanbul, Turkey*
[2]*Institut für Aerodynamik und Gasdynamik, Universität Stuttgart, Pfaffenwaldring 21, D-70550 Stuttgart, Germany*

(Received 18 February 2008; accepted 24 September 2008; published online 31 October 2008)

Kriging method for data reconstruction and spatial enhancement of stereo-particle image velocimetry (S-PIV) data for a transitional boundary layer with a laminar separation bubble is investigated. Particularly, the effect of various variogram models and their parameters are studied in detail. In addition, we show that missing data clusters, or black zones, which often occur in PIV measurements, can be estimated using kriging provided the data are well correlated. An important issue in PIV measurements is that built-in PIV data processing software might have problems to detect or to correct spurious erroneous vectors called "outliers." It is shown that these outliers can be eliminated or greatly alleviated using kriging. λ_2 isosurface and stream traces show that noisy vortical structures are eliminated but the main coherent structures are well preserved and smoothed, thus procedures for the detection and tracking of vortex core lines can be effectively applied on kriged data. Analytical test data for a more quantitative evaluation of the performance of kriging are given in the Appendices. © *2008 American Institute of Physics*. [DOI: 10.1063/1.3003069]

I. INTRODUCTION

With the ever increasing development of new and better optical and microelectronic devices, particle image velocimetry (PIV) has been increasingly used in experimental fluid dynamics as a nonintrusive, multipoint optical technique to obtain quantitative data for complex flows. Nowadays, stereoscopic, scanning/multiplane and tomographic PIV measurements provide fully time-resolved three-dimensional (3D) flow fields similar to the flow fields obtained by numerical simulation (see, for example, Refs. 1–5). While PIV measurements supply researchers with experimental quantitative information for flow fields as never obtained previously, the acquisition of "real" flow data by the method is complicated and requires certain conditions to be satisfied for a reliable image quality. If the image quality is poor, for example, due to a poor setup and alignment involving insufficient illumination, low levels of contrast, low or inhomogeneous seed density, etc., the cross-correlation evaluation of moving particle displacements from the low-quality images often results in spurious/erroneous velocity vectors. To alleviate this problem, PIV processing software is available to detect and remove erroneous vectors automatically (see, for example, edPIV reported in Ref. 6). However, because the erroneous vectors are identified by comparing the results of immediate neighbors, they may not be detected if the "bad" vectors have very few immediate neighbors especially near the boundaries. In addition, for a certain number of reasons (e.g., shadowing, insufficient illumination, obstructed view, etc.) snapshots with missing data clusters (region of "black zones") may occur in some regions of the measurement domain. Therefore, it is essential that before analyzing and postprocessing of PIV data using various coherent structure extraction and vortex identification techniques such as λ_2 visualization and proper orthogonal decomposition (POD), the above mentioned erroneous vectors (outliers) and missing data clusters (black zones) need to be repaired/estimated. In fact, in order to compute the standard POD modes of the velocity field, each snapshot must be complete [except for gappy POD applications which can be used to estimate missing data (see Refs. 7–9)].

Moreover, even in the absence of missing data clusters and outliers, often the resolution of PIV measurements are not sufficiently high for many of the data visualization methods that usually involve the evaluation of the derivatives of the field variables, which further increase the noise level of the experimental data because they amplify it.

In this paper, we demonstrate that kriging can be used as an effective tool (i) to remove/alleviate outliers, (ii) to estimate the missing data clusters (black zones), and finally (iii) to reconstruct the PIV data for further coherent structure and vortex extraction analyses on a much finer and smoother mesh.

II. KRIGING METHOD

Kriging is an unbiased estimation procedure that uses known values and a variogram to determine unknown values. Based on the variogram, optimal weights are assigned to known values in order to estimate the data at unknown points. The variogram characterizes the spatial continuity or roughness/smoothness of a data set.[10] The variogram analysis consists of first constructing an experimental variogram from the available data and fitting a suitable variogram

[a)]Telephone: +90 212 293 1300. Fax: +90 212 245 0795. Electronic mail: guneshasa@itu.edu.tr.
[b)]Telephone: +49 711 685 3432. Fax: +49 711 685 3438. Electronic mail: rist@iag.uni-stuttgart.de.

model to the experimental variogram. The experimental variogram is calculated by averaging one-half the difference squared of the values over all pairs of observations with the specified separation distance h and possible directions as follows:

$$\gamma(h_{i,j}) = \frac{1}{2L_i}\sum_{i}^{L_i}(z_i - z_j)^2, \tag{1}$$

where z_i is the value of the variable at point i, z_j $(j=i+h)$ is the value of the variable at h separation distance away from point i, L_i is the total number of data points for a given separation distance, and γ is the variance. Important characteristic features of the variogram are the *range* and the *sill* (also the minimum variance value is called the *nugget* effect but it will not be useful in PIV data smoothing). The range represents the distance at which there is no longer a correlation between the points, whereas the sill is the average variance of points at such a distance away from the point in question that there is no correlation between the points. When the variance is normalized using the value of the sill, the correlation between points decreases as $\gamma \to 1$ and increases as $\gamma \to 0$.

The variogram model is usually chosen from a set of mathematical functions that describe the spatial relationship of the data. Often used functions are polynomials, exponentials, spherical, spline and the Gaussian functions.[10,11] The selection of a suitable variogram model is a crucial step of the kriging procedure, as it has an important effect on the weights and estimation error. In addition, each model function contains a correlation parameter by which the estimation/smoothing level can be controlled as desired.

Using the variogram, the weights W_i can be found by solving the following system of linear equations:

$$\underset{\sim}{\gamma}(h_{i,j})\underset{\sim}{W}_i = \underset{\sim}{\gamma}(h_{i,p}), \tag{2}$$

where the coefficient matrix $\gamma(h_{i,j})$ can be calculated using the available (design) data set via Eq. (1) once and for all, while the right hand side vector $\gamma(h_{i,p})$ needs to be evaluated for each unknown point, p. Then, a standard LU decomposition or since the coefficient matrix is symmetric and positive definite, a more efficient Cholesky factorization can be used to solve Eq. (2) for each unknown point. Note that Eq. (2) corresponds to simple kriging, in which the mean (i.e., the expectation in probability theory) is assumed to be zero. On the other hand, ordinary kriging, being the most commonly used type of kriging, assumes a constant but unknown expectation. This leads to the introduction of an additional parameter λ (a Lagrange multiplier), which is used for the minimization of the kriging error. This, in turn, allows us to have a constraint on the weights as follows:

$$\sum_{i=1}^{n} W_i = 1. \tag{3}$$

Then, the modified linear system for the ordinary kriging consisting of n data points can be written as follows:

$$\begin{bmatrix} \gamma(h_{11}) & \gamma(h_{12}) & \dots & \gamma(h_{1n}) & 1 \\ \gamma(h_{21}) & \gamma(h_{22}) & \dots & \gamma(h_{2n}) & 1 \\ : & : & \dots & : & : \\ \gamma(h_{n1}) & \gamma(h_{n2}) & \dots & \gamma(h_{nn}) & 1 \\ 1 & 1 & \dots & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} W_1 \\ W_2 \\ : \\ W_n \\ \lambda \end{bmatrix} = \begin{bmatrix} \gamma(h_{1P}) \\ \gamma(h_{2P}) \\ : \\ \gamma(h_{nP}) \\ 1 \end{bmatrix}. \tag{4}$$

Having obtained weights through solving the linear system, each new (unknown) point, $Z(x,y)$ can be estimated simply as linear combination of the available data as follows:

$$Z(x,y) = \sum_{i=1}^{n} W_i z_i. \tag{5}$$

In this paper, we employ ordinary kriging and refer to Refs. 10–13 for theory, details, and types of kriging.

For the past several decades, application of kriging has been exclusively used with success mainly in geostatistics, meteorology, and environmental sciences. Recently, application of the procedure is extended to process engineering, health sciences, and thermal and fluid sciences. For example, kriging has been successfully applied to data recovery and reconstruction of randomly generated laminar gappy flow fields of uniform flow past a circular cylinder (Refs. 7 and 9). It has been shown that kriging has a number of advantages over other reconstruction techniques such as gappy POD when temporal resolution of the flow data is not sufficiently high. More recently, an investigation of spatial resolution enhancement/smoothing of stereo-PIV data for a transitional flat-plate boundary layer flow involving a laminar separation bubble has been presented using both POD and kriging methods.[14] In the current paper, while we extend our application of kriging concentrating on the effect of various variogram models, we particularly investigate outlier resistance of the procedure and the treatment of missing data clusters which often occur in PIV data. In addition, we compare smoothing capabilities of kriging to low-pass digital filters commonly used in data smoothing.

III. STEREO-PIV DATA

The stereo-PIV experiments for a transitional boundary layer flow containing a laminar separation bubble were carried out by Lang.[15] For detailed information on experimental setup, findings, etc., we refer to Refs. 3 and 15. Here, we briefly inform the reader about the investigated flow which contains nonlinear transition stages in a laminar separation bubble. The basic transition scenario is that when a laminar boundary layer separates in a region of adverse pressure gradient on a flat plate such that a laminar separation bubble forms and the flow reattaches to form a turbulent boundary layer due to sudden development of 3D disturbances.

The experiments were performed in a laminar water tunnel facility at the Institute of Aerodynamics and Gas Dynamics in Stuttgart University. A flat plate is mounted on the free stream $(U_\infty = 0.125\ \text{m/s})$ of the test section of a low-disturbance water tunnel. To generate a pressure induced laminar separation bubble, a displacement body (length of the body $L=0.69\ \text{m}$) was positioned in the test section above

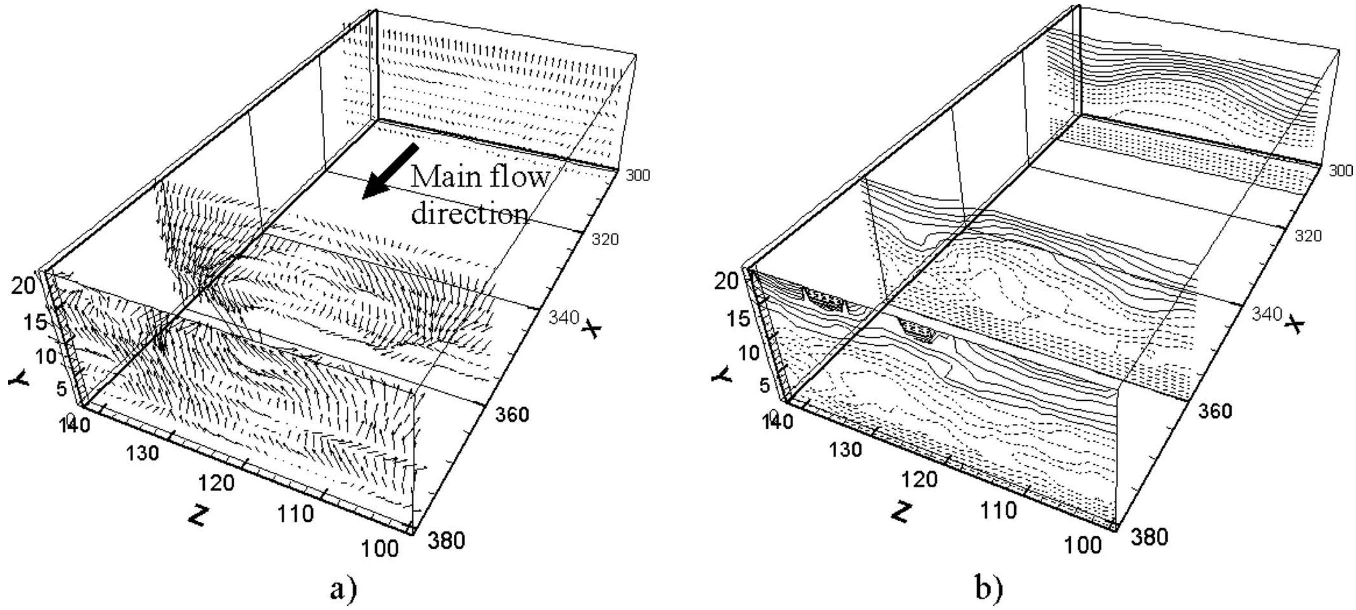

FIG. 1. (a) Velocity vectors (tangent to the plane) at planes $x=310$, $x=360$, and $x=381$. (b) Contours of streamwise component of velocity at the same planes (dotted lines correspond to reverse flow). Dimensions are in millimeters.

the flat plate. The global Reynolds number based on the displacement body length and the reference velocity ($U_\infty = 0.125$ m/s) is $\mathrm{Re}=10^5$ in water. In order to compare experimental results, direct numerical simulations are performed by Marxen.[16] In simulations, general physical parameters of the flow are chosen to match the experimental setup as closely as possible. It is shown that instantaneous experimental and numerical results closely agree with each other (see Ref. 17). Figure 1 shows the 3D PIV measurement domain with instantaneous (phase averaged) velocity vectors (tangent to the plane) at downstream distances of $x=310$, $x=360$, and $x=381$ mm as well as the streamwise velocity component at the same planes. Coordinates X, Y, and Z correspond to streamwise, cross-flow, and spanwise flow directions, respectively, and dimensions are in millimeters. There is a large reverse flow as represented by dotted lines in Fig. 1(b) due to the separation bubble. While the flow is mainly one directional in the upstream part of the separation bubble (e.g., at $x=310$ mm), it is very complex and fully 3D for the downstream part of the separation bubble ($x>360$). Table I shows the minimum and maximum normalized velocity components as well as the difference of the velocity components for selected downstream planes. It is seen that the cross-flow and spanwise components of velocity increases an order of magnitude along the downstream flow direction. In the experiment the shear layer instability has been controlled by an oscillating wire upstream of laminar separation and a steady spanwise roughness elements (see Refs. 3 and 15) such that the *unsteady* nature of the flow became accessible via phase averaging of individual snapshots. The disturbance cycle has been divided into 18 equally spaced phase intervals and each phase average has been computed from 25 instantaneous measurements at the according phase of the oscillating wire. Thus, a faithful representation of the 3D instantaneous flow field became available.

IV. SMOOTHING/ENHANCEMENT OF STEREO-PIV DATA

It has been shown in Ref. 14 that kriging based on the Gaussian variogram model can be used to smooth the spanwise component of the vorticity field and the smoothness can be controlled by variogram/correlation parameters. By smoothing we mean to merely "re-evaluate" the PIV data at its measurement points using kriging. Enhancement refers to, on the other hand, interpolation and smoothing of the data. In

TABLE I. The maximum and minimum values of velocity components at selected downstream planes, X (in mm). The velocity components are normalized using the maximum speed in 3D field [$V_{\max}=\sqrt{(u^2+v^2+w^2)}=92.5$ mm/s].

	$u_{\max}$	$u_{\min}$	Δu	$v_{\max}$	$v_{\min}$	Δv	$w_{\max}$	$w_{\min}$	Δw
$X=301$	0.995	−0.612	1.607	0.096	−0.026	0.122	0.035	−0.011	0.046
$X=310$	0.950	−0.599	1.549	0.108	−0.036	0.144	0.053	−0.021	0.074
$X=330$	0.953	−0.687	1.640	0.220	0.0	0.220	0.104	−0.142	0.246
$X=360$	0.754	−0.682	1.436	0.081	−0.335	0.416	0.264	−0.122	0.386
$X=381$	0.972	−0.682	1.654	0.346	−0.375	0.721	0.222	−0.210	0.432

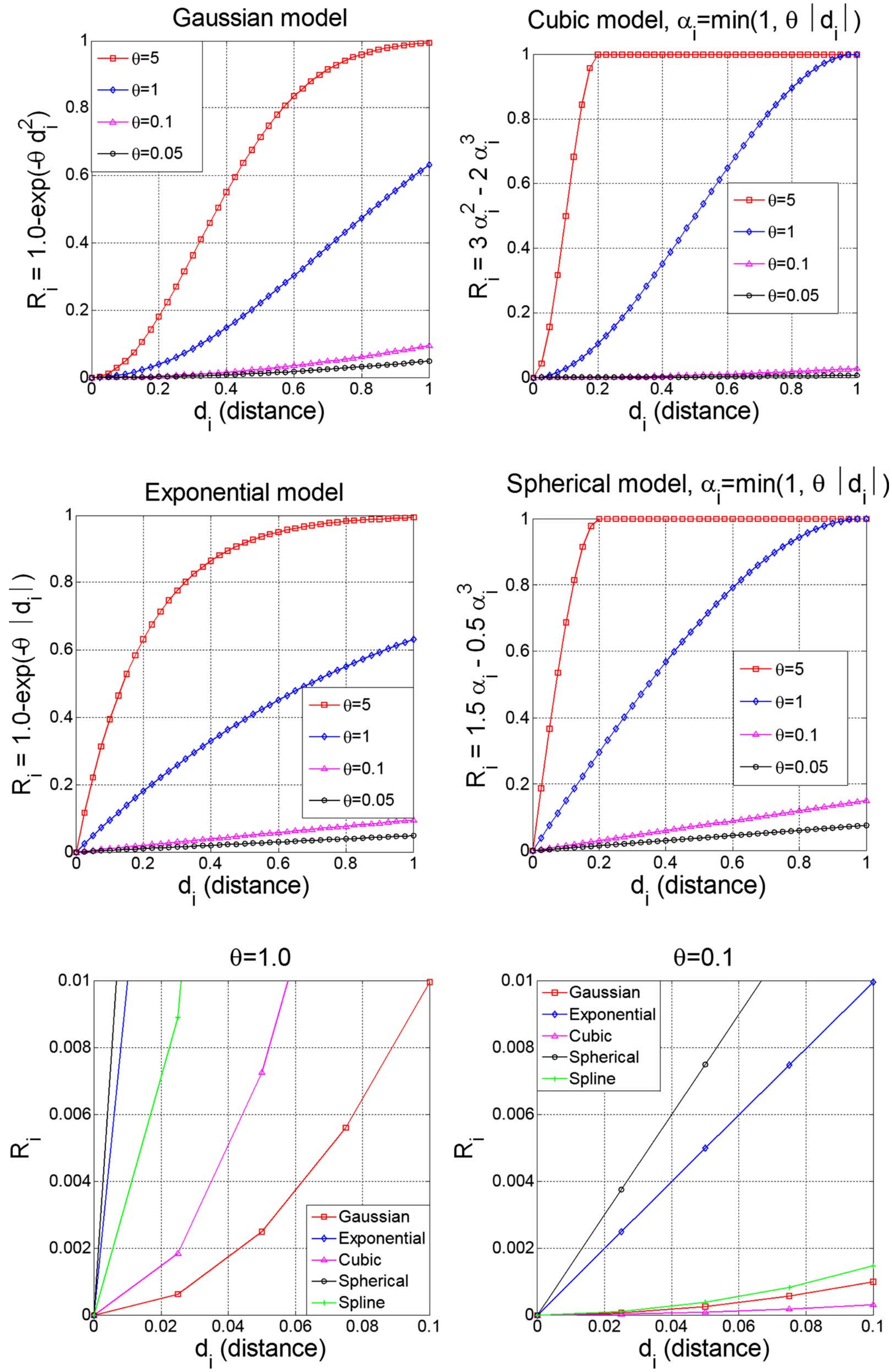

FIG. 2. (Color online) Common variogram models with the correlation parameter, θ. Gaussian, cubic, and spline models show a parabolic behavior near the origin.

this study, we have investigated several variogram models (e.g., Gaussian, cubic, spline, exponential, and spherical) in detail. Figure 2 shows the common variogram models with the correlation parameter, θ. As seen from Fig. 2, for all variogram models with decreasing values of θ the correlation increases (i.e., for a selected separation distance, γ decreases as θ decreases). More importantly, however, is the behavior of the variogram function near the origin (i.e., initial slope of the curve), which reflects the rate of decrease in the correlation. As reported by Lophaven *et al.*,[18] the variogram func-

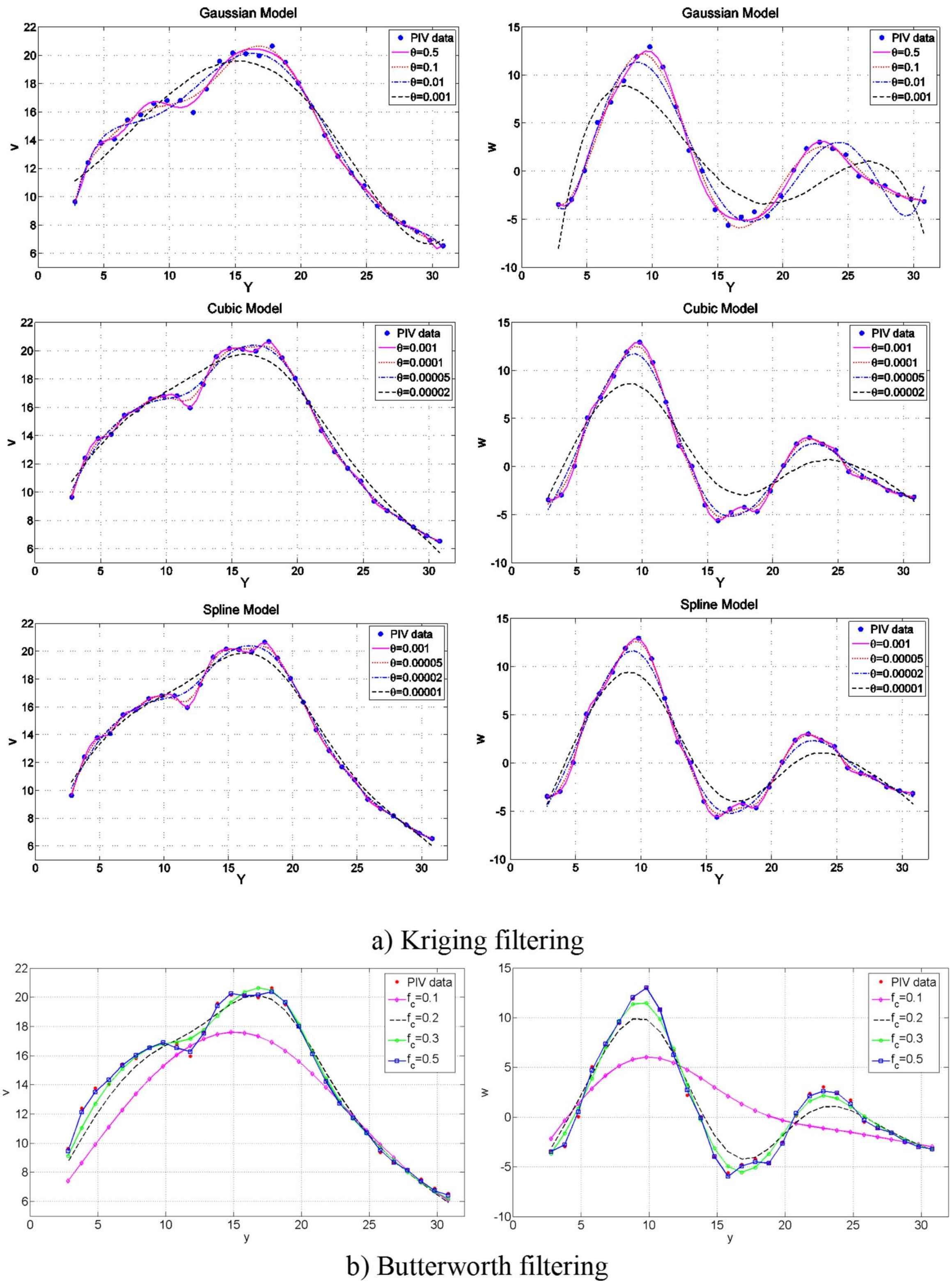

FIG. 3. (Color online) PIV data filtering using kriging with various variogram models and Butterworth IIR filtering. Data are extracted at $x=360$ plane ($z=100$ line).

tions that show parabolic behavior near the origin (e.g., Gaussian, cubic, and spline) are well correlated and can be used for smoothing of data. In fact, by reducing the correlation parameter to a sufficient value (specific for a given model), we are able to show that Gaussian, cubic, and spline models can *all* be used to smooth PIV data as shown in Fig. 3, where cross-flow (v) and spanwise velocity components (w) at the $z=100$ line ($x=360$ plane) are extracted for a detailed comparison. In addition, we also compare kriging smoothing with common low-pass digital filters [e.g., fast-Fourier-transform–based FIR filter and Butterworth IIR digital filter]. Note that the variogram parameter (θ) in kriging

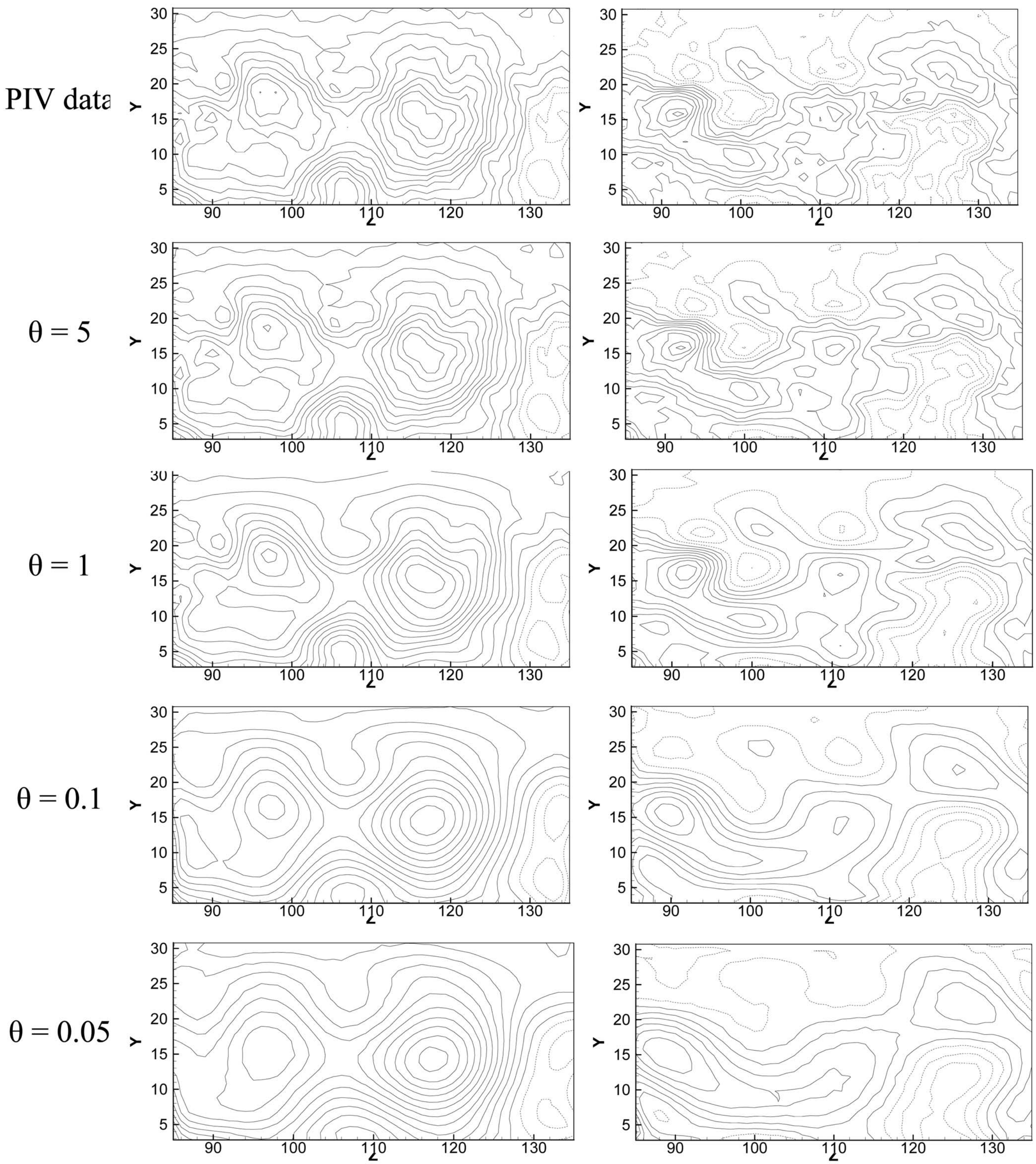

FIG. 4. The effect of Gaussian correlation parameter in smoothing/filtering of experimental data. Left column: the cross-flow velocity, right column: the spanwise velocity.

and the cutoff frequency (f_c) in various digital filters work similarly, i.e., the smoothing increases as θ or f_c is decreased (see Fig. 3). The data in Fig. 3 indicate that kriging follows the original data points more closely than the digital filter, i.e., it is more adaptive to local gradients, while the filter seems to average these out. The latter effect will also occur when the smoothing parameter of kriging is too small. Thus, each subfigure shows the acceptable limits for this parameter for each method.

Using either exponential or spherical variogram models it was not possible to smooth the data as these have a too sharp linear increase in γ (and hence a too fast drop off of correlation) near the origin as clearly depicted in Fig. 2. However, such "linear" variogram models can be used for

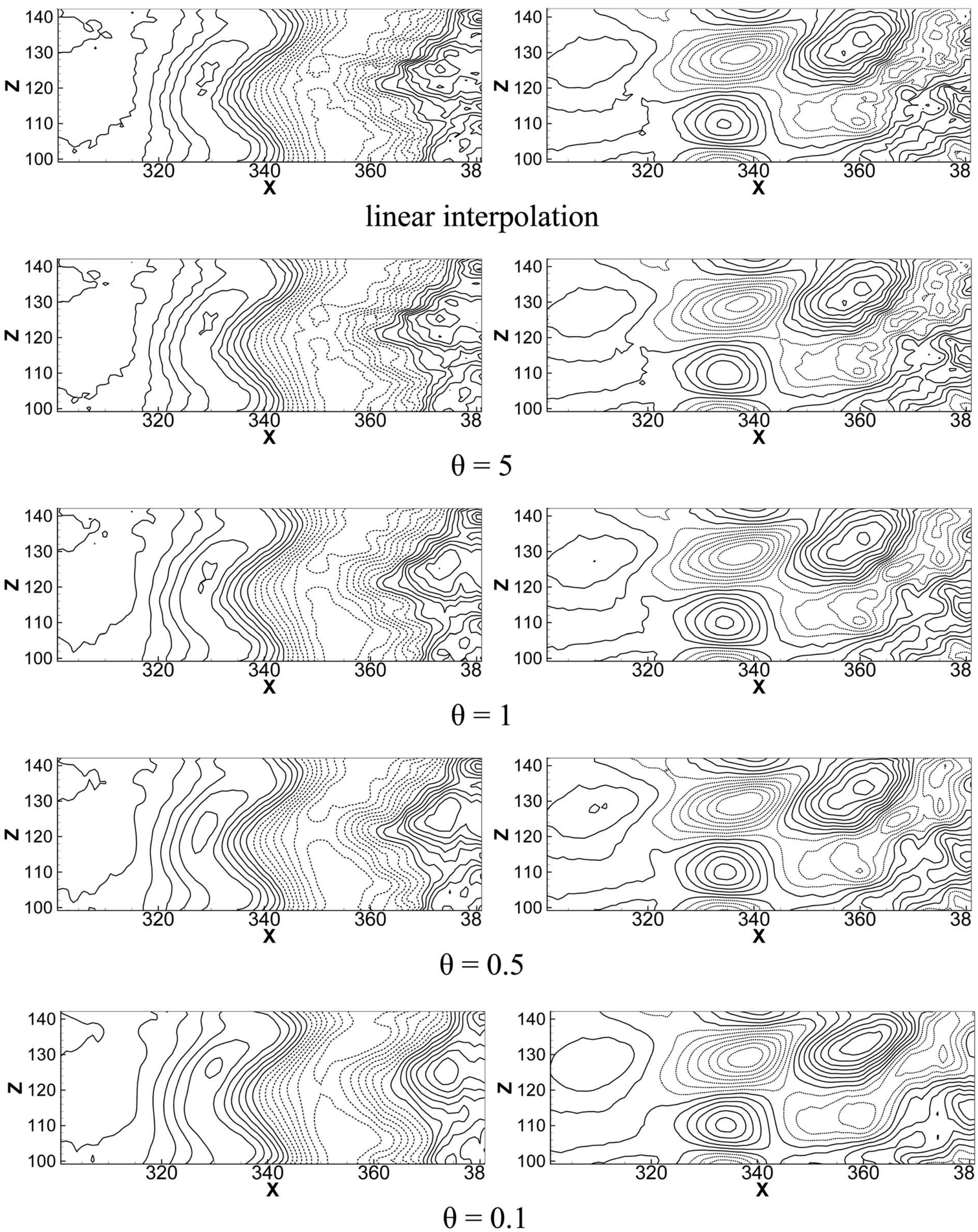

FIG. 5. The effect of the Gaussian correlation parameter in estimating experimental data at $y=15$ plane (new plane). Left column: cross-flow velocity, right column: the spanwise velocity.

the interpolation of exact data, i.e., "free" from measurement errors and noise (e.g., numerical data or highly accurate point measurements such as laser Doppler anemometry). For most of the measurements including the present PIV data, a so-called "smoothing" variogram model (e.g., Gaussian, cubic, or spline) shall be used for a certain level of smoothing. In the remaining of the paper, we present results based on the *Gaussian* variogram model as similar smoothing results are obtained for cubic and spline models as long as the variogram parameter is chosen suitably.

Figure 4 shows typical contours of our stereo-PIV data set ($x=360$) and smoothing using the Gaussian variogram

TABLE II. Divergence of the 3D normalized velocity field for original and smoothed PIV data.

		$\nabla . \vec{V}$
PIV data	⋯	0.0528
	Correlation parameter, θ	⋯
Kriging smoothing (Gaussian variogram model)	5	0.0283
	1	0.0280
	0.5	0.0275
	0.1	0.0238
	0.05	0.0254
	0.01	0.0224
	0.001	0.0183
	0.0005	0.0209

model with different correlation parameters (note that these are much higher than the lower limit observed in Fig. 3). The left column presents the cross-flow velocity, while the right column shows the spanwise velocity component. In order to obtain 3D smooth/enhanced data, we perform kriging smoothing/interpolation for all constant x-planes separately as described in Ref. 14. Then, the estimated velocity components at an (unknown) plane $y=15$ are shown in Fig. 5 as a function of the Gaussian variogram parameter, θ. Note that large values of θ corresponds to linear interpolation while gradual smoothing is obtained with decreasing θ. Note that we have concentrated on kriging of the cross-flow (v and w components of the velocity) only. The reason is that these components have finer structures (compared to u) and are therefore more difficult to interpolate, estimate, or smooth. Since the flow under consideration is incompressible, the ve-

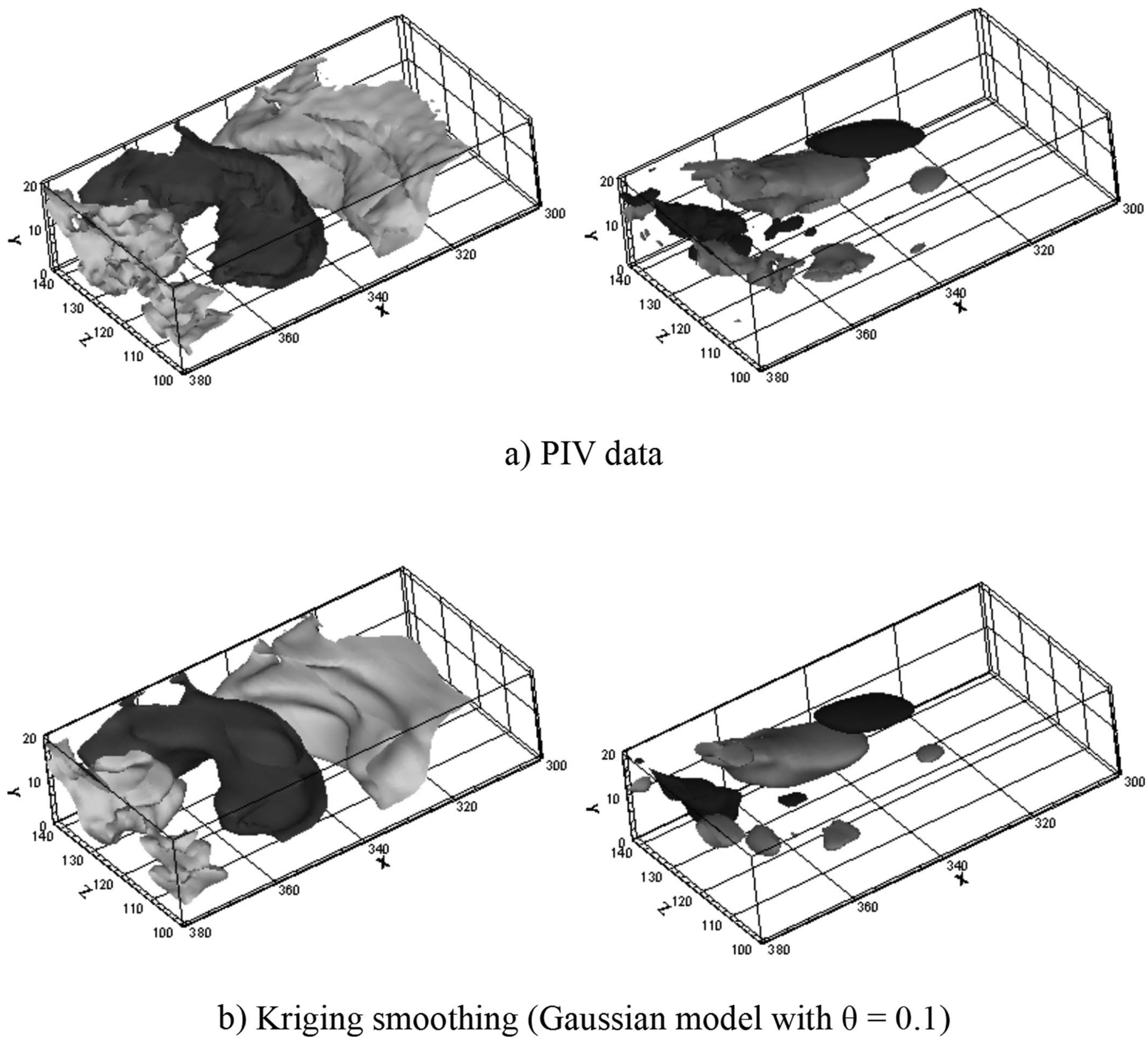

FIG. 6. Isocontours of cross-flow (left column) and spanwise (right column) normalized velocity at oscillation phase 180°. Dark gray v, $w=-0.11$, light gray v, $w=0.11$.

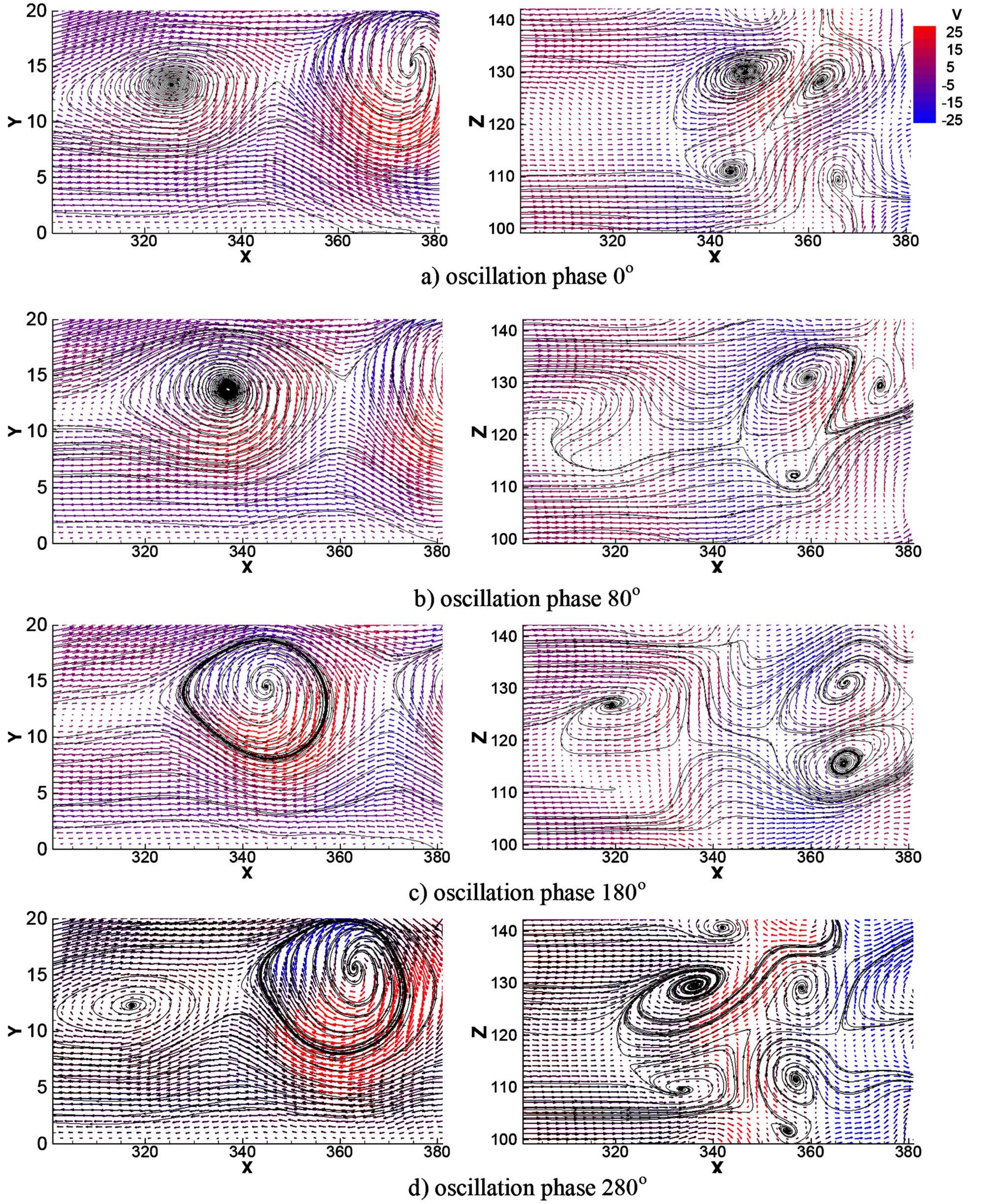

FIG. 7. (Color online) Time evolution of the streamwise moving vortical structures of stereo-PIV data obtained by kriging. Left column: streamwise plane ($z=130$ mm), the vector color/shade denotes the out-of-plane component w (spanwise component). Instantaneous stream lines in a frame of reference that moves with the fluid are added for the purpose of illustration. Right column: plane $y=15$ mm parallel to the flat plate. Vector color/shade denotes the out-of-plane component v (cross-flow component). Velocities are in mm/s.

locity vector field should theoretically satisfy the divergence-free condition (solenoidal field) but this constraint cannot generally be imposed on flow measurements.

In Table II, we present divergence of the 3D normalized flow field for original as well as smoothed PIV data. It is seen that for all smoothing cases, the incompressibility/divergence-free constraint is generally improved by a factor of 2 even though we did not impose it as part of kriging

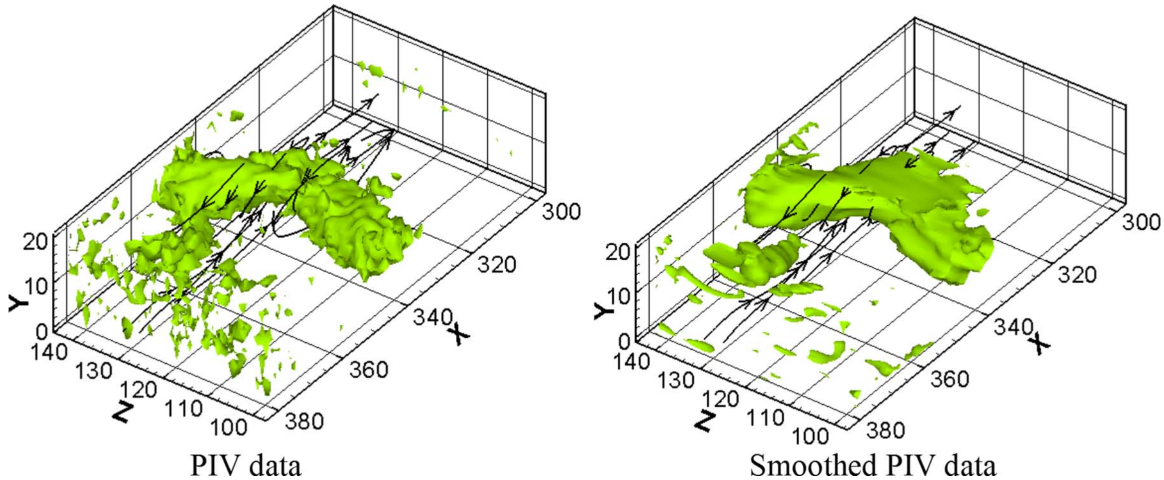

FIG. 8. (Color online) Vortices extracted using λ_2-method for original and smoothed PIV data at oscillation phase 180° ($\lambda_2=-18$).

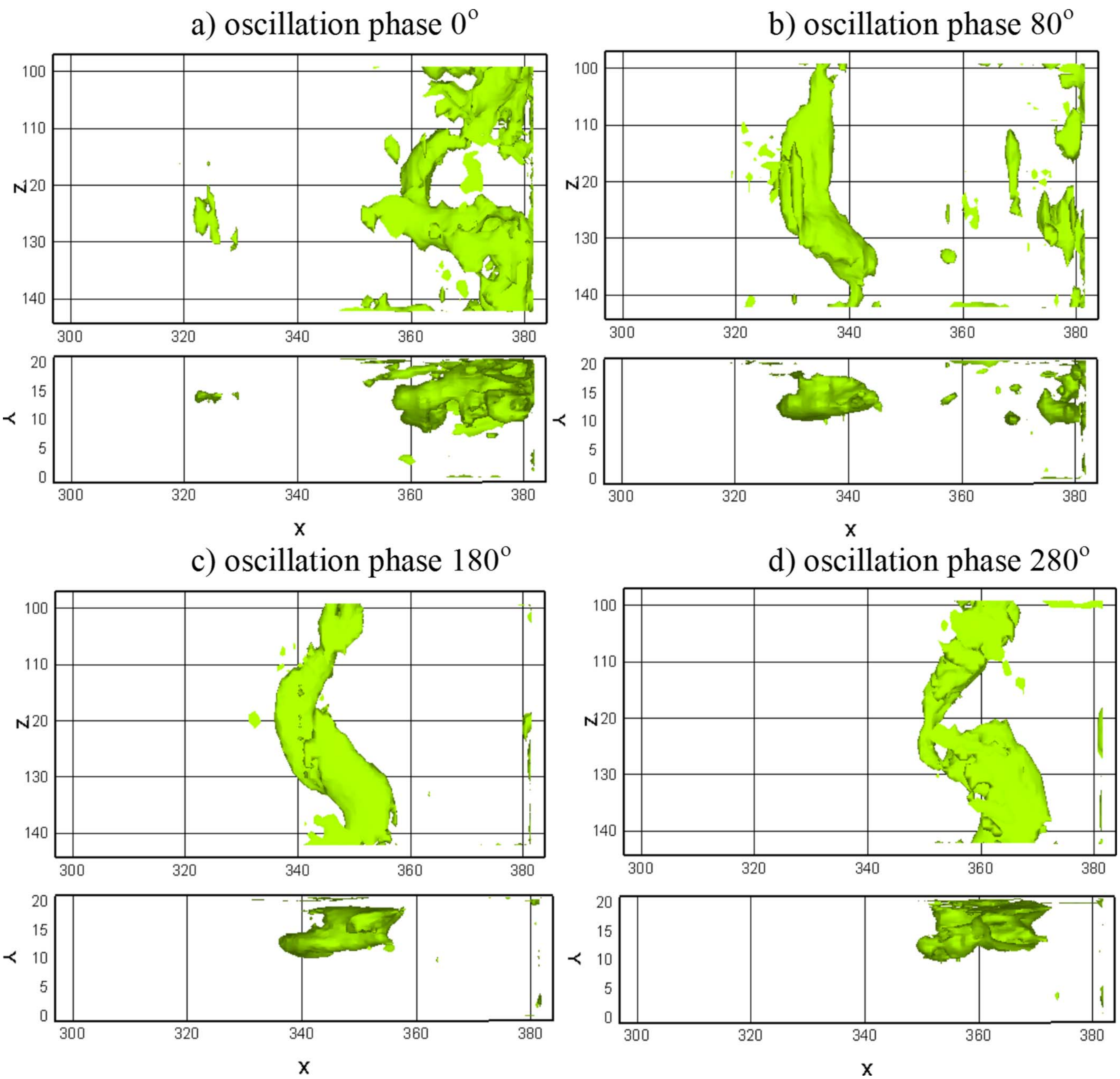

FIG. 9. (Color online) Time evolution of vortices extracted using λ_2-method for smoothed PIV data. ($\lambda_2=-22$). (Gaussian model with $\theta=0.1$.)

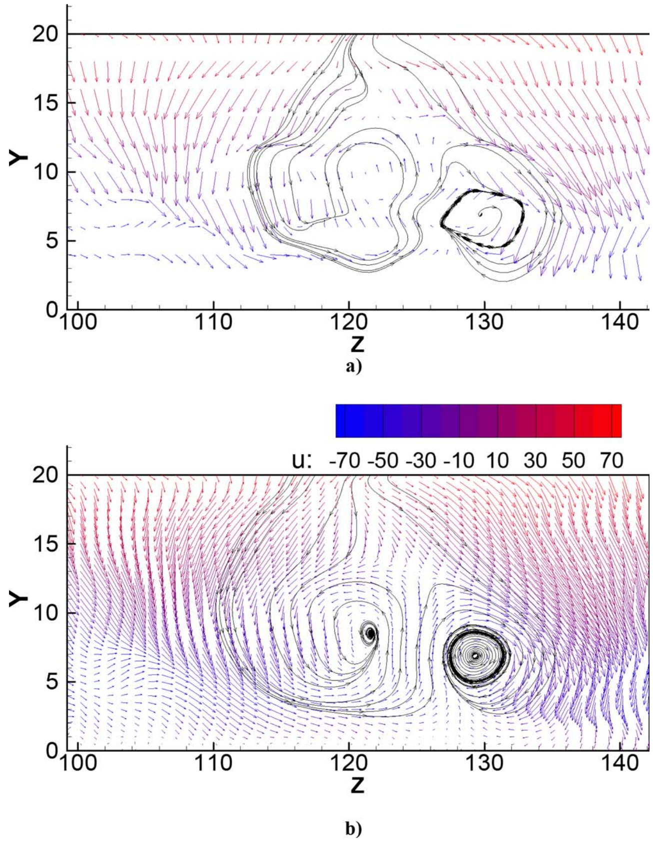

FIG. 10. (Color online) (a) Original stereoscopic PIV measurement in the y-z plane ($x=360$ mm). (b) Enhancement of the stereoscopic PIV data by kriging interpolation. Note that the instantaneous counter-rotating vortex pairs with the strong back flow in the transition region are accurately captured [see Lang *et al.* (Ref. 3)]. The vector color/shade shows the streamwise component of the velocity.

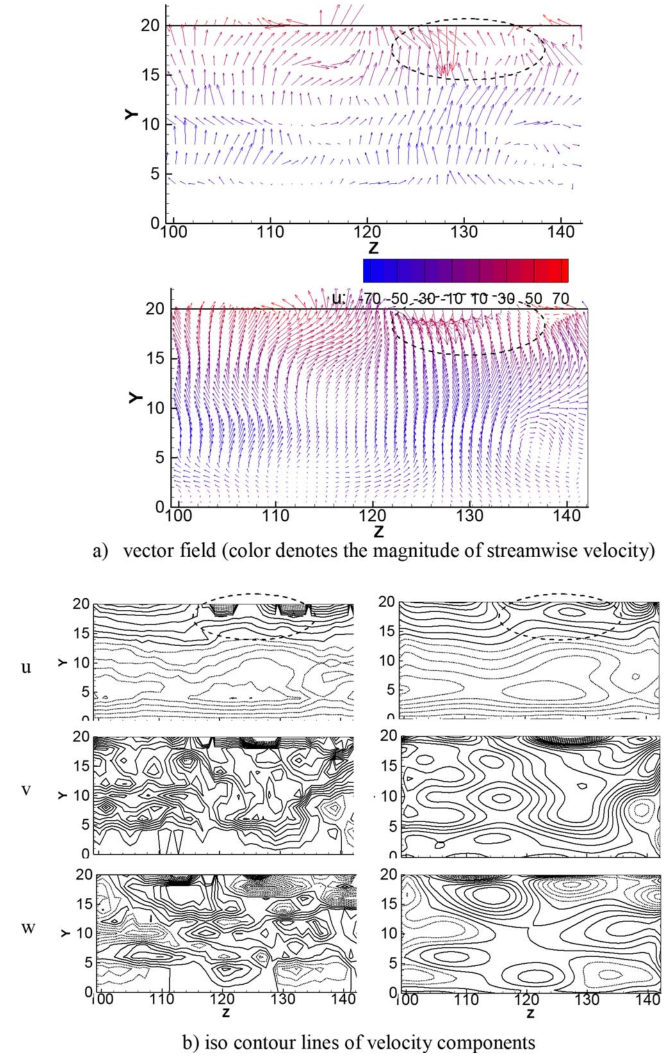

FIG. 11. (Color online) Enhancing/smoothing of experimental data containing outliers ($x=381$ plane).

implementation. As the correlation parameter decreases, divergence values decrease and the data are gradually more smoothed. However, there is a limit for the lowest useful correlation parameter. The divergence starts to oscillate when θ falls below 0.1. This observation correlates very well with the investigations shown in Fig. 3, where increasingly large deviations from the given data point occur for lower correlation parameters as well. If the correlation parameter is too low we seem to have a problem of "overcorrelation" of the data. For example, for $\theta=0.05$, unexpected wiggles are observable in the data in Fig. 4.

Figure 6 shows two representative isocontours of crossflow (left column) and spanwise velocity (right column) for original and smoothed PIV data. Negative and positive v in these illustrations correspond to the downward and upward motions, respectively, at the downstream and upstream ends of a spanwise-oriented vortex that is generated by roll up of the separated boundary layer. This latter motion is illustrated in Fig. 7. A reference speed of 80 mm/s has been subtracted from the streamwise velocity component in order to visualize it in a qualitative manner in a frame of reference that approximately moves with the flow. The left column in Fig. 7 shows streamwise cuts with instantaneous streamlines as an additional visual clue to illustrate the qualitative structure. Figure 6 and the right column of Fig. 7 demonstrate that the according structures have a crescentlike or zigzag shape which becomes more pronounced as the vortices travel downstream.

Thus, the structures get "pushed forward" at $z=140$ and 100 mm compared to $z=120$ mm. This leads to an increase in spanwise gradients and amplification of the spanwise velocity component w, as can be seen in the right column of Fig. 6, especially for $z>130$ mm. In order extract the dominant vortices and to illustrate their 3D structure, we use the λ_2-method of Jeong and Hussain.[19] According illustrations will be shown in Figs. 8 and 9. Since this method relies on the velocity gradient tensor, i.e., the spatial derivatives of the velocity components, it is necessary to use smoothed data instead of raw data for it. The difference is well illustrated in Fig. 8 by a comparison of the structures obtained with and without kriging. Some exemplary streamlines are also shown in order to relate the 3D structure to Fig. 7.

The dominant structure is either V- or Λ-shaped, depending on the direction of view. Traditionally, if the tip of such a structure points into downstream direction it is called a Λ-vortex. The opposite structure should then be called a V-vortex. Recently, such vortical structures have been ob-

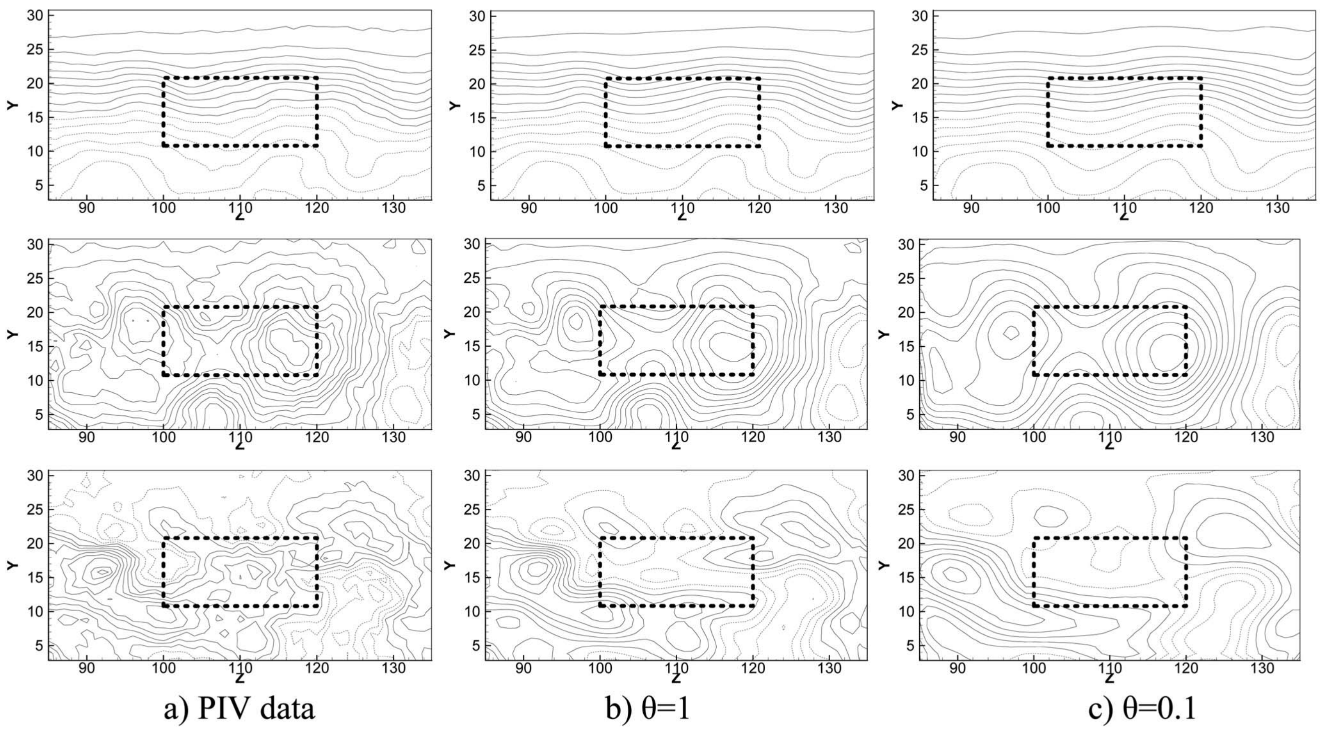

FIG. 12. Kriging estimation of black zone for PIV data at plane $x=360$. (a) PIV data (with missing data in a rectangular zone shown). Estimation of missing zone data using Gaussian variogram model with (b) $\theta=1$ and (c) $\theta=0.1$. From top to bottom: contours of the streamwise, cross-flow, and spanwise velocity components.

served by Burgmann *et al.*[5] for a transitional separation bubble of an airfoil. They used the term "C-shaped vortex"[20] which agrees with the rather obtuse angle of the V's in the present visualizations. However, note that the complete structure is of zigzag shape, i.e., a spanwise sequence of V- and Λ-shaped vortices. These types of vortices are common in transitional boundary layer flows as reported by Meyer *et al.*[21]

In the following figure we shall try to convey the unsteady development of the dominant vortices using vortex extraction and visualization output (λ_2 isosurfaces) of the smoothed 3D data sets. The isovalue ($\lambda_2=-22$) has been selected such that only the vortex centers appear. This can be controlled by comparing the side views in Fig. 9 with the streamlines in Fig. 7. Thus, the initial vortex at $x \approx 325$ mm in Fig. 9(a) is just about to cross the chosen threshold. As it becomes stronger in Fig. 9(b) it develops into a C-shaped structure when viewed from above. At its lower part, i.e., at $z>130$ mm where the spanwise velocity in Fig. 6 was already observed to be strongest, we may detect the left part of a Λ-vortex (whose right part is outside the measurement region). This Λ-shaped vortex becomes more acute and stretched into streamwise direction in the course of time depicted through Figs. 9(c), 9(d), and 9(a) (recall the periodicity of the data because of periodic forcing and phase averaging).

The breakdown process at $x>360$ mm is rather complex and consists of simultaneous mixing in spanwise and wall-normal directions with typical length scales around 10–20 mm, which are still too difficult to faithfully extract them from the existing data. Nevertheless, one such structure

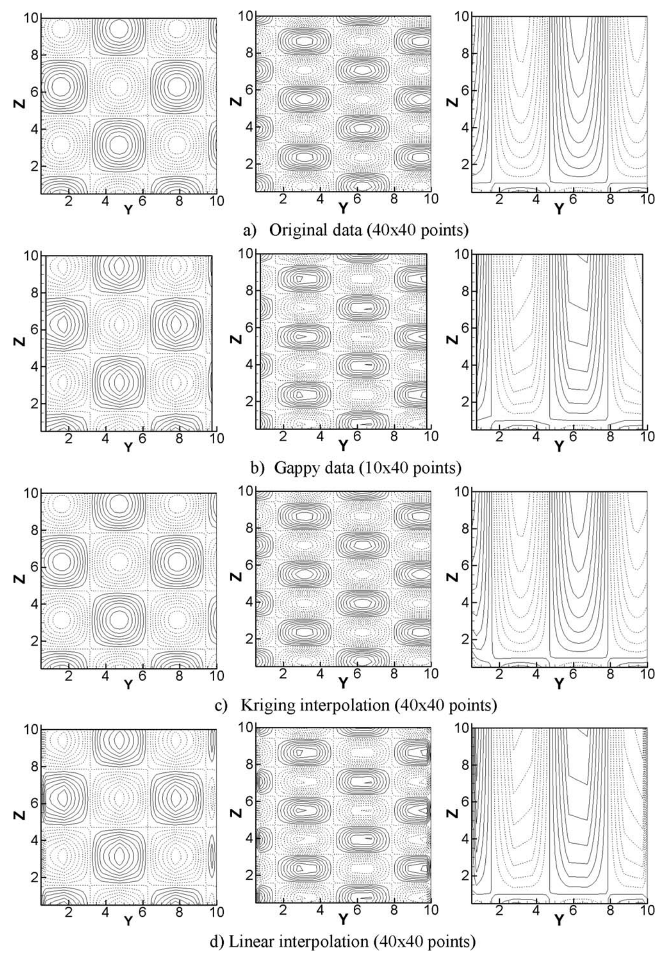

FIG. 13. Reconstruction of periodic functions using kriging and linear interpolations. Left column: $f_1(y,z)=\sin(y)\cos(z)$, middle column: $f_2(y,z)=\cos(y)\sin(2z)$, and right column: $f_3(y,z)=\cos(y)\ln(z)$.

TABLE III. Comparison of kriging and linear interpolation for reconstructing missing lines.

rms error	Kriging interpolation	Linear interpolation
$f_1(y,z)$	3.58×10^{-2}	1.95×10^{-1}
$f_2(y,z)$	6.65×10^{-3}	3.51×10^{-1}
$f_3(y,z)$	0.9×10^{-1}	3.53×10^{-1}

is shown in Fig. 10 in a cross-stream (spanwise) cut. The raw data and the interpolated data are shown for comparison. A streamwise vortex pair appears as $z=120-130$ mm under (i.e., closer to the wall than) the dominant spanwise zigzag structure. We also show it here to prove its existence in addition to Ref. 3. Unfortunately, more and finer-resolved measurements are needed to fully clarify its spatial structure and its connection to the structures further away from the wall.

V. TREATMENT OF OUTLIERS

It is known that in PIV measurements, built-in data processing software often fails to detect or might be unable to correct the spurious erroneous vectors called outliers. On the other hand, through simple examples, Armstrong and Boufassa,[22] for example, compared the robustness of ordinary and log-normal kriging with respect to the outlier resistance. They found that although log-normal kriging is more resistant to outliers, its estimate is very sensitive to slight changes in the sill of the variogram of the logs. In addition, log-normal kriging is suitable for log-normally distributed data so we report here the result of ordinary kriging with a Gaussian variogram model.

As the stereo-PIV data is phase averaged, in most of the flow field the outliers are minimal or nonexistent. For this reason, in order to investigate the treatment of outliers using kriging specifically, we have selected the very downstream

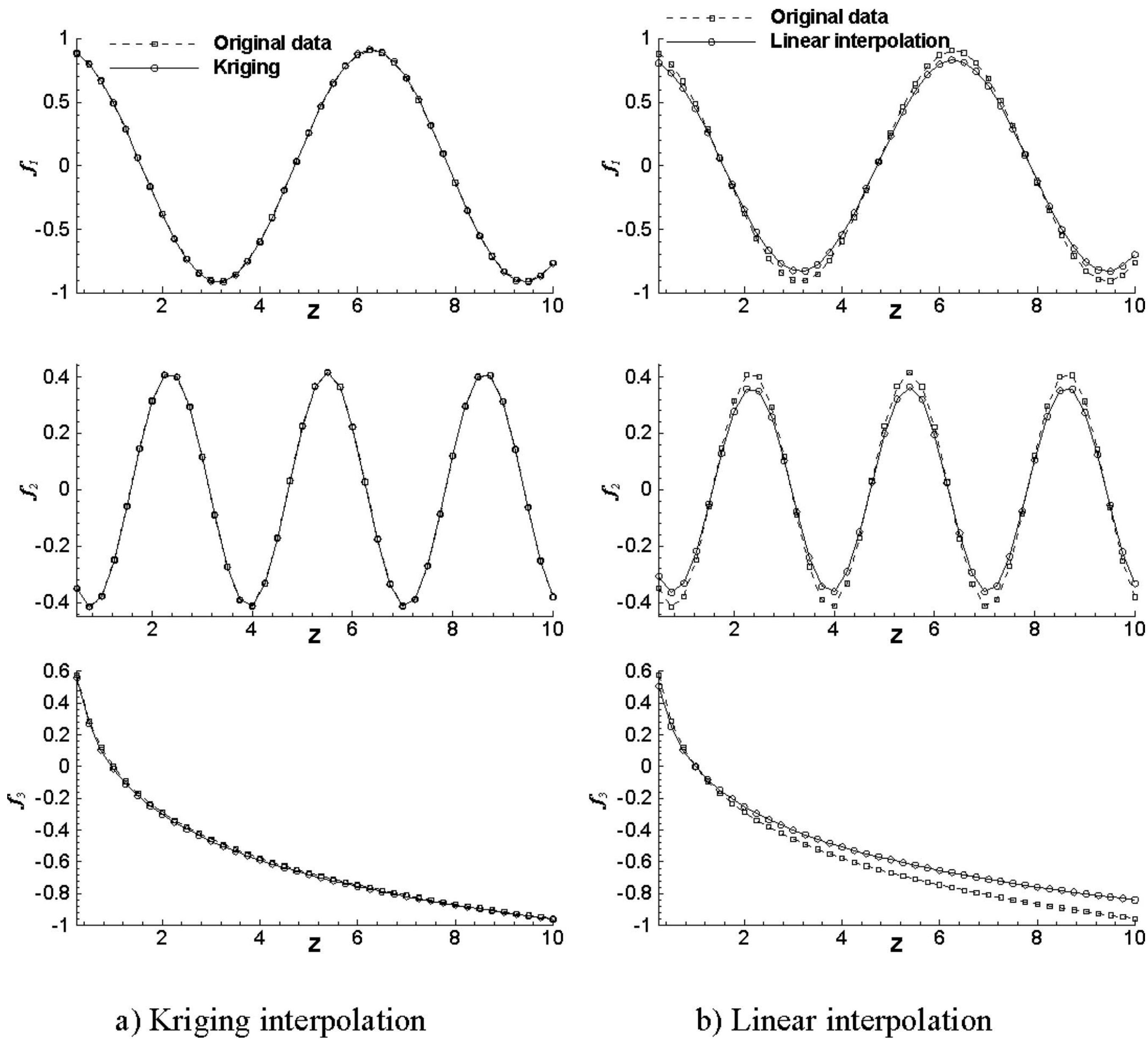

FIG. 14. Comparison of kriging and linear interpolations for selected periodic functions at missing line $y=2$.

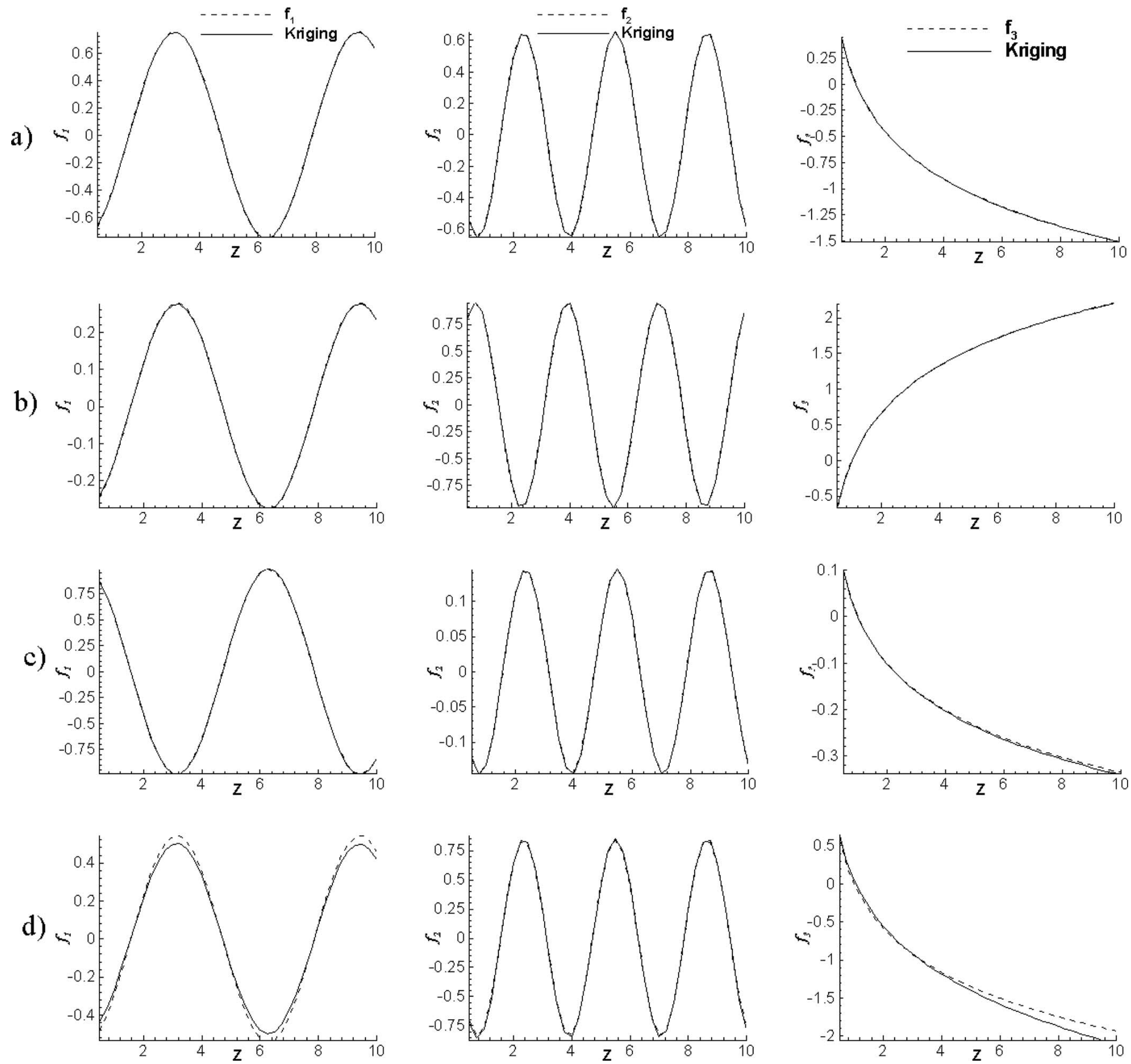

FIG. 15. The performance of kriging interpolation at different "missing" y-locations: (a) $y=4$, (b) $y=6$, (c) $y=8$, and (d) $y=10$.

part of the separation bubble ($x=381$, i.e., the last measured plane, see Fig. 1) which contains noticeable outliers as shown in Fig. 11. It is clearly seen from the vector field as well as the isocontour lines of velocity components that outliers can be effectively eliminated or greatly reduced using kriging. We also mention here that POD, which employs the spatiotemporal data, is also very resistant to outliers as reported in Ref. 23.

VI. RECONSTRUCTION OF LARGE MISSING DATA CLUSTERS (BLACK ZONES)

In this section, we extend our investigation on the capability of kriging for reconstructing large missing data clusters or black zones. Kriging has been shown recently to be an effective data recovery tool for fluid flow problems with *smooth* data, obtained from a two-dimensional DNS flow past a circular cylinder (Refs. 7 and 9). However, the most common reality is that these black zones usually occur in experimental data (e.g., PIV) for several reasons mentioned in Sec. I. In addition, in the experimental investigations, the black zones may usually occur for all snapshots in a measurement campaign, so it is not ordinarily possible to apply the gappy POD to estimate missing data. Kriging is not however restricted by this problem since it uses only spatial correlation. In order to gain experience with the kriging reconstruction for *nonsmooth* data, we apply it here to predict large missing data clusters of *noisy* experimental data.

Figure 12 shows PIV data (left column) and its kriging reconstructions (middle column, $\theta=1$; right column, $\theta=0.1$) for streamwise, cross-flow, and spanwise components of velocity, from top to bottom, respectively. The rectangular missing data clusters indicated by dashed lines are actually omitted from the PIV data, but shown in Fig. 12 for comparison with the kriging reconstruction. First of all, we performed kriging with a Gaussian variogram model, which resulted in a "smooth" reconstruction as expected but this lead to interface problems as the known "outer" data were noisy. To eliminate this problem, before reconstruction of the black zone, we first smooth the outer data and then carry out the reconstruction of the black zone as shown in Fig. 12. Then, it is seen that the interface problem is eliminated and a complete, smooth data sheet is obtained. Another important point is that, while the streamwise component and the cross-flow component of the velocity vector are reconstructed faithfully, kriging interpolation fails to accurately reconstruct the spanwise component. This may be attributed to a very complicated, low correlation spatial distribution of the spanwise velocity component and a complete loss of information through the black zone.

VII. CONCLUSIONS

We have shown that kriging can be effectively used for data smoothing and spatial enhancement of stereo-PIV data for a transitional boundary layer with laminar separation bubble. In addition, outliers (unrealistic vectors) that are often observed in PIV measurements can be eliminated and/or greatly reduced. Another common problem in PIV measurements is that for a certain number of reasons (e.g., shadowing, insufficient illumination, obstructed view, nonhomogeneous seeding concentrations, etc.) information may be missing in clusters (or black zones). Kriging can be used to estimate these black zones when the data are well correlated.

We have demonstrated that a meaningful and unambiguous selection of variogram model and related correlation parameter is possible. Both are crucial for estimation: When the values of the designed (available) data set are not exact, i.e., the source points have some uncertainty or contain background noise as in experimental data, smoothing variogram models such as Gaussian, cubic, and spline that show parabolic behavior near the origin shall be used. On the other hand, linear variogram models (exponential and spherical) can be used for interpolation of "exact" data sets such as those obtained from numerical data.

An important aspect of PIV data smoothing/interpolation is that the resulting velocity fields should obey the fundamental laws that govern the fluid dynamics. At present, work is underway to incorporate the incompressibility/divergence-free constraint as part of our kriging implementation.

ACKNOWLEDGMENTS

We gratefully acknowledge the financial support of Scientific and Technical Research Council of Turkey (TÜBİTAK) and International Bureau of the German Federal Ministry of Education and Research (IB-BMBF). The authors would like to thank Dr. M. Lang for providing his latest stereo-PIV data. We also like to thank Dipl.-Ing. K. Baysal for his help in implementing the λ_2-method.

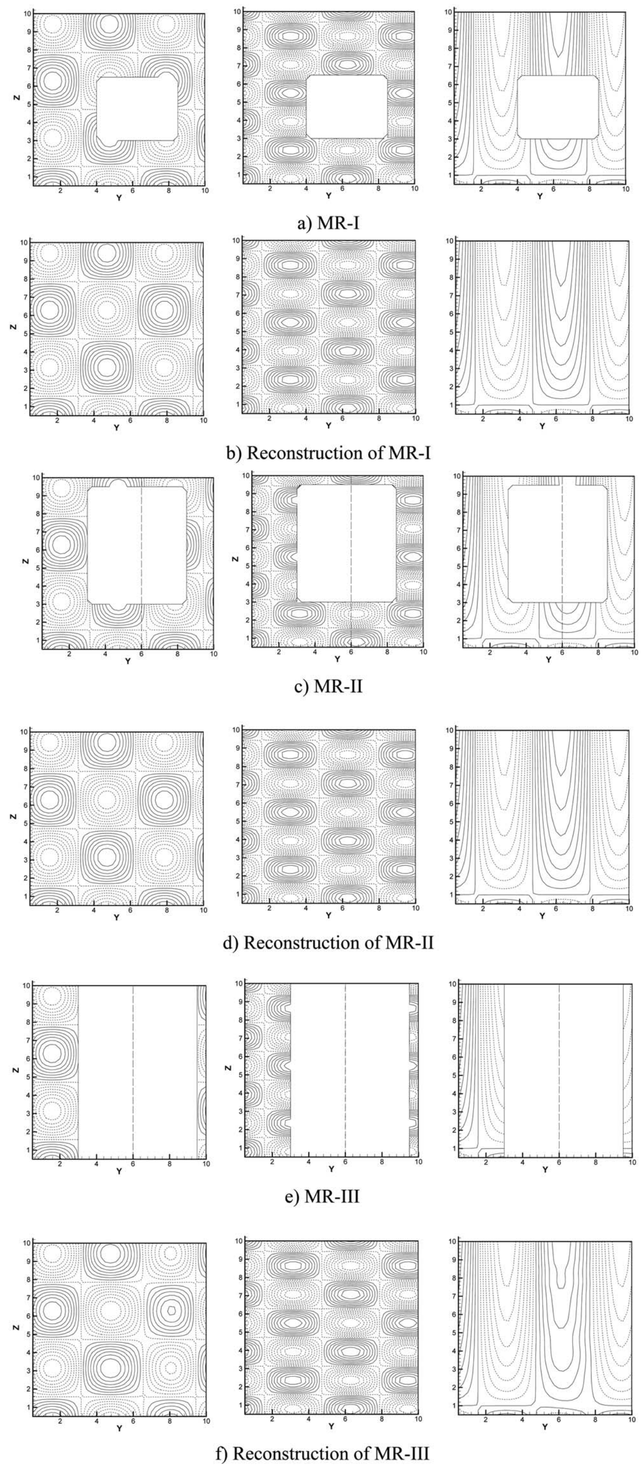

FIG. 16. Reconstruction of MRs (black zones) using kriging interpolation. Left column: $f_1(y,z)=\sin(y)\cos(z)$, middle column: $f_2(y,z)=\cos(y)\sin(2z)$, and right column: $f_3(y,z)=\cos(y)\ln(z)$.

APPENDIX A: RECONSTRUCTION OF PERIODIC FUNCTIONS

In this appendix, we investigate the kriging interpolation in detail for selected analytical functions in order to acquaint the reader further with the procedure. We consider both periodic and nonperiodic known functions to create certain data sets and study the performance of kriging to predict the func-

TABLE IV. The rms errors of reconstruction of the MRs for selected test functions (see Fig. 16).

rms error	MR-I	MR-II	MR-III
$f_1(y,z)$	8.64×10^{-7}	2.50×10^{-4}	1.84×10^{-1}
$f_2(y,z)$	2.23×10^{-5}	1.43×10^{-4}	1.53×10^{-2}
$f_3(y,z)$	5.86×10^{-6}	4.65×10^{-5}	1.73×10^{-1}

tion values at unknown design points. Since these data sets are created by known functions, these functions are independently evaluated at unknown design points in order to compare with kriging interpolation. In addition, kriging interpolation is compared with linear interpolation.

The specific periodic functions we consider are as follows:

$$f_1(y,z) = \sin(y)\cos(z), \tag{A1}$$

$$f_2(y,z) = \cos(y)\sin(2z), \tag{A2}$$

$$f_3(y,z) = \cos(y)\ln(z). \tag{A3}$$

First of all, we generate measurement data using Eqs. (A1)–(A3) on a domain $y=[0,10]$ and $z=[0,10]$ using an equally distributed grid of (40×40) points $(dy=dz=0.25)$. Next, in order to create a "gappy data set" we discard the values of constant y-lines so that we obtain a much coarser but still equally distributed grid of (10×40) points (with $dy=1$, and $dz=0.25$). Figure 13 shows the reconstruction of coarser gappy data by kriging interpolation. Actual contours of functions obtained by Eqs. (A1)–(A3) on a (40×40) grid points are also shown for comparison. Table III shows the root mean square (rms) error of the reconstruction for the three functions. We note that using a *relative error* is not particularly suitable here because at some (but only few) points the field values are so small that small deviations give very large relative errors, which make a reasonable comparison impossible. In order to normalize with respect to the "variance" of the field, we evaluate the variance as

$$\sqrt{\frac{1}{N}\sum_{i=1}^{N}[f_k(y,z)-\bar{f}_k]^2}, \tag{A4}$$

where $\bar{f}_k$ is the average field and $k=1$, 2, and 3, respectively.

Using the variance, the normalized rms error can be evaluated as follows:

$$\mathrm{rms}(f_k) = \frac{\sqrt{\frac{1}{N}\sum_{i=1}^{N}[f_{k,C}(y,z)-f_k(y,z)]^2}}{\sqrt{\frac{1}{N}\sum_{i=1}^{N}[f_k(y,z)-\bar{f}_k]^2}}, \tag{A5}$$

where $f_{k,C}$ is the reconstructed field using kriging or linear interpolation.

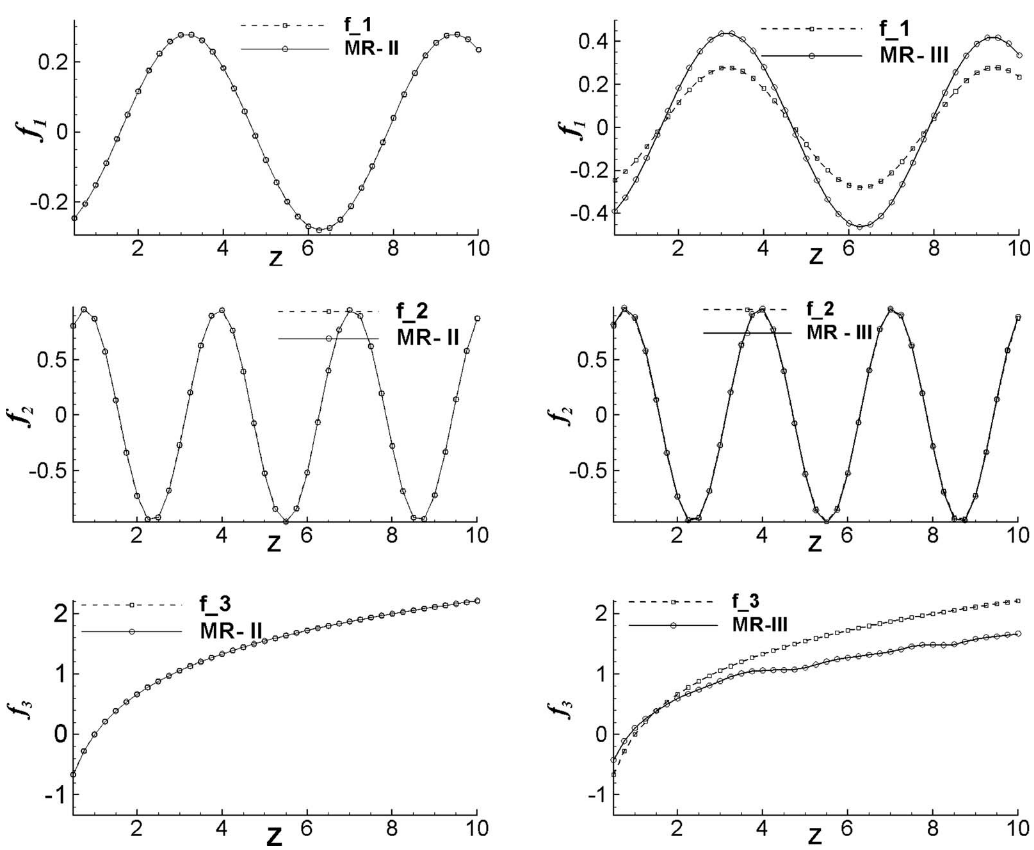

FIG. 17. The comparison of reconstructed data from MR-II and MR-III with that of the actual data at line $y=6$ [for left column see Fig. 16(c) and right column see Fig. 16(e)].

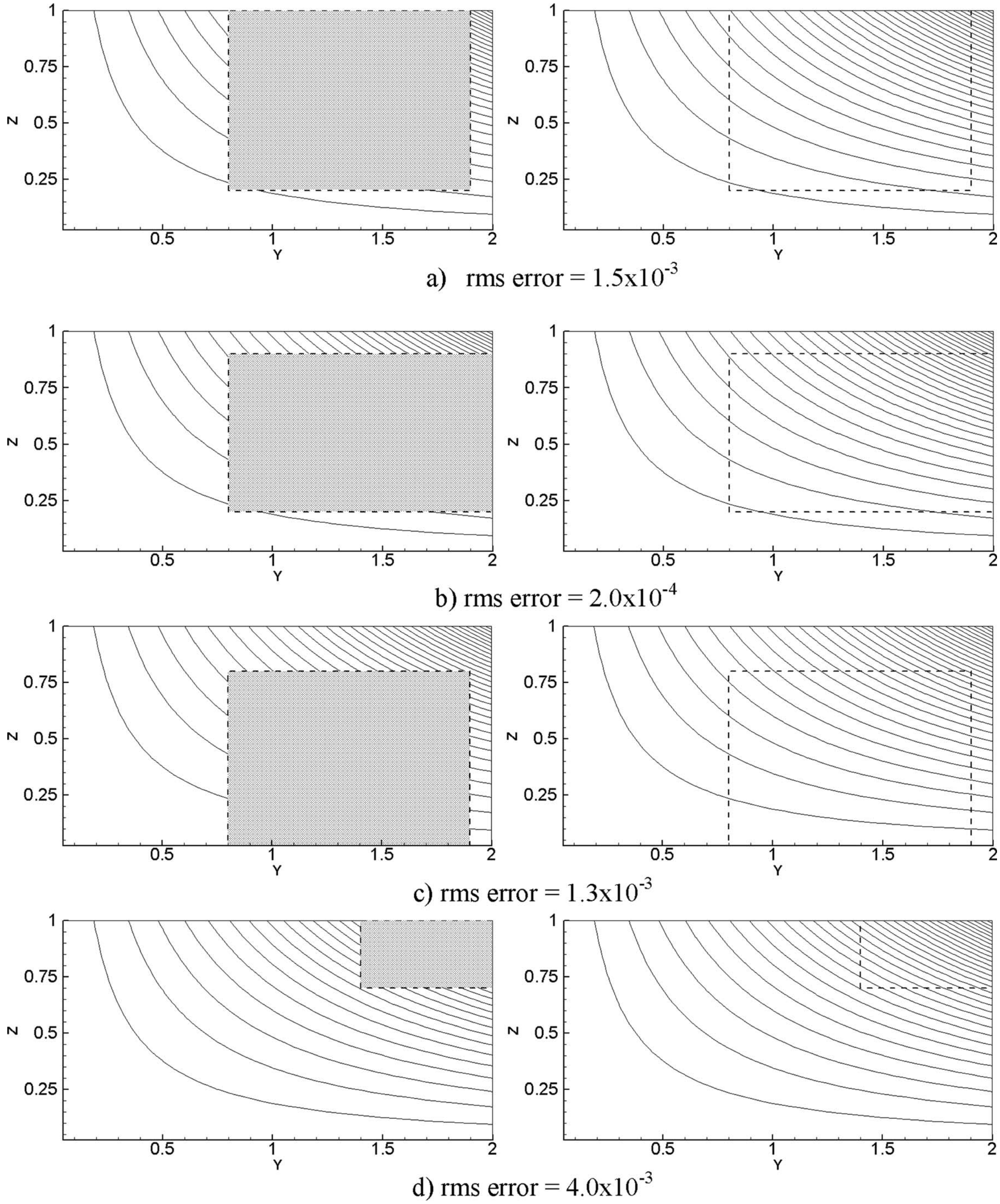

FIG. 18. Various MRs and their reconstruction by kriging interpolation. Left column: analytical solution with MR (Poisson equation). Right column: reconstructed MR by kriging interpolation.

It is seen that for all functions kriging interpolation is much more accurate than the linear interpolation, almost two orders of magnitude for periodic functions with smaller wavelength, like the function f_2 in our example. In order to see deviations from the original functions, we compare kriging and linear interpolation for selected periodic functions at the "missing line" $y=2$ (see Fig. 14). While kriging interpolation is virtually indistinguishable from the true data, for the linear interpolation considerable deviations from the actual data are observable especially for the maximum/minimum values of the functions. The prediction of kriging interpolation at other missing lines (shown in Fig. 15) has been extremely good. Note that at the bottom of Fig. 15, the missing line $y=10$ corresponds to the right domain boundary. Since the $y=10$ plane is outside of the data set on which the kriging model is constructed, we see some deviations from the actual functions. This and other numerical investigations we have performed show that the kriging method is not to be used for extrapolation since our numerical experiments show that the accuracy sharply deteriorates outside the design data set.

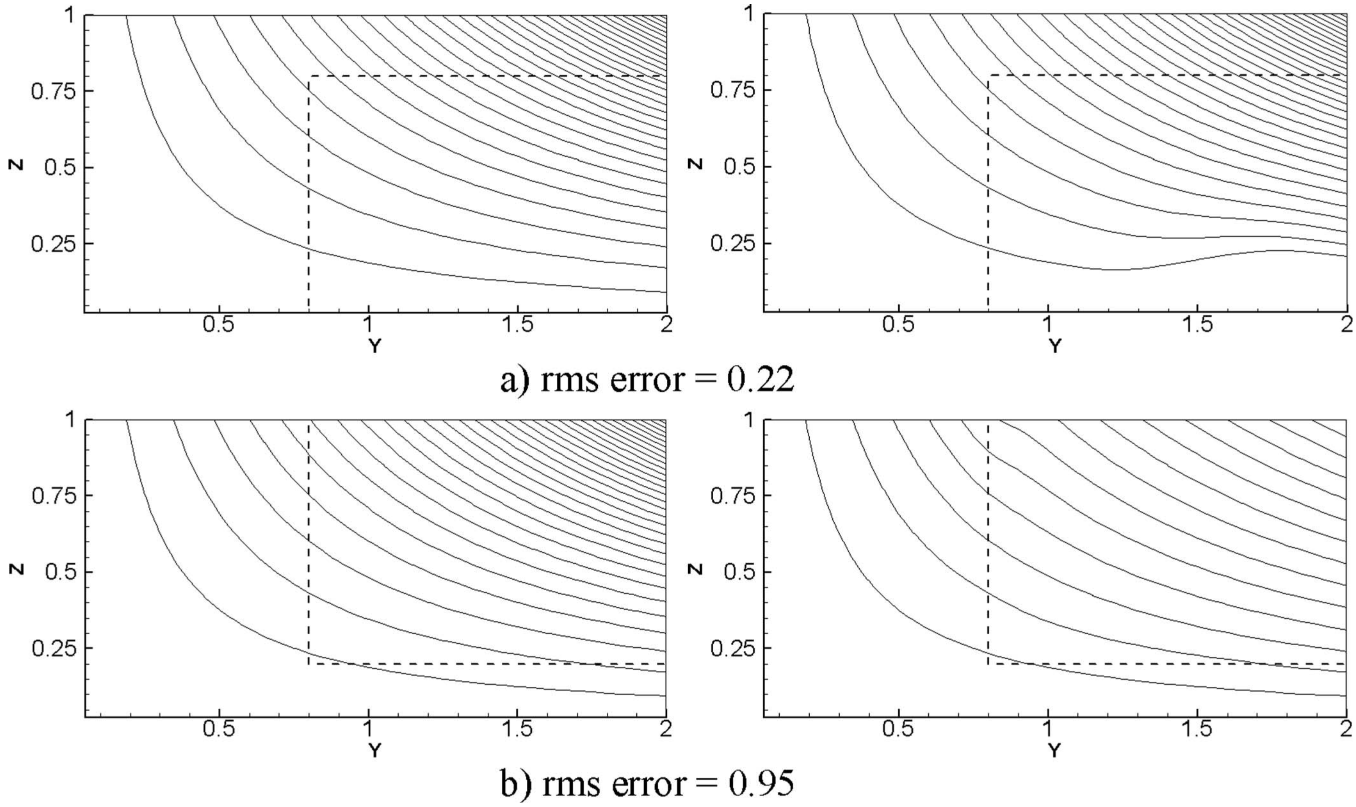

FIG. 19. Various MRs (shown as dotted rectangle) and their reconstruction by kriging interpolation. Left column: analytical solution (Poisson's equation). Right column: reconstructed MR by kriging interpolation.

APPENDIX B: RECONSTRUCTION OF LARGE MISSING DATA CLUSTERS (BLACK ZONES) BY KRIGING INTERPOLATION

In this appendix we simulate the case when large missing data clusters or black zones occur. For this, we created three missing regions (MR-I–MR-III) with varying rectangular black zone locations and sizes, as shown in Fig. 16 for all the three functions. The kriging model is constructed based on the known data set around the missing regions (MRs). Then, kriging interpolation is used to "fill" the MRs. Figure 16 compares the original data and results of kriging interpolation for the three different functions and black zones. The reconstruction (rms) errors [similarly defined based on the variance as in Eq. (A5)] of the MRs for all the three periodic functions are given in Table IV. For both MRs, MR-I and MR-II, we see that the reconstruction error is extremely small and all the functions are recovered very accurately as shown in Figs. 16(b) and 16(d). Figure 17 shows the comparison of reconstructed data from MR-II and MR-III with that of the actual data at line $y=6$. Note that line $y=6$ crosses the middle of the MR so that most of the deviations from the actual solution are expected along this line [see Figs. 16(c) and 16(e)]. We see that, while variations along this line are practically zero for MR-II, there are very large deviations for MR-III for functions f_1 and f_3. We should also note here that considering MR-III, it is crucial to keep a small portion of the data [on the right boundary of the domain as in Fig. 17(e)] so that MR-III constitutes an interpolation problem rather than an extrapolation one. In fact, by removing the "strip" of the data on the right boundary, we completely fail to reconstruct the MR-III by kriging.

APPENDIX C: RECONSTRUCTION OF NONPERIODIC FUNCTIONS

In this appendix we use kriging interpolation for a nonperiodic function. We consider the following elliptic partial differential equation:

$$\frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = (y^2 + z^2)e^{yz}, \quad 0 < y < 2, \quad 0 < z < 1 \qquad \text{(C1)}$$

subject to the boundary conditions,

$$u(0,z) = 1, \quad u(2,z) = e^{2z}, \quad 0 \le z \le 1,$$

$$u(y,0) = 1, \quad u(y,1) = e^{y}, \quad 0 \le y \le 2.$$

The analytical solution for the given problem is $u(y,z) = e^{yz}$. Figure 18 shows several MRs (black zones) and their reconstruction using kriging interpolation. Wee see that results obtained from kriging interpolation are in good agreement with the analytical solution for MRs shown in Fig. 18. Note that in Figs. 18(a)–18(c), the domains include an informative strip of data (whether vertically or horizontally). We have experimented with varying size and locations of the MR and we conclude that at least a strip of data with sharp gradients should be kept in the "left-over" data in order to reconstruct it correctly. Figure 19 shows two cases where kriging interpolation cannot reconstruct the analytical solution because the above stated condition is not fulfilled. Note that for these cases, the reconstruction (rms) error is several orders of magnitude higher than for the MRs shown in Fig. 18. For example, in Fig. 19(b) the gradient at $z=1$ is not captured by kriging interpolation since the most crucial information is removed by the MR for this case.

[1] Ch. Brücker, "Digital-particle-image-velocimetry (DPIV) in a scanning light-sheet: 3D starting flow around a short cylinder," Exp. Fluids **19**, 255 (1995).
[2] C. J. Kähler and J. Kompenhans, "Fundamentals of multiple plane stereo particle image velocimetry," Exp. Fluids **29**, 70 (2000).
[3] M. Lang, U. Rist, and S. Wagner, "Investigations on controlled transition development in a laminar separation bubble by means of LDA and PIV," Exp. Fluids **36**, 43 (2004).
[4] G. E. Elsinga, F. Scarano, B. Wieneke, and B. W. van Oudheusden, "Tomographic particle image velocimetry," Exp. Fluids **41**, 933 (2006).
[5] S. Burgmann, C. Brücker, and W. Schröder, "Scanning PIV measurements of a laminar separation bubble," Exp. Fluids **41**, 319 (2006).
[6] N. E. Murray and L. S. Ukeiley, "An application of gappy POD: For subsonic cavity flow PIV data," Exp. Fluids **42**, 79 (2007).
[7] D. Venturi and G. E. Karniadakis, "Gappy data and reconstruction procedures for flow past a cylinder," J. Fluid Mech. **519**, 315 (2004).
[8] K. Willcox, "Unsteady flow sensing and estimation via the gappy proper orthogonal decomposition," AIAA Paper No. 2004-2415, 2004.
[9] H. Gunes, S. Sirisup, and G. E. Karniadakis, "Gappy data: To krig or not to krig?" J. Comput. Phys. **212**, 358 (2006).
[10] J. C. Davis, *Statistics and Data Analysis in Geology*, 3rd ed. (Wiley, New York, 2002).
[11] E. H. Isaaks and R. M. Srivastava, *Introduction to Applied Geostatics* (Oxford University Press, New York, 1989).
[12] J. Sacks, W. J. Welch, T. J. Mitchell, and H. P. Wynn, "Design and analysis of computer experiments," Stat. Sci. **4**, 409 (1989).
[13] M. Stein, *Interpolation of Spatial Data: Some Theory for Kriging* (Springer-Verlag, Berlin, 1999).
[14] H. Gunes and U. Rist, "Spatial resolution enhancement/smoothing of stereo-PIV data using POD-based and kriging interpolation methods," Phys. Fluids **19**, 064101 (2007).
[15] M. Lang, "Experimentelle Untersuchungen zur Transition in einer laminaren Ablöseblase mit Hilfe der LDA und der PIV," Ph.D. Thesis, Universität Stuttgart, 2005.
[16] O. Marxen, "Numerical studies of physical effects related to the controlled transition process in laminar separation bubbles," Ph.D. Thesis, Universität Stuttgart, 2005.
[17] O. Marxen and U. Rist, "Direct numerical simulation of non-linear transitional stages in an experimentally investigated laminar separation bubble," in *High Performance Computing in Science and Engineering '05*, edited by W. Nagel, W. Jager, and M. Resch (Springer, Berlin, 2005), pp. 103–117.
[18] S. Lophaven, H. Nielsen, and J. Sondergaard, "DACE—A Matlab kriging toolbox," Version 2.0, DTU Report No. IMM-REP-2002-12, 2002.
[19] J. Jeong and F. Hussain, "On the identification of a vortex," J. Fluid Mech. **285**, 69 (1995).
[20] S. Burgmann, J. Dannemann, and W. Schröder, "Time resolved and volumetric PIV measurements of a transitional separation bubble of an SD7003 airfoil," Exp. Fluids **44**, 609 (2008).
[21] D. Meyer, U. Rist, and S. Wagner, "Direct numerical simulation of secondary Λ-vortices in a transitional boundary layer," in *Advances in Turbulence VII*, Proceedings of the Seventh European Turbulence Conference, Saint Jean Cap Ferrat, France, edited by U. Frisch (Kluwer, Dordrecht, 1998), pp. 97–100
[22] M. Armstrong and A. Boufassa, "Comparing the robustness of ordinary kriging and lognormal kriging: Outlier resistance," Math. Geol. **20**, 447 (1988).
[23] H. E. Cekli, "Enhancement and smoothing methods for experimental data: Application to PIV measurements of a laminar separation bubble," M.Sc. thesis, Istanbul Technical University, 2007.