15th International Symposium on Particle Image Velocimetry – ISPIV 2023
June 19–21, 2023, San Diego, California, USA

Automated parameter optimization for LPT measurements using multi-objective evolutionary algorithms

Philipp Godbersen[1]*, Andreas Schröder[1]

[1] German Aerospace Center (DLR), Institute of Aerodynamics and Flow Technology, Göttingen, Germany

* Philipp.Godbersen@dlr.de

Abstract

The selection of suitable evaluation parameters for Iterative Particle Reconstruction (IPR) is an important aspect when processing particle tracking experiments. Such parameters generally need to be selected individually for each experiment setup and require some experience to adjust. We propose an optimization procedure based on evolutionary algorithms tailored specifically to the IPR approach that can select suitable parameters automatically. Besides assisting with routine evaluations such an optimization can be especially useful when investigating novel components or improvements to the standard IPR approach. Existing experience might no longer apply in such situations due to the complex interplay of individual components in the iterative algorithm. In such situations an optimized set of parameters can serve as a more objective evaluation of any modifications as opposed to a manual tuning of parameters which might miss unexpected synergies and effects. We implement such an optimization approach and utilize it to investigate the effect of an improved peak detection scheme within an IPR evaluation at high seeding densities. In addition to the synthetic experiments we also present a pathway towards application on real-world experiment data.

1 Introduction

An important component for Lagrangian particle tracking (LPT) measurements is the reconstruction of 3D particle clouds for individual time-steps from measurement images. Iterative particle reconstruction (IPR, Wieneke (2012), Jahn et al. (2021)) is a state of the art method that enables triangulation even at high particle image densities and is a component of LPT measurement techniques such as Shake-The-Box (STB, Schanz et al. (2016)).

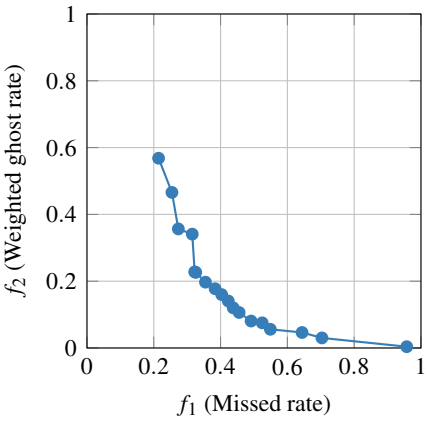
The reconstruction can be broken down into several stages. First 2D particle locations, the particle peaks, are identified for all cameras by means of a suitable peak detector. Together with a calibrated camera model these peaks are then used to triangulate 3D particle positions in the measurement volume. At this stage only a limited number of particles is triangulated for highly seeded images due to overlapping or indistinguishable particles. The IPR tackles this problem by back-projecting the successfully found particles onto the cameras, optimizing their position, and subtracting their image by means of an optical transfer function which completes one iteration of the scheme. The so formed residual image now has a reduced particle image density and allows the next iteration to find additional particles that were not recoverable on the initial image. Each iteration reduces the particle image density on the residual image this way until ideally it is empty at the end and all particles have been found. A detailed introduction is given by Jahn et al. (2021).

2 IPR parameter optimization

The application of IPR is controlled by a set of parameters such as triangulation radii, intensity filter values and peak detection thresholds that adjust the behavior of individual stages in the above mentioned scheme. Due to the iterative nature of the method not just a single set of parameters needs to be determined. Each iteration is provided with its individual set of parameter values due to the changing nature of the residual image. A parameter combination that is best for the densely seeded image of the initial iteration step is not

IPR parameter	search range
mlThresh	0.2 ... 0.9
minIntensity	10 ... 1000
omitBrightest	0 ... 3
maxShakeStep	0.1 ... 0.7
triangulationRadius	0.1 ... 1.0
maxNumCantSee	0 ... 2
minParticleDistance	0.0 ... 1.0

(a) IPR parameters to be optimized and their permitted value ranges. For a detailed description on the role of these parameters see Jahn et al. (2021)



(b) Pareto front for the initial IPR iteration. Each point corresponds to a full IPR parameter set as shown in table 1(a)

Figure 1: Optimization result for a single IPR iteration discovered by use of the multi objective optimizer

necessarily optimal for the later steps where the residual particle image density has significantly diminished and more aggressive values might be possible.

For the current study we have selected 7 parameters in an individual IPR iteration as relevant for an optimization. As each iteration gets its own individual parameter values the number of optimization variables directly scales with the number of iterations. An evaluation using 10 IPR iterations would then lead to 70 individual parameters needing to be optimized. In addition to the somewhat large amount of parameters we also desire an optimization procedure that provides flexibility in the design of the cost function. The selection of parameters for IPR iteration is driven by trade-offs in terms of the number of ghost particles and the number of found particles where increasing the percentage of found particles generally also increases the number of ghosts. This relationship is nonlinear and not necessarily known in advance so the selection of an optimal point to optimize towards is challenging. This motivates the use of a multi-objective optimization algorithm which by its design is able to discover these trade-off relationships between multiple objectives. Instead of a single optimal solution a pareto front within the objective space is returned. Such a front is the bounding curve of all tested parameter combinations. All other parameter combinations produce results that are dominated by the values of those on the curve.

Within this work we selected the percentage of missed particles and the percentage of ghost particles as the two objectives to be minimized. Keeping the objective two-dimensional has the advantage of easier visualization of the resulting pareto front as well as a reduced effort for the optimizer as opposed to higher dimensions. The two chosen variables are the key high-level metrics for judging the performance of an IPR evaluation and many secondary metrics correlate with them, e.g. an evaluation that produces a low amount of ghost particles while finding almost all true particles tends to be one that has a low positioning error for the found particles as well. Figure 1(b) shows an example for such a pareto front resulting from optimizing the parameters of a single IPR iteration for a synthetic experiment.

In practice we do not just perform a single IPR iteration but multiple which makes the optimization more complex. A joint optimization of all iterations together would greatly increase the number of parameters to optimize as well as making each cost function evaluation more expensive. Not just a single IPR iteration would need to be considered but always the full number. To greatly reduce the effort required we propose a different solution that avoids the fully joint optimization described above. The individual iterations can not be optimized completely separate as they depend on one another, however we can still exploit the sequential property of the iterations by noting that each only depends on the performance of the one directly preceding it. Optimal parameters for a given iteration strongly depend on the residual problem given to them by the previous iteration. A certain parameter combination might work better if the prior iteration was rather conservative with a low number of erroneous particles found. Another might benefit from a higher number of particles already identified even if at the cost of a higher number of ghost particles. The iteration must therefore not just choose its own optimal parameters but also the ones for the one directly preceding it, a recursion that ends with the first iteration having no prior.

Fortunately the prior multi-objective optimizations of the prior iterations already contain a lot of infor-

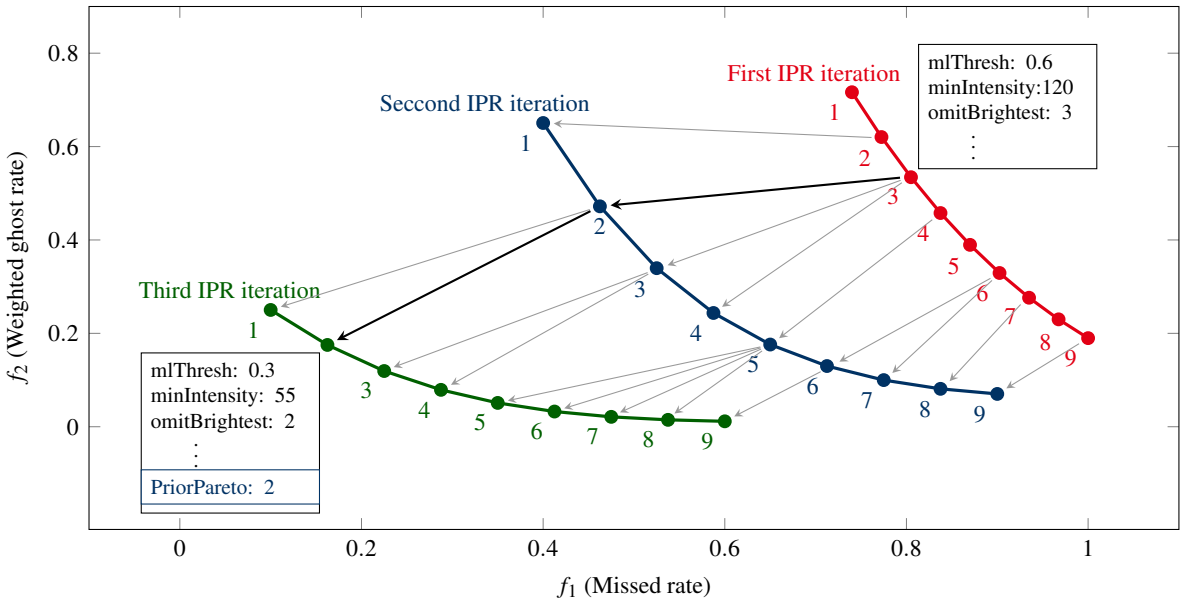


Figure 2: Illustration of the proposed sequential optimizing scheme. Three successive IPR iterations are shown with their respective pareto fronts in the objective space. Each optimizes its own set of parameters separately. The optimal choice of parameters for a given iteration depends on the choices made in the one before it. Instead of needing to perform a joint optimization of all parameter sets we recognize that any optimal parameter combination of a prior iteration must lie on its pareto curve. Each optimization therefore has an additional parameter that selects an index of point on the prior curve. We can therefore construct a full set of parameters by walking back from a selected point on the last curve along the path towards the first curve. An example of such path is shown in black.

mation that can be exploited. An iteration that needs to select parameters for a preceding iterations does not need to select them freely but already knows that any worthwhile parameters must lie on that iterations pareto curve. Instead of the full amount of values necessary to describe an IPR iteration it only needs to select one value, a parameter that walks along the pareto curve. For each IPR iteration we therefore only need to optimize its fixed set of IPR parameters plus one single extra parameter selecting a prior iteration instance from the pareto curve. This approach is visualized in figure 2 This means the number of parameters for each optimizer run stays constant irrespective of the total number of IPR iterations that are performed. Furthermore since the iterations build on one another we can avoid having to rerun earlier iterations over and over again by storing the results for those on the pareto curve. The pareto parameter then simply decides which prior result to load and within the cost function evaluation only a single IPR iteration building on that data needs to be considered. Whereas the full joint optimization had a rather quadratic effort in terms of number of IPR iterations this approach here now reduces this to a more linear effort making it feasible to optimize over several iterations.

The optimization problem for a specific iteration can be framed within this general structure:

$$\begin{aligned} \min \quad & f_m(x) && m = 1,..,M \\ \text{s.t.} \quad & g_j(x) \leq 0 && j = 1,..,J \\ & x_i^L \leq x_i \leq x_i^U && i = 1,..,N \\ & x \in \Omega \end{aligned} \tag{1}$$

with M cost functions f_m, J constraint functions g_j, lower and upper bounds for the N parameters x_i^L and x_i^U, and parameter space Ω. The constraint functions are not a necessity for the approach described above but they can be useful to communicate additional knowledge to the optimizer. In our case we added constraints to remove configurations with a long evaluation time as well as those with a very high number of ghosts after triangulation prior to filtering. These represent defective parameterizations that would not come close to the pareto front anyway but explicitly rejecting them via this constraint helps speed up the early part of the optimization by discouraging the optimizer from further exploring that region of the parameter space.

While the selection of the rate of missed particles and the rate of ghost particles of the current IPR iteration is a sensible choice for the cost functions we modify this slightly to take the sequential nature of optimizations into account. Instead of just optimizing for the ghost rate at the current IPR iteration we also take the ghost rate of the selected prior iteration steps into account. The idea is to provide a more stable goal for the optimizer as this forces it to maintain a low ghost level for the entire IPR processing, not just the last iteration. This enforces a more well behaved parameter set over the entire processing to penalize results where a high ghost level is maintained for the entire evaluation and then only somehow pressed towards lower values in the final IPR iteration. Reaching a low ghost rate in the final iteration is still the primary concern so a decaying weighted average over past iterations is used. One could also simply include this concern by adding an additional cost function dimension to the optimizer so that both the average and the last value are optimized for and a selection can be made at the end. We decided against this since increasing the dimensionality of the objective space makes the optimization problem harder and the resulting pareto surfaces are not as easily visualized as the pareto fronts in 2d. The two objective functions as defined in the general framework in (1) are therefore:

$$f_1(x) = \text{missrate}_n, \tag{2}$$

$$f_2(x) = \frac{\sum_{i=1}^{n} w_i \, \text{ghostrate}_i}{\sum_{i=1}^{n} w_i} \quad \text{with} \quad w_i = i^{-\frac{1}{4}}, \tag{3}$$

with n as the current optimized IPR iteration index where i then iterates over prior index values.

We treat this as a black-box optimization problem as no gradient information is readily available and utilize a multi-objective evolutionary algorithm to provide solutions. For the results within this paper we utilize the USNAGI-3 algorithm (Seada and Deb, 2016) with 18 reference directions as implemented in the pymoo software package (Blank and Deb, 2020) but other algorithms are also viable and we intend to further investigate the performance of different algorithms for this optimization in the future. Evolutionary algorithms approach the optimization problem by creating a population of candidates that are initially drawn from a random sampling of parameter space. Candidates are judged by a fitness function with some being removed in the process. These are replaced by new candidates which are determined from other members in the population by operators mimicking genetic evolution. For our optimization we utilize the default operators defined in pymoo which consist of simulated binary crossover and polynomial mutation (Deb et al., 2007). We are currently working towards usage of the Borg algorithm (Hadka and Reed, 2013) for the optimization as we expect a slightly more suitable placement of reference points on the pareto fronts with this. To limit optimization time the optimizer is currently limited to 100 optimizer iteration as well as an early termination threshold of 0.01 in the objective space. This terminates optimization if no significant change in the objective is observed over several past optimizer iterations. With our settings this corresponds to approximately 1000 individual evaluations of the objective function and thereby individual IPR evaluations. A single IPR iteration required approximately 2.5 seconds resulting in roughly 40 minutes of optimization time per pareto front. Limited work has gone into optimizing for time so this can likely be much improved on either by tuning of the optimizer or more restrictive selection of the parameter search ranges. To gather impressions of the capabilities of the optimizer the parameter ranges were set rather wide and did not change between the different iterations. This could be further optimized by selecting tighter bounds and introducing further knowledge into their selection. For example it is unlikely, that a large triangulation radius together with aggressive peak detection and only few cameras required is a sensible choice for the very first iteration due to the still high particle image density.

3 Application to synthetic data set

In order to demonstrate the proposed optimization approach we apply it to the evaluation of a novel peak detection scheme and its impact on IPR performance on a synthetic case. The novel peak detector is yet unpublished and not the target of this paper. It is what partially motivated the optimization approach presented here as the change in peak detection performance possibly requires different IPR parameterizations than used for a conventional evaluation and provides an interesting test for the optimizer.

We use the Case II taken from Jahn et al. (2021) at 0.16 ppp which is the highest density considered there and was not fully solved. A maximum number of 16 IPR iterations was selected to tackle this challenging problem. We then optimize IPR parameters for this case using the optimization scheme described above. Figure 3 shows the pareto fronts discovered by the optimizer for each of the individual IPR iterations. Curves

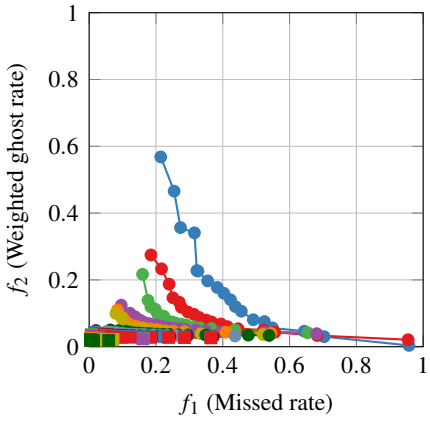
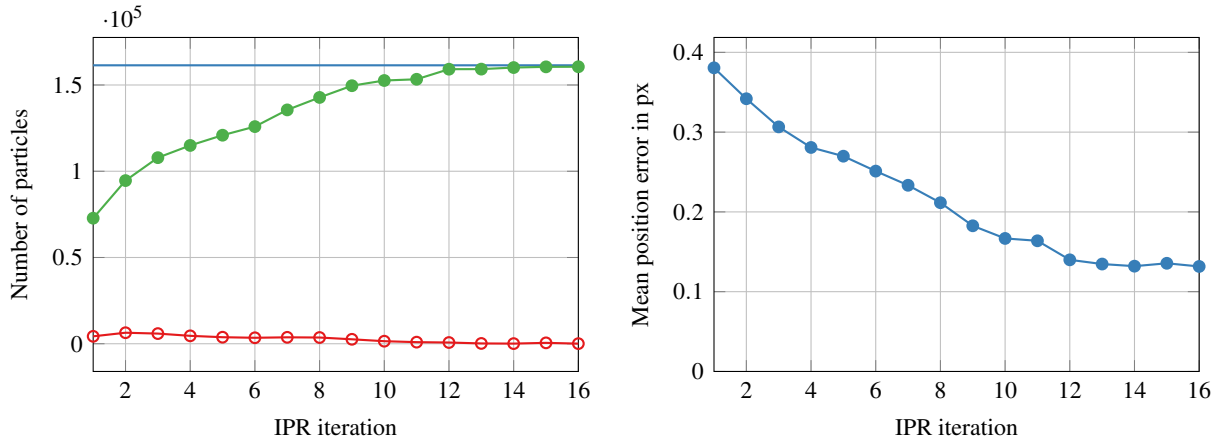


Figure 3: Pareto curves resulting from optimization of multiple IPR iterations. Different iterations are shown by different colors and markers (colorset cycles). Markers for later iterations overlap towards the origin but the trend of successive iterations recovering more and more particles is apparent.



(a) Amount of ghost particles (hollow circles) and true particles (solid circles) found. The total number of particles in the ground truth is shown by the solid line.

(b) Mean positioning error of found particles

Figure 4: Performance of the optimized parameters when used in IPR evaluation of an 0.16 ppp case

for later iterations start to visually overlap as they approach the minimum but the steady trend towards an optimal solution is apparent. An optimized parameter set can now be obtained by selecting the point closest to the origin on the final curve and walking back trough the curves as described in figure 2. Figure 4(a) shows the results of a IPR evaluation on this case using the optimized parameter set together with the novel peak detection. With this combination the problem can be fully solved in the sense that essentially all particles are found and no ghosts remain. Notable is also the very low amount of ghost particles even in early iterations. The average position accuracy of the correctly recovered particles is shown in figure 4(b). As more an more particles are found this value improves until it converges towards less than 0.15 px average error. This compares favorable with the results shown in Jahn et al. (2021) where for this high density only much less particles are correctly recovered. There, with the old peak detector and a hand-tuned parameter set, roughly 65% of the true particles are found after 20 IPR iterations, with a ghost particle rate of approx. 25%. Those evaluation were not necessarily optimized towards such high densities as they put more focus onto selecting good general purpose values but this still shows the large improvement the new peak detection and individual optimization of parameters can offer.

The plot in figure 4(a) suggests that less than the 16 iteration initially selected might be necessary. The iterative nature of the optimization with pareto fronts available for each successive stage simplifies such a study. Instead of selecting some optimal value on the final pareto curve, which then defines parameter for a

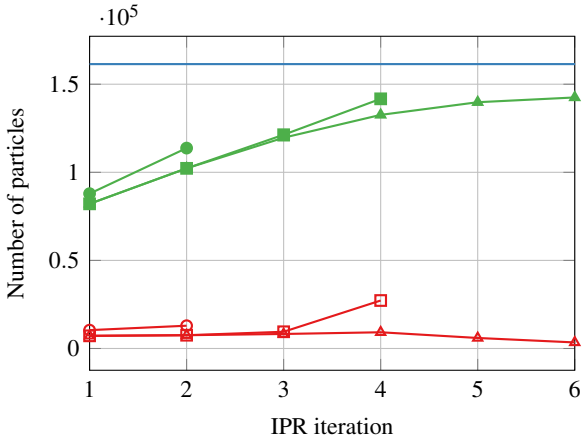


Figure 5: Instead of selecting an optimal point from final optimized iteration once can also select an optimal point from any of the previous pareto fronts resulting in an optimized scheme for a lower number of IPR iterations. An optimal paramtetrization for the first iteration is not necessarily the same when doing 10 total iteration versus just two. This plot shows such shorter evaluations for two, four, and six total iterations. One can see that evaluations using more iterations can utilize less aggressive settings that more slowly discover all the particles while producing less ghosts.

full 16 iterations IPR evaluation one can simply select an earlier pareto front and choose a value there. This then gives values for all prior steps too but now with the expectation that this selected iteration is to be the final one. An optimal paramtetrization for the first iteration is not necessarily the same when doing 10 total iteration versus just two. If the optimizer has more IPR iterations available to reach its final goal it can select less aggressive settings in the early iteration with the knowledge that together they still reach an optimal result at the end and with less creation of ghosts along the way. This can be seen in figure 5 where a selection of such reduced evaluations are plotted in one figure. One can see that the optimizer selects more aggressive setting for the first iteration if only two are to be performed in total than when using more.

We can investigate the parameters the optimizer has found for the iterations. Figure 6 shows a selection of the parameters used by the IPR for each iteration. Some patterns can be recognized in these graphs. The triangulation radius initially follows a ramping behavior where starting from a conservative value in the first iteration this radius is then increased. For the fourth iteration the radius is then suddenly decreased down to its initial value again. As can be seen in the lower graph this coincides with a jump in the maxNumCantSee parameter from zero to one. This parameter selects the amount of cameras that a particle must be visible on for a triangulation to be considered valid. A value of zero restricts to only triangulating particles that are seen by all four cameras. A value of one correspondingly allows for particles not visible on one of the cameras This less restrictive value for the triangulation allows more particles to be found but would also introduce more ghosts, an effect that is countered by the reduction in allowed triangulation radius by the optimizer. This behavior matches the approach one would take if selecting these parameters by hand. Interesting is that the value for the maxNumCantSee parameter is only briefly increased and then immediately reduced back to zero until it is increased again briefly at a later iteration again. This differs from the otherwise used approach where typically one would then keep this value at the increased level and might be due to the increased performance of the peak detection so that this brief increase is enough. The triangulation radius is also reduced in the later iterations where particles are allowed to be missing from up to two cameras. A majority of the particles is found in the early iterations though and especially the parameters for the very last iterations might not be particularly insightful as essentially all particles have been found at that stage.

A more consistent trend can be seen in the value for the parameter minIntensity which governs the threshold particle intensity in counts below which a particle is considered a ghost and filtered away. This value quite consistently decreases from the first to the last iteration resulting in less and less aggressive filtering. Such a behavior can be explained by the reducing complexity of the triangulation for later iterations. As more and more particles are found the effective particle image density on the residual images becomes less and less resulting in an easier triangulation problem. Such easier triangulation provide less of a chance for ghost particles to appear so less aggressive filtering can be applied in order to recover those true particles which where erroneously filtered away before. To gain further insight into the selection of the parameters by the optimizer, a sensitivity study could be performed in further investigations in order to identify which

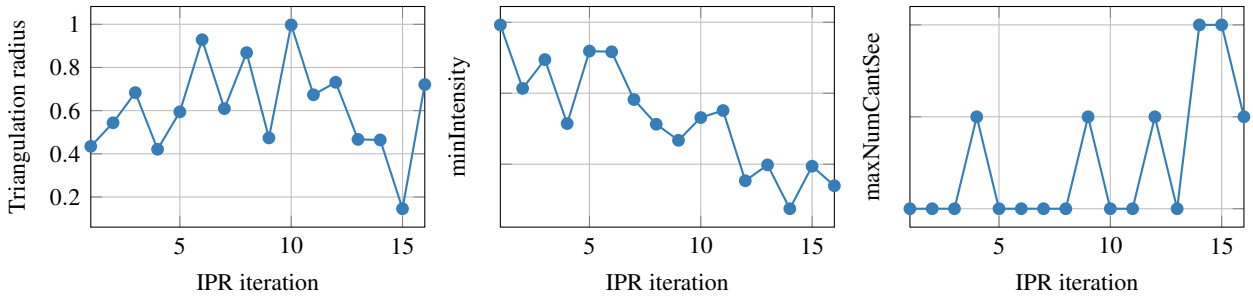


Figure 6: A selection of IPR parameters found by the optimizer for this synthetic case. Values for last few iterations are most likely ambiguous as the problem is essentially perfectly solved at this point and no meaningful improvement is possible regardless of parameter values.

choices of the optimizer are the most significant. Based on a completed optimization this could be archived by perturbing the parameters around the found minimum.

4 Pathways towards real world data

As a final point we wish to address the application to real world data. Even just on synthetic data the approach provides interesting capabilities for comparing modifications of IPR algorithms, but the method is not necessarily restricted in this way. The optimization scheme requires ground truth data which generally is not readily available for actual experiments. We propose the use of time-resolved STB results as pseudo ground truth data source. A well converged STB evaluation barely relies on IPR performance at all when processing a given additional image and as such can provide a semi independent source of information for pseudo ground truth generation. The additional context provided by the temporal information over many images is simply much more robust than the individual image IPR targeted here.

An attractive target for IPR optimization are experiments using multi-pulse STB (Novara et al., 2016) as these evaluations rely much more on individual IPR performance since much less temporal information is available. However in these cases it is generally not feasible to conduct time-resolved STB evaluations due to the speed of the flow. To circumvent this we propose to conduct a time-resolved evaluation on flow off images of the multi-pulse experiment where the slow naturally convecting flow is slow enough to be temporally resolved even by the slow repetition rate of the multi-pulse system. The optimization can then be performed on these flow-off images as now pseudo ground truth data is available, the results of which can then be applied to the actual flow-on measurement as both flow-on and flow-off cases feature an identical optical setup. Figure 7 shows such an evaluation on Volume Self Calibration (VSC, Wieneke (2008)) images of an actual multi-pulse measurement (Manovski et al., 2021). VSC images were used as these were the available flow-off images from this existing measurement. Future experiments could involve recording a short additional image sequence at flow off conditions similar to VSC images but at the actual target particle image density as this is typically reduced for VSC runs. All necessary components of the optimization pipeline shown in figure 8 are therefore available. Future work will consist of fully applying such a pipeline to an real world measurement.

5 Conclusion

We have described a novel optimization scheme for IPR parameter selection and demonstrated its effectiveness on a synthetic test case. By adapting the optimization approach specifically to the sequential nature of the problem, the application to an IPR with a large number of iterations was possible. Using optimized parameters together with a not further detailed improved peak detection scheme we were able to achieve significantly better reconstruction performance on the synthetic case from Jahn et al. (2021). The improved performance is only partly through the optimization scheme presented here, the primary cause is the improved peak detection itself. However due to the interdependent nature of the individual components of an IPR evaluation any changes also require fresh parameter combinations in order to fully exploit the benefits of the improvements. This contribution in this work enable a given IPR algorithm to work to its full potential for a given dataset.

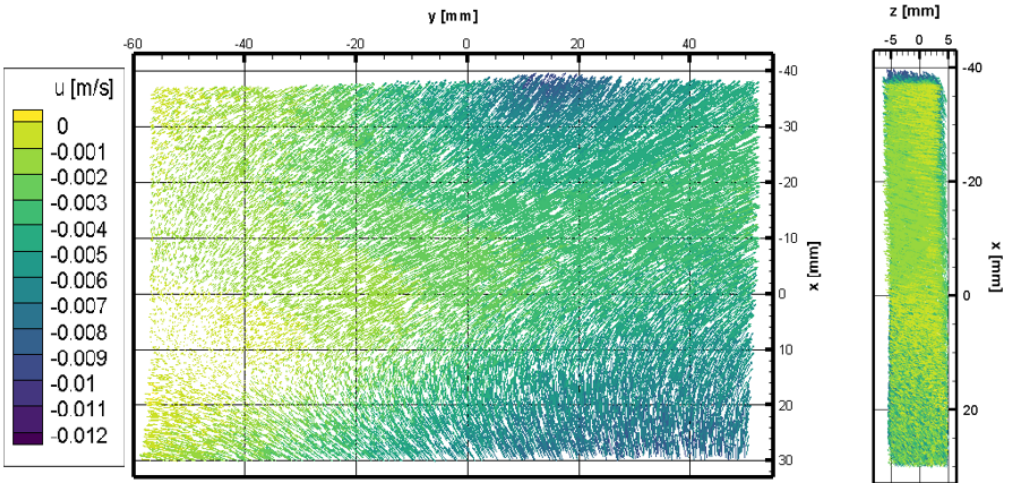


Figure 7: Particle tracks recovered from VSC images of a multi-pulse STB measurement by use of time-resolved STB in the flow off state. Source for pseudo ground truth data.

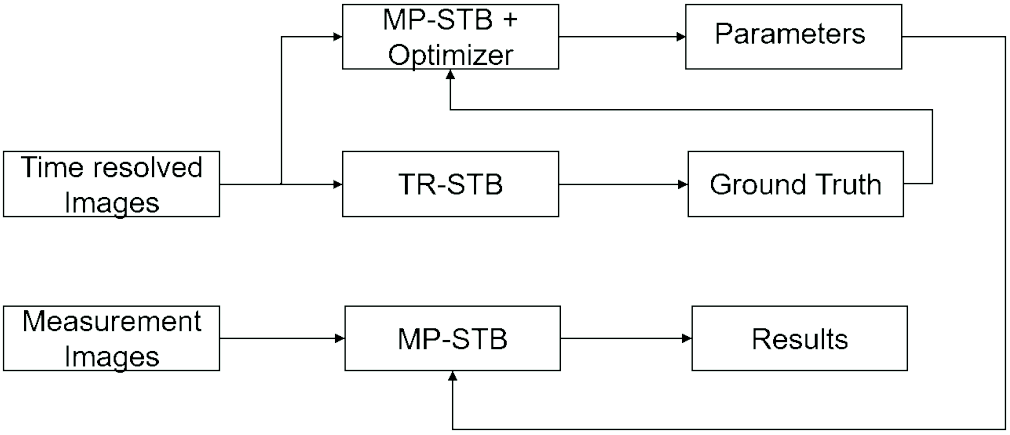


Figure 8: Schematic for the use of time-resolved measurement data on flow-off images for multi-pulse applications.

Beyond the use on synthetic experiments, the method also provides an approach for individually optimized IPR parameters for real world measurements.

References

Blank J and Deb K (2020) Pymoo: Multi-objective optimization in python. *IEEE Access* 8:89497–89509

Deb K, Sindhya K, and Okabe T (2007) Self-adaptive simulated binary crossover for real-parameter optimization. in *Proceedings of the 9th Annual Conference on Genetic and Evolutionary Computation*. GECCO '07. page 1187–1194. Association for Computing Machinery, New York, NY, USA

Hadka D and Reed P (2013) Borg: An Auto-Adaptive Many-Objective Evolutionary Computing Framework. *Evolutionary Computation* 21:231–259

Jahn T, Schanz D, and Schröder A (2021) Advanced iterative particle reconstruction for lagrangian particle tracking. *Experiments in Fluids* 62:1–24

Manovski P, Novara M, Mohan NKD, Geisler R, Schanz D, Agocs J, Godbersen P, and Schröder A (2021) 3d lagrangian particle tracking of a subsonic jet using multi-pulse shake-the-box. *Experimental Thermal and Fluid Science* 123:110346

Novara M, Schanz D, Reuther N, Kähler CJ, and Schröder A (2016) Lagrangian 3d particle tracking in high-speed flows: Shake-The-Box for multi-pulse systems. *Experiments in Fluids* 57:128

Schanz D, Gesemann S, and Schröder A (2016) Shake-The-Box: Lagrangian particle tracking at high particle image densities. *Experiments in Fluids* 57:70

Seada H and Deb K (2016) A unified evolutionary optimization procedure for single, multiple, and many objectives. *IEEE Transactions on Evolutionary Computation* 20:358–369

Wieneke B (2008) Volume self-calibration for 3d particle image velocimetry. *Experiments in fluids* 45:549–556

Wieneke B (2012) Iterative reconstruction of volumetric particle distribution. *Measurement Science and Technology* 24:024008