A. Additional numerical results

A.1 LGSSM

Tables 1 and 2 display the matrices $A$, $B$, $RR^\top$, and $SS^\top$ used for all experiments in the LGSSM model context. In Figures 1a, 2a and 3a we display boxplots of bias estimates, where each estimate is obtained by averaging $10^4$ independent runs of the corresponding algorithm and each box is based on $10^3$ replications of this bias estimator. The PARIS is compared to the PPG for different algorithmic configurations $(N, k, k_0)$ and for different computational budgets $C = kN$ of sizes $10^3$ (Figure 1), $2 \times 10^3$ (Figure 2), and $5 \times 10^3$ (Figure 3). Each experiment is carried through for each of the different designs $k_0 = \lfloor 2^{-1}k \rfloor$, $k_0 = \lfloor (3/4)C/N \rfloor$, and $k_0 = k - 1$ of the burn-in.

Table 1: The $A$ (left) and $B$ (right) matrices in the LGSSM.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.4119</td>
<td>0.0013</td>
<td>0.001010</td>
<td>0.0019</td>
<td>0.0018</td>
</tr>
<tr>
<td>2</td>
<td>0.245</td>
<td>0.0012</td>
<td>0.00169</td>
<td>0.0019</td>
<td>0.0018</td>
</tr>
<tr>
<td>3</td>
<td>0.349</td>
<td>0.0012</td>
<td>0.00205</td>
<td>0.0021</td>
<td>0.0020</td>
</tr>
<tr>
<td>4</td>
<td>0.2572</td>
<td>-0.0012</td>
<td>-0.00269</td>
<td>-0.0020</td>
<td>-0.0019</td>
</tr>
<tr>
<td>5</td>
<td>0.795</td>
<td>-0.0012</td>
<td>-0.0033</td>
<td>-0.0031</td>
<td>-0.0030</td>
</tr>
</tbody>
</table>

Table 2: The covariance matrices $RR^\top$ (left) and $SS^\top$ (right) for the state and measurement noises, respectively, in the LGSSM.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.0026</td>
<td>0.0006</td>
<td>0.0006</td>
<td>0.0006</td>
<td>0.0006</td>
</tr>
<tr>
<td>2</td>
<td>0.0006</td>
<td>0.0006</td>
<td>0.0006</td>
<td>0.0006</td>
<td>0.0006</td>
</tr>
<tr>
<td>3</td>
<td>0.0006</td>
<td>0.0006</td>
<td>0.0006</td>
<td>0.0006</td>
<td>0.0006</td>
</tr>
<tr>
<td>4</td>
<td>0.0006</td>
<td>0.0006</td>
<td>0.0006</td>
<td>0.0006</td>
<td>0.0006</td>
</tr>
<tr>
<td>5</td>
<td>0.0006</td>
<td>0.0006</td>
<td>0.0006</td>
<td>0.0006</td>
<td>0.0006</td>
</tr>
</tbody>
</table>

A.2 Stochastic volatility

In this section we repeat the same experiments in Appendix A.1 in the context of the StoVol model described in Section 5. Figure 4, Figure 5, and Figure 6 display boxplots of bias estimates for the PARIS and the PPG for different algorithmic configurations $(N, k, k_0)$ and different computational budgets $C = kN$ of sizes $10^2$ (Figure 4), $5 \times 10^2$ (Figure 5), and $10^3$ (Figure 6). The bias of each algorithm is estimated by averaging $10^3$ independent runs of the same, and each box is based on $10^3$ independent replications of this bias estimator. Again, in each plot, the PARIS and PPG share the same computational budget (regardless configuration of the PPG).
A.2 Stochastic volatility

Figure 1: PARIS and PPG outputs for the LGSSM with $C = 10^3$ and different designs of the burn-in $k_0$.

![Graphs showing PARIS and PPG outputs for different $k_0$ values.](image)

(a) $k_0 = \lfloor 2^{-1}C/N \rfloor$
(b) $k_0 = \lfloor (3/4)C/N \rfloor$
(c) $k_0 = k - 1$

Figure 2: PARIS and PPG outputs for the LGSSM with $C = 2.5 \times 10^3$ and different designs of the burn-in $k_0$.

![Graphs showing PARIS and PPG outputs for different $k_0$ values.](image)

(a) $k_0 = \lfloor 2^{-1}C/N \rfloor$
(b) $k_0 = \lfloor (3/4)C/N \rfloor$
(c) $k_0 = k - 1$

**Choice of $(N, k, k_0)$**. Designing the configuration $(N, k, k_0)$ is challenging, since the upper bound $\kappa_{N,n}$ on the mixing rate is known to be conservative. As clear from Figure 4–Figure 6, the best configuration also depends on $C$; indeed, we see that for a smaller budget it is better to let the particle sample size $N$ be large. Nevertheless, for more generous budgets it seems to be better to use a large number $k$ of iterations at the expense of $N$.

Concerning the burn-in parameter $k_0$, the choice depends mainly on the bias–variance trade-off. In applications where minimising the bias is important one would choose $k_0 = k - 1$, which gives the smallest possible bias. Otherwise, a trade-off that provides an improvement in bias at the cost of an increase in MSE over the PARIS by only a factor of 2 is to choose
A.2 Stochastic volatility

Figure 3: PARIS and PPG outputs for the LGSSM with $C = 5 \times 10^3$ and different designs of the burn-in $k_0$.

$k_0 = \lfloor k/2 \rfloor$; recall the discussion in Section 4.2

Figure 4: PARIS and PPG outputs for the stovol model with $C = 10^2$ and different designs on the burn-in $k_0$.

A.2.1 Comparison with the Rhee–Glynn-type estimator of [21]

We now compare the proposed PPG estimator with the unbiased Rhee–Glynn-type smoothing estimator $H_{k_0,k,N}$ defined in [21] Eq. 2], where the parameter $k_0$ is the burn-in phase length, $k$ the minimum number of Gibbs iterations, and $N$ the number of particles used in the coupled conditional particle filter. This estimator is based on the coupled conditional particle filter with ancestor sampling proposed in [21]; see Algorithm 4 for details.
A.2 Stochastic volatility

Figure 5: PARIS and PPG outputs for the stovol model with $C = 5 \times 10^2$ and different designs of the burn-in $k_0$.

Figure 6: PARIS and PPG outputs for the stovol model with $C = 10^3$ and different designs of the burn-in $k_0$.

Since the number of particles used in the algorithm is itself a random variable, we first perform $3 \times 10^3$ independent runs of the same and report the average meeting time (i.e., number of iterations of Algorithm 4 until the conditional paths $\zeta_{0:n}$ and $\zeta'_{0:n}$ become identical) for three different choices of the hyperparameters in Table 3. We deduce from Table 3 that the average total number of particles generated is about $3 \times 10^3$. Therefore, we compare the Rhee–Glynn estimator induced by the coupled conditional particle filter with the PPG estimator with $(N, k_0, k) = (10, 150, 300)$. Figure 7 shows histograms of estimates produced using the Rhee–Glynn-type procedure, for the three different configurations, along with histograms of the estimates produced by the PPG. Each histogram is based on $3 \times 10^3$ in-
A.2 Stochastic volatility

<table>
<thead>
<tr>
<th>$N$</th>
<th>$k_0$</th>
<th>$k$</th>
<th>Meeting time</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>5</td>
<td>10</td>
<td>30.4</td>
</tr>
<tr>
<td>250</td>
<td>2</td>
<td>4</td>
<td>12.6</td>
</tr>
<tr>
<td>500</td>
<td>1</td>
<td>2</td>
<td>7.1</td>
</tr>
</tbody>
</table>

Table 3: Coupled conditional particle filter meeting times for three different configurations with $Nk = 10^3$.

dependent replications. We find that the variance and empirical bias of the Rhee–Glynn-type estimator is about 10 and 20 times larger, respectively, than for the PPG for the same computational effort.

![Histograms](image)

(a) $(N, k) = (100, 10)$  (b) $(N, k) = (250, 4)$  (c) $(N, k) = (500, 2)$

Figure 7: Histograms of estimates produced using the Rhee–Glynn-type smoothing estimator of [21] for three different configurations and the PPG estimator with $(N, k_0, k) = (10, 150, 300)$. Each box is based on 3000 independent replications. The plot also provides the corresponding 95% coverage asymptotic confidence intervals.

Another way of obtaining Rhee–Glynn-type smoothing estimator would be to consider the coupling of the conditional backward sampling particle filter, as proposed in [22]. In the case of the bootstrap particle filter, the conditional particle filter with backward sampling is probabilistically equivalent to the conditional particle filter with ancestor sampling. Furthermore, [22, Section 7] also show that for $n = 10^3$, both the conditional particle filter with backward sampling and the conditional particle filter with ancestor sampling have similar performance. Thus, we expect the results in this section to translate to the estimators proposed in [22].
B. Algorithms

The following section provides pseudocode for the algorithms discussed in Section 3, namely: the original PARIS algorithm (Algorithm 1) proposed in [24], the conditional PARIS update (Algorithm 2), and the PPG (Algorithm 3). In addition, we provide a pseudocode for the coupled conditional particle filter with ancestor sampling (Algorithm 4), being the key ingredient of the unbiased Rhee–Glynn-type estimator proposed in [21] against which the PPG is benchmarked in Appendix A.2.1. Note that the conditional PARIS update described in Algorithm 2 differs somewhat from that described in Section 3 in the way the underlying conditional dual process \( \{\xi_m\}_{m \in \mathbb{N}} \) is propagated; more precisely, in Algorithm 2, each conditional dual process update \( \xi_{m+1} \sim M_m(\xi_{m+1}|\xi_m, \cdot) \), where the value of \( \zeta_{m+1} \) is inserted into a randomly chosen position in \( \xi_{m+1} \) (whereas the remaining elements of \( \xi_{m+1} \) are sampled independently from \( \Phi_m(\mu(\xi_m)) \)) is replaced by deterministic assignment of \( \zeta_{m+1} \) to \( \xi_{N_{m+1}} \). Of course, this change has no impact as long as we are interested in integrating functions that are permutation invariant with respect to the produced many-body systems, which is the case throughout our work. Still, as this derandomization technique simplifies somewhat the implementation of the PPG, we have chosen to include it in our pseudocode.

\[
\text{Data: } \{(\xi^n_i, \beta^n_i)\}_{i=1}^N \quad \text{Result: } \{(\xi'^n_{i+1}, \beta'^n_{i+1})\}_{i=1}^N \\
\text{for } i \leftarrow 1 \text{ to } N \text{ do } \\
\quad \text{draw } I_{n+1}^{i} \sim \text{cat}(\{g_n(\xi^n_\ell)\}_{\ell=1}^N); \\
\quad \text{draw } \xi'^n_{i+1} \sim M_n(\xi'^n_{n+1}, \cdot); \\
\quad \text{for } j \leftarrow 1 \text{ to } M \text{ do } \\
\quad \quad \text{draw } J_{n+1}^{(i,j)} \sim \text{cat}(\{q_n(\xi^n_\ell, \xi'^n_{i+1})\}_{\ell=1}^N) \\
\quad \text{end} \\
\quad \text{set } \beta'^n_{i+1} \leftarrow \frac{1}{M} \sum_{j=1}^M \left( \beta_{n+1}^{(i,j)} + \bar{h}_n(\xi'^n_{n+1}; \xi^n_{i+1}) \right); \\
\text{end}
\]

Algorithm 1: One update of the PARIS.

Coupling algorithms. Algorithm 4 provides a more detailed description of (the predictive variant of) the coupled conditional particle filter
Data: $v_n$, $\zeta_{n+1}$

Result: $v_{n+1}$

for $i \leftarrow 1$ to $N-1$ do
  draw $I_{m+1}^i \sim \text{cat}(\{q_m(\xi_m^\ell)|_{\ell=1}^N\}$;
  draw $\xi_{m+1|m+1}^i \sim M_m(\xi_{m+1}^\cdot, \cdot)$;
end

set $\xi_{m+1|m+1}^N \leftarrow \zeta_{m+1}$;

for $i \leftarrow 1$ to $N$ do
  for $j \leftarrow 1$ to $M$ do
    draw $J_{m+1}^{(i,j)} \sim \text{cat}(\{q_m(\xi_m^\ell, \xi_{m+1|m+1}^i)|_{\ell=1}^N\}$;
  end

set $\beta_{m+1}^i \leftarrow \frac{1}{M}\sum_{j=1}^M \left(\beta_{m+1}^{(i,j)} + \tilde{h}_m(\xi_{m+1}^i, \xi_{m+1|m+1}^i)\right)$;

set $\xi_{0:m+1|m+1}^i \leftarrow (\xi_{0:m+1|m+1}^i, \xi_{m+1|m+1}^i)$;
end

set $v_{n+1} \leftarrow ((\xi_{0:n+1|n+1}^1, \beta_{n+1}^1), \ldots, (\xi_{0:n+1|n+1}^N, \beta_{n+1}^N))$;

**Algorithm 2:** One conditional PARIS update, expressed in a short form as “$v_{n+1} \leftarrow \text{CondPaRIS}(v_n, \zeta_{n+1})$”.

proposed in [21, Algorithm 1], and we focus here on the version of this algorithm where the iteratively produced particle paths underlying the resulting estimator are generated by means of ancestor sampling [23]. If $\{\omega_i\}_{i=1}^N$ and $\{\omega_i'\}_{i=1}^N$ are possibly unnormalised event probabilities, we denote by $M(\{\omega_i\}_{i=1}^N, \{\omega_i'\}_{i=1}^N)$ the maximal coupling between the distributions $\text{cat}(\{\omega_i\}_{i=1}^N)$ and $\text{cat}(\{\omega_i'\}_{i=1}^N)$. In our implementations, we used the maximum coupling given in [?]. In order to couple two conditional particle filters, we assume, following [21, Algorithm 1], that for every $m \in \mathbb{N}$ we are able to simulate a random variable $\varepsilon_m$, defined on some measurable space $(S_m, \mathcal{S}_m)$ and distributed according $\mu_m \in M_1(S_m)$, such that there exists some measurable function $\phi$ on $(X_m \times S_m, \mathcal{X}_m \otimes \mathcal{S}_m)$ such that for every $x_m \in X_m$, $\mu_m \circ \phi^{-1}(x_m, \cdot)$ (the pushforward of $\mu_m$ through $\phi_m(x_m, \cdot)$) equals $M_m(x_m, \cdot)$.
Data: $\zeta_{0:n}$
Result: $v_n$, $\zeta'_{0:n}$
draw $(\xi_{0|0}^1, \ldots, \xi_{0|N-1}^1) \sim \eta_0 \otimes (N-1)$;
set $\xi_{0|0}^N \leftarrow \zeta_0$;
set $\beta_0 \leftarrow (0, \ldots, 0)$;
for $m \leftarrow 0$ to $n-1$ do
  run $((\xi_{m+1|m+1}^1, \beta_{m+1}^1), \ldots, (\xi_{N|m+1}^N, \beta_{N|m+1}^N)) \leftarrow$
  $\text{CondPaRIS}((\xi_{m|m}^1, \beta_{m}^1), \ldots, (\xi_{N|m}^N, \beta_{m}^N), \zeta_{m+1})$;
end
set $v_n \leftarrow ((\xi_{n|n}^1, \beta_{n}^1), \ldots, (\xi_{n|n}^N, \beta_{n}^N))$;
draw $J \sim \text{cat}\{1\}^N$;
set $\zeta'_{0:n|n} \leftarrow \xi_{0:n|n}^J$;

Algorithm 3: One iteration of the Parisian particle Gibbs (PPG)
Data: $\zeta_{0:n}, \tilde{\zeta}_{0:n}$
Result: $\zeta'_{0:n}, \tilde{\zeta}'_{0:n}$
set $(\xi_0^1, \ldots, \xi_0^{N-1}) \sim \eta_0^{(N-1)}$;
set $(\tilde{\xi}_0^1, \ldots, \tilde{\xi}_0^{N-1}) \leftarrow (\xi_0^1, \ldots, \xi_0^{N-1})$;
set $(\xi_0^N, \tilde{\xi}_0^N) \leftarrow (\zeta_0, \tilde{\zeta}_0)$;
for $m \leftarrow 0$ to $n - 1$ do
  for $i \leftarrow 1$ to $N - 1$ do
    draw $(I_{m+1}^i, \tilde{I}_{m+1}^i) \sim M(\{g_m(\xi_\ell^i)\}_{\ell=1}^N, \{g_m(\tilde{\xi}_\ell^i)\}_{\ell=1}^N)$;
  end
  draw $(I_{m+1}^N, \tilde{I}_{m+1}^N) \sim M(\{q_m(\xi_\ell^N, \zeta_{m+1})\}_{\ell=1}^N, \{q_m(\tilde{\xi}_\ell^N, \tilde{\zeta}_{m+1})\}_{\ell=1}^N)$;
  for $i \leftarrow 1$ to $N$ do
    draw $\varepsilon_m \sim \mu_m$;
    set $(\xi_{m+1}^i, \tilde{\xi}_{m+1}^i) \leftarrow (\phi_m(I_{m+1}^i, \varepsilon_m), \phi_m(\tilde{I}_{m+1}^i, \varepsilon_m))$;
  end
end
draw $J_n \sim \text{cat}(\{1\}_{\ell=1}^N)$;
set $\tilde{J}_n \leftarrow J_n$;
set $(\xi_n, \tilde{\xi}_n) \leftarrow (\xi_{J_n}, \tilde{\xi}_{\tilde{J}_n})$;
for $m \leftarrow n - 1$ to $0$ do
  set $(J_m, \tilde{J}_m) \leftarrow (I_{m+1}^{J_m}, \tilde{I}_{m+1}^{\tilde{J}_m})$;
  set $(\xi_m, \tilde{\xi}_m) \leftarrow (\xi_{J_m}^m, \tilde{\xi}_{\tilde{J}_m}^m)$;
end

Algorithm 4: Coupled conditional particle filters [21].
C. Additional proofs

C.1 Proof of Proposition 2

First, note that, by definitions (3.1) and (3.2),

\[
H_n(x_{0:n}) := \int \mathcal{S}_n(x_{0:n}, dy_{n}) \mu(x_{0:n}|n) h
\]

\[
= \int \cdots \int \left( \frac{1}{N} \sum_{j_n=1}^{N} h(x_{0:n-1|n}^j, x_n^j) \right)
\]

\[
\times \prod_{m=0}^{n-1} \prod_{i_{m+1}=1}^{N} \sum_{j_{m+1}=1}^{N} \sum_{j_m=1}^{N} q_m(x_m^j, x_{m+1}^{j_m+1}) \delta_{x_0:0|m}^j \left( dx_{0:m|m+1} \right),
\]

where \( x_{0:0|0} = \emptyset \) for all \( i \in [1, N] \) by convention. We will show that for every \( k \in [0, n] \), \( H_{k,n} \equiv H_n \), where

\[
H_{k,n}(x_{0:n}) := \frac{1}{N} \sum_{j_n=1}^{N} \cdots \sum_{j_{k+1}=1}^{N} \prod_{j_k=1}^{N} q_k(x_k^j, x_{k+1}^{j_k+1}) a_{k,n}(x_0, \ldots, x_{k-1}, x_k^j, \ldots, x_n^j)
\]

with

\[
a_{k,n}(x_0, \ldots, x_{k-1}, x_k^j, \ldots, x_n^j)
\]

\[
= \int \prod_{m=0}^{k-2} \prod_{i_{m+1}=1}^{N} \sum_{j_{m+1}=1}^{N} \sum_{j_m=1}^{N} q_m(x_m^{j_m}, x_{m+1}^{j_m+1}) \delta_{x_{0:m|m}}^j \left( dx_{0:m|m+1} \right) h(x_{0:k-1|k}, x_k^j, \ldots, x_n^j).
\]

Since, by convention, \( \prod_{\ell=0}^{n-1} = 1 \), \( H_{n,n}(x_{0:n}) = N^{-1} \sum_{j_n=1}^{N} a_{n,n}(x_0, \ldots, x_{n-1}, x_n^j) \), and we note that \( H_n \equiv H_{n,n} \). We now show that \( H_{k,n} \equiv H_{k-1,n} \) for every \( k \in [1, n] \); for this purpose, note that

\[
a_{k,n}(x_0, \ldots, x_{k-1}, x_k^j, \ldots, x_n^j)
\]

\[
= \int \prod_{m=0}^{k-2} \prod_{i_{m+1}=1}^{N} \sum_{j_{m+1}=1}^{N} \sum_{j_m=1}^{N} q_m(x_m^{j_m}, x_{m+1}^{j_m+1}) \delta_{x_{0:m|m}}^j \left( dx_{0:m|m+1} \right) h(x_{0:k-1|k}, x_k^j, \ldots, x_n^j)
\]

\[
\times \int \prod_{i_k=1}^{N} \sum_{j_{k-1}=1}^{N} q_{k-1}(x_{k-1}^{j_{k-1}}, x_k^{i_k}) \delta_{x_{0:k-1|k-1}}^{j_{k-1}} \left( dx_{0:k-1|k} \right) h(x_{0:k-1|k}, x_k^j, \ldots, x_n^j),
\]
and since \(x_{0:k-2|k-1}^{j_{k-1}} = (x_{0:k-2|k-1}^{j_{k-1}}, x_{0:k-1|k-1}^{j_{k-1}})\), it holds that

\[
\int \prod_{i_k=1}^{N} \prod_{j_{k-1}=1}^{N} q_{k-1}(x_{k-1}^{j_{k-1}}, x_{k}^{i_k}) \delta_{x_{0:k-1|k-1}^{j_{k-1}}} (dx_{0:k-1|k-1}^{i_k}) h(x_{0:k-1|k-1}^{j_{k-1}}, x_{k}^{i_k}, \ldots, x_{j_n}^{i_n})
\]

\[= \sum_{j_{k-1}=1}^{N} \frac{q_{k-1}(x_{k-1}^{j_{k-1}}, x_{k}^{j_k})}{\sum_{j_{k-1}=1}^{N} q_{k-1}(x_{k-1}^{j_{k-1}}, x_{k}^{j_k})} h(x_{0:k-2|k-1}^{j_{k-1}}, x_{k-1}^{j_k}, x_{k}^{j_k}, \ldots, x_{j_n}^{j_k}).\]

Therefore, we obtain

\[
a_{k,n}(x_0, \ldots, x_{k-1}, x_{k}^{j_k}, \ldots, x_{j_n}^{j_n})
\]

\[= \int \prod_{m=0}^{k-2} \prod_{i_m=1}^{N} \prod_{j_m=1}^{N} q_{m}(x_{m}^{j_m}, x_{m+1}^{i_m}) \delta_{x_{m}^{j_m}} (dx_{m}^{i_m}) h(x_{m}^{j_m}, x_{m+1}^{i_m})
\]

\[\times \sum_{j_{k-1}=1}^{N} \frac{q_{k-1}(x_{k-1}^{j_{k-1}}, x_{k}^{j_k})}{\sum_{j_{k-1}=1}^{N} q_{k-1}(x_{k-1}^{j_{k-1}}, x_{k}^{j_k})} h(x_{0:k-2|k-1}^{j_{k-1}}, x_{k-1}^{j_k}, x_{k}^{j_k}, \ldots, x_{j_n}^{j_k}).\]

Now, changing the order of summation with respect to \(j_{k-1}\) and integration on the right hand side of the previous display yields

\[
a_{k,n}(x_0, \ldots, x_{k-1}, x_{k}^{j_k}, \ldots, x_{j_n}^{j_n})
\]

\[= \sum_{j_{k-1}=1}^{N} \frac{q_{k-1}(x_{k-1}^{j_{k-1}}, x_{k}^{j_k})}{\sum_{j_{k-1}=1}^{N} q_{k-1}(x_{k-1}^{j_{k-1}}, x_{k}^{j_k})} a_{k-1,n}(x_0, \ldots, x_{k-1}^{j_{k-1}}, \ldots, x_{j_n}^{j_k}).\]

Thus,

\[
H_{k,n}(x_{0:n})
\]

\[= \frac{1}{N} \sum_{j_{k-1}=1}^{N} \frac{q_{k-1}(x_{k-1}^{j_{k-1}}, x_{k}^{j_k})}{\sum_{j_{k-1}=1}^{N} q_{k-1}(x_{k-1}^{j_{k-1}}, x_{k}^{j_k})} a_{k-1,n}(x_0, \ldots, x_{k-1}^{j_{k-1}}, \ldots, x_{j_n}^{j_k})
\]

\[= \frac{1}{N} \sum_{j_{k-1}=1}^{N} \frac{q_{k-1}(x_{k-1}^{j_{k-1}}, x_{k}^{j_k})}{\sum_{j_{k-1}=1}^{N} q_{k-1}(x_{k-1}^{j_{k-1}}, x_{k}^{j_k})} a_{k-1,n}(x_0, \ldots, x_{k-2}^{j_{k-2}}, \ldots, x_{n}^{j_n})
\]

\[= H_{k-1,n}(x_{0:n}).\]
which establishes the recursion. Therefore, $H_n \equiv H_{0,n}$ and we may now conclude the proof by noting that $B_nh \equiv H_{0,n}$.

### C.2 Proof of Theorem 5

In order to establish Theorem 5 we will prove the following more general result, of which Theorem 5 is a direct consequence.

**Proposition 1.** For every $n \in \mathbb{N}$ and $M \in \mathbb{N}^*$ there exist $c_n > 0$ and $d_n > 0$ such that for every $N \in \mathbb{N}^*$, $z_0:n \in X_0:n$, $(f_n, \tilde{f}_n) \in F(\mathcal{X}_n)^2$, and $\varepsilon > 0$,

\[
\int C_n S_n(z_0:n, db_n) \times 1 \left\{ \left| \frac{1}{N} \sum_{i=1}^{N} \{ b_n(x_i^n) + \tilde{f}_n(x_i^n) \} - \eta_n(z_0:n)(f_n B_n(z_{0:n-1})h_n + \tilde{f}_n) \right| \geq \varepsilon \right\} \leq c_n \exp \left( -\frac{d_n N \varepsilon^2}{2\kappa_n^2} \right),
\]

where

\[
\kappa_n := \|f_n\|_{\infty} \sum_{m=0}^{n-1} \|\tilde{h}_m\|_{\infty} + \|\tilde{f}_n\|_{\infty}. \tag{C.1}
\]

To prove Proposition 1 we need the following technical lemma.

**Lemma 1.** For every $n \in \mathbb{N}$, $(f_{n+1}, \tilde{f}_{n+1}) \in F(\mathcal{X}_{n+1})^2$, $z_{0:n+1} \in X_{0:n+1}$, and $N \in \mathbb{N}^*$,

\[
\gamma_{n+1}(z_{0:n+1})(f_{n+1} B_{n+1}(z_{0:n})h_{n+1} + \tilde{f}_{n+1})
\]

\[
= \left( 1 - \frac{1}{N} \right) \gamma_n(z_{0:n}) \{ Q_n f_{n+1} B_n(z_{0:n-1})h_n + Q_n (\tilde{h}_n f_{n+1} + \tilde{f}_{n+1}) \}
\]

\[
+ \frac{1}{N} \gamma_n(z_{0:n}) g_n \left( f_{n+1}(z_{n+1}) B_{n+1}(z_{0:n})h_{n+1}(z_{n+1}) + \tilde{f}_{n+1}(z_{n+1}) \right).
\]

**Proof.** Since Lemma 2 holds also for the Feynman–Kac model with a frozen path, we obtain

\[
\gamma_{n+1}(z_{0:n+1})(f_{n+1} B_{n+1}(z_{0:n})h_{n+1} + \tilde{f}_{n+1})
\]

\[
= \gamma_n(z_{0:n}) \{ Q_n(z_{n+1}) f_{n+1} B_n(z_{0:n})h_n + Q_n(z_{n+1})(\tilde{h}_n f_{n+1} + \tilde{f}_{n+1}) \}.
\]
Thus, the proof is concluded by noting that for every $x_n \in \mathcal{X}_n$ and $h \in \mathcal{F}(\mathcal{X}_{n+1})$,

$$Q_n(z_{n+1})h(x_n) = \left(1 - \frac{1}{N}\right) Q_n h(x_n) + \frac{1}{N} g(x_n) h(x_n, z_{n+1}).$$

Finally, before proceeding to the proof of Proposition 4 we introduce the law of the PARIS evolving conditionally on a frozen path $z = \{z_m\}_{m \in \mathbb{N}}$. Define, for $m \in \mathbb{N}$ and $z_{m+1} \in \mathcal{X}_{m+1}$,

$$P_m(z_{m+1}) : \mathcal{Y}_m \times \mathcal{Y}_{m+1} \ni (y_m, A) \mapsto \int M_m(z_{m+1})(x_m|y_m, d x_{m+1}) S_m(y_m, x_{m+1}, A).$$

For any given initial distribution $\psi_0 \in \mathcal{M}_1(\mathcal{Y}_0)$, let $\mathbb{P}_{\psi_0}^{P_z}$ be the distribution of the canonical Markov chain induced by the Markov kernels $\{P_m(z_{m+1})\}_{m \in \mathbb{N}}$ and the initial distribution $\psi_0$. By abuse of notation we write $\mathbb{P}_{\psi_0}^{P_z}$ instead of $\mathbb{P}_{\psi_0[\eta_0(z_0)]}^{P_z}$, where the extension $\psi_0[\eta_0]$ is defined in Section 6.3.

**Proof of Proposition 4** We proceed by forward induction over $n$. Let the $\sigma$-fields $\mathcal{F}_n$ and $\mathcal{F}_n$ be defined as in the proof of Theorem 3 but for the conditional PARIS dual process. Then, under the law $\mathbb{P}_{\psi_0}^{P_z}$, reusing (6.11),

$$\mathbb{E}_{\psi_0}^{P_z} \left[ \beta_n f_n(\xi_n^1) + \tilde{f}_n(\xi_n^1) \mid \mathcal{F}_{n-1} \right]$$

$$= \mathbb{E}_{\psi_0}^{P_z} \left[ \mathbb{E}_{\psi_0}^{P_z} \left[ \beta_n f_n(\xi_n^1) + \tilde{f}_n(\xi_n^1) \mid \mathcal{F}_{n-1} \right] \right]$$

$$= \mathbb{E}_{\psi_0}^{P_z} \left[ \begin{array}{c}
\mathbb{E}_{\psi_0}^{P_z} \\
\beta_n f_n(\xi_n^1) + \tilde{f}_n(\xi_n^1) \mid \mathcal{F}_{n-1} \end{array} \right]$$

$$= \mathbb{E}_{\psi_0}^{P_z} \left[ f_n(\xi_n^1) \sum_{\ell=1}^N \sum_{\ell' = 1}^N \sum_{n'=1}^N q_{n-1}(\xi_{n'-1}^\ell, \xi_{n'-1}^1) \left( \beta_n^{\ell} + \tilde{h}_n(\xi_{n-1}^\ell, \xi_n^1) \right) + \tilde{f}_n(\xi_n^1) \mid \mathcal{F}_{n-1} \right].$$

Using (2.6), we get

$$\mathbb{E}_{\psi_0}^{P_z} \left[ \begin{array}{c}
\beta_n f_n(\xi_n^1) + \tilde{f}_n(\xi_n^1) \mid \mathcal{F}_{n-1} \\
\end{array} \right]$$

$$= \left(1 - \frac{1}{N}\right) \sum_{\ell=1}^N \sum_{\ell' = 1}^N \left( \beta_n^{\ell} Q_{n-1} f_n(\xi_{n-1}^\ell) + Q_{n-1}(\tilde{h}_n(\xi_{n-1}^\ell) + \tilde{f}_n(\xi_{n-1}^\ell)) \right)$$

$$+ \frac{1}{N} \left( f_n(z_n) \sum_{\ell=1}^N \sum_{\ell' = 1}^N \sum_{n'=1}^N q_{n-1}(\xi_{n'-1}^\ell, z_n) \left( \beta_n^{\ell} + \tilde{h}_n(\xi_{n-1}^{\ell'}, z_n) \right) + \tilde{f}_n(z_n) \right).$$

(C.2)
In order to apply the induction hypothesis to each term on the right-hand side of the previous identity, note that

\[ B_n(z_{0:n-1})h_n(z_n) = \frac{\eta_{n-1}(z_{0:n-1})[q_{n-1}(\cdot, z_n)\{B_{n-1}(z_{0:n-2})h_{n-1}(\cdot) + \tilde{h}_{n-1}(\cdot, z_n)\}]}{\eta_{n-1}(z_{0:n-1})[q_{n-1}(\cdot, z_n)]}. \]

Therefore, using Lemma \[\text{(C.1)}\] and noting that \( \gamma_n(z_{0:n})\mathbb{1}_{X_n}/\gamma_{n-1}(z_{0:n})\mathbb{1}_{X_{n-1}} = \eta_{n-1}(z_{0:n-1})g_{n-1} \) yields

\[ \eta_{n}(z_{0:n})\left(f_nB_n(z_{0:n-1})h_n + \tilde{f}_n\right) = \frac{1}{N} \left(f_n(z_n)B_n(z_{0:n-1})h_n(z_n) + \tilde{f}_n(z_n)\right) + \left(1 - \frac{1}{N}\right) \frac{\eta_{n-1}(z_{0:n-1})\{Q_{n-1}f_nB_{n-1}(z_{0:n-2})h_n + Q_{n-1}(\tilde{h}_{n-1} + \tilde{f}_n)\}}{\eta_{n-1}(z_{0:n-1})g_{n-1}}. \] (C.3)

By combining (C.2) with (C.3), we decompose the error according to

\[ \frac{1}{N} \sum_{i=1}^{N} \{\beta_n^i f_n(\xi_{n|n}^i) + \tilde{f}_n(\xi_{n|n}^i)\} - \eta_{n}(z_{0:n})\left(f_nB_n(z_{0:n-1})h_n + \tilde{f}_n\right) = \frac{1}{N} \sum_{i=1}^{N} \{\beta_n^i f_n(\xi_{n|n}^i) + \tilde{f}_n(\xi_{n|n}^i)\} - \mathbb{E}_{\eta_0}^{P^z} \left[ \beta_n^1 f_n(\xi_{n|n}^1) + \tilde{f}_n(\xi_{n|n}^1) \mid \tilde{F}_{n-1} \right] + \mathbb{E}_{\eta_0}^{P^z} \left[ \beta_n^1 f_n(\xi_{n|n}^1) + \tilde{f}_n(\xi_{n|n}^1) \mid \tilde{F}_{n-1} \right] - \eta_{n}(z_{0:n})\left(f_nB_n(z_{0:n-1})h_n + \tilde{f}_n\right) = I_{N}^{1(1)} + \left(1 - \frac{1}{N}\right) I_{N}^{2(2)} + \frac{1}{N} I_{N}^{3(3)}. \] (C.4)

where

\[ I_{N}^{1(1)} := \frac{1}{N} \sum_{i=1}^{N} \{\beta_n^i f_n(\xi_{n|n}^i) + \tilde{f}_n(\xi_{n|n}^i)\} - \mathbb{E}_{\eta_0}^{P^z} \left[ \beta_n^1 f_n(\xi_{n|n}^1) + \tilde{f}_n(\xi_{n|n}^1) \mid \tilde{F}_{n-1} \right], \]

\[ I_{N}^{2(2)} := \frac{\sum_{\ell=1}^{N} \{\beta_{n-1}^{\ell} Q_{n-1} f_n(\xi_{n-1}^\ell) + Q_{n-1}(\tilde{h}_{n-1} f_n + \tilde{f}_n)(\xi_{n-1}^\ell)\}}{\sum_{e=1}^{N} g_{n-1}(\xi_{n-1}^e)} - \eta_{n-1}(z_{0:n-1})\{Q_{n-1}f_nB_{n}(z_{0:n-1})h_n + Q_{n-1}(\tilde{h}_{n-1} f_n + \tilde{f}_n)\} \frac{1}{\eta_{n-1}(z_{0:n-1})g_{n-1}} \] (C.5)
C.2 Proof of Theorem 5

and

\[
I_N^{(3)} := f_n(z_n) \sum_{\ell=1}^{N} \frac{q_{n-1}(\xi_{n-1}^{\ell}, z_n)}{\sum_{\ell'=1}^{N} q_{n-1}(\xi_{n-1}^{\ell'}, z_n)} \left( \beta_{n-1}^\ell + \tilde{h}_{n-1}(\xi_{n-1}^{\ell}, z_n) \right) - f_n(z_n) \eta_{n-1}(z_{0:n-1}) \{ g_{n-1}(\cdot, z_n) \{ B_{n-1}(z_{0:n-2}) h_{n-1}(\cdot) + \tilde{h}_{n-1}(\cdot, z_n) \} \}. \tag{C.6}
\]

The proof is now completed by treating the terms \(I_N^{(1)}\), \(I_N^{(2)}\), and \(I_N^{(3)}\) separately, using Hoeffding’s inequality and its generalisation in [15, Lemma 4]. Choose \(\varepsilon > 0\); then, by Hoeffding’s inequality,

\[
P_{\eta_0}^{P,z} \left( |I_N^{(1)}| \geq \varepsilon \right) \leq 2 \exp \left( -\frac{1}{2} \frac{\varepsilon^2}{\kappa_n^2} N \right). \tag{C.7}
\]

To treat \(I_N^{(2)}\), we apply the induction hypothesis to the numerator and denominator, each normalised by \(1/N\), yielding, since \(\|Q_{n-1} h\|_\infty \leq \bar{\tau}_{n-1} \|h\|_\infty\) for all \(h \in F(\mathcal{X}_{n-1} \otimes \mathcal{X}_n)\),

\[
P_{\eta_0}^{P,z} \left( \left| \frac{1}{N} \sum_{\ell=1}^{N} \{ \beta_{n-1}^\ell Q_{n-1} f_n(\xi_{n-1}^{\ell}) + Q_{n-1}(\tilde{h}_{n-1} f_n + \tilde{f}_n)(\xi_{n-1}^{\ell}) \} \right. \right.
\]

\[-\left. \eta_{n-1}(z_{0:n-1}) \{ Q_{n-1} f_n B_n(z_{0:n-1}) h_n + Q_{n-1}(\tilde{h}_{n-1} f_n + \tilde{f}_n) \} \right| \geq \varepsilon \right) \leq c_{n-1} \exp \left( -d_{n-1} \frac{\varepsilon^2}{\tau_{n-1}^2 \kappa_n^2} N \right)
\]

and

\[
P_{\eta_0}^{P,z} \left( \left| \frac{1}{N} \sum_{\ell=1}^{N} g_{n-1}(\xi_{n-1}^{\ell}) - \eta_{n-1}(z_{0:n-1}) g_{n-1} \right| \geq \varepsilon \right) \leq c_{n-1} \exp \left( -d_{n-1} \frac{\varepsilon^2}{\tau_{n-1}^2 \kappa_n^2} N \right).
\]

Combining the previous two bounds with the generalised Hoeffding inequality in [15, Lemma 4] yields, using also the bounds

\[
\frac{\sum_{\ell=1}^{N} \{ \beta_{n-1}^\ell Q_{n-1} f_n(\xi_{n-1}^{\ell}) + Q_{n-1}(\tilde{h}_{n-1} f_n + \tilde{f}_n)(\xi_{n-1}^{\ell}) \}}{\sum_{\ell'=1}^{N} g_{n-1}(\xi_{n-1}^{\ell'})} \leq \kappa_n
\]
and $\eta_{n-1}(z_{0:n-1})g_{n-1} \geq \tau_{n-1}$, the inequality
\[
\mathbb{P}_{\eta_0}^z \left( |I_N^{(2)}| \geq \varepsilon \right) \leq c_{n-1} \exp \left( -d_{n-1} \frac{\varepsilon^2}{\tau_{n-1}^2 \kappa_n^2} N \right). \tag{C.8}
\]
The last term $I_N^{(3)}$ is treated along similar lines; indeed, by the induction hypothesis, since $\|q_{n-1}\|_{\infty} \leq \bar{\tau}_{n-1}\bar{\sigma}_{n-1}$,
\[
\mathbb{P}_{\eta_0}^z \left( \left| \frac{1}{N} \sum_{\ell=1}^N q_{n-1}(\xi_{n-1}^\ell, z_n) \left( \beta_{n-1}^\ell + \tilde{h}_{n-1}(\xi_{n-1}^\ell, z_n) \right) - \eta_{n-1}(z_{0:n-1})[q_{n-1}('), z_n]) \{B_{n-1}(z_{0:n-1})h_{n-1}(') + \tilde{h}_{n-1}('), z_n)\} \right| \geq \varepsilon \right) 
\leq c_{n-1} \exp \left( -d_{n-1} \left( \frac{\varepsilon}{\bar{\tau}_{n-1}\bar{\sigma}_{n-1} \sum_{m=0}^{n-1} \|h_m\|_{\infty}} \right)^2 N \right)
\] and
\[
\mathbb{P}_{\eta_0}^z \left( \left| \frac{1}{N} \sum_{\ell=1}^N q_{n-1}(\xi_{n-1}^\ell, z_n) - \eta_{n-1}(z_{0:n-1})[q_{n-1}('), z_n]) \right| \geq \varepsilon \right) 
\leq c_{n-1} \exp \left( -d_{n-1} \left( \frac{\varepsilon}{\bar{\tau}_{n-1}\bar{\sigma}_{n-1}} \right)^2 N \right).
\]
Thus, since
\[
\sum_{\ell=1}^N q_{n-1}(\xi_{n-1}^\ell, z_n) \left( \beta_{n-1}^\ell + \tilde{h}_{n-1}(\xi_{n-1}^\ell, z_n) \right) \leq \sum_{m=0}^{n-1} \|h_m\|_{\infty}
\]
and $\eta_{n-1}(z_{0:n-1})[q_{n-1}('), z_n]) \geq \tau_{n-1}$, the generalised Hoeffding inequality provides
\[
\mathbb{P}_{\eta_0}^z \left( |I_N^{(3)}| \geq \varepsilon \right) \leq c_{n-1} \exp \left( -d_{n-1} \left( \frac{\tau_{n-1} \varepsilon}{2\bar{\tau}_{n-1}\bar{\sigma}_{n-1}\|f_n\|_{\infty} \sum_{m=0}^{n-1} \|h_m\|_{\infty}} \right)^2 N \right). \tag{C.9}
\]
Finally, combining the bounds \((C.7)\) \((C.9)\) completes the proof. \qed
C.3 Proof of Proposition 3

The statement of Proposition 3 is implied by the following more general result, which we will prove below.

**Proposition 2.** For every \(n \in \mathbb{N}, M \in \mathbb{N}^*, N \in \mathbb{N}^*, z_{0:n} \in X_{0:n}, (f_n, \tilde{f}_n) \in F(X_n)^2\), and \(p \geq 2\), it holds that

\[
N \int C_n S_n(z_{0:n}, d\mathbf{b}_n) \left\| \frac{1}{N} \sum_{i=1}^{N} \{b_n^i f_n(x_{n|n}^i) + \tilde{f}_n(x_{n|n}^i)\} - \eta_n(z_{0:n}) (f_n B_n(z_{0:n-1}) h_n + \tilde{f}_n) \right\|^p \leq c_n (p/d_n)^{p/2} N^{-p/2} \kappa_n^p,
\]

where \(c_n > 0\), \(d_n > 0\) and \(\kappa_n\) are defined in Proposition 1 and (C.1), respectively.

Before proving Proposition 2, we establish the following result.

**Lemma 2.** Let \(X\) be an \(\mathbb{R}^d\)-valued random variable, defined on some probability space \((\Omega, \mathcal{F}, \mathbb{P})\), satisfying \(\mathbb{P}(|X| \geq t) \leq c \exp(-t^2/(2\sigma^2))\) for every \(t \geq 0\) and some \(c > 0\) and \(\sigma > 0\). Then for every \(p \geq 2\) it holds that \(\mathbb{E}[|X|^p] \leq cp^{p/2}\sigma^p\).

**Proof.** Using Fubini’s theorem and the change of variable formula,

\[
\mathbb{E}[|X|^p] = \int_0^\infty pt^{p-1} \mathbb{P}(|X| \geq t) \, dt = cp^{p/2-1}\sigma^p \Gamma(p/2),
\]

where \(\Gamma\) is the Gamma function. It remains to apply the bound \(\Gamma(p/2) \leq (p/2)^{p/2-1}\) (see [?]), which holds for \(p \geq 2\) by [2, Theorem 1.5].

**Proof of Proposition 2.** By combining Proposition 1 and Lemma 2 we obtain

\[
N \int C_n S_n(z_{0:m}, d\mathbf{b}_n) \left\| \frac{1}{N} \sum_{i=1}^{N} \{b_n^i f_n(x_{n|n}^i) + \tilde{f}_n(x_{n|n}^i)\} - \eta_n(z_{0:n}) (f_n B_n(z_{0:n-1}) h_n + \tilde{f}_n) \right\|^2 \leq c_n (p/d_n)^{p/2} N^{-p/2} \left( \|f_n\|_\infty \sum_{m=0}^{n-1} \|\tilde{h}_m\|_\infty + \|\tilde{f}_n\|_\infty \right)^p,
\]

which was to be established.
C.4 Proof of Proposition 4

Like previously, we establish Proposition 4 via a more general result, namely the following.

**Proposition 3.** For every $n \in \mathbb{N}$, there exists $\bar{c}_{n}^{bias} < \infty$ such that for every $M \in \mathbb{N}^*$, $N \in \mathbb{N}^*$, $z_{0:n} \in X_{0:n}$, and $(f_n, \tilde{f}_n) \in F(X_n)^2$,

$$\left| \int \mathbb{C}_n \mathbb{S}_n(z_{0:n}, \mathbb{d} b_n) \frac{1}{N} \sum_{i=1}^{N} \{ b_n^i f_n(x_{n|n}^i) + \tilde{f}_n(x_{n|n}^i) \} - \eta_n(z_{0:n}) (f_n B_n(z_{0:n-1}) h_n + \tilde{f}_n) \right| \leq \bar{c}_{n}^{bias} \kappa_n N^{-1},$$

where $\kappa_n$ is defined in (C.1).

We preface the proof of Proposition 3 by a technical lemma providing a bound on the bias of ratios of random variables.

**Lemma 3.** Let $\alpha$ and $\beta$ be (possibly dependent) random variables defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and such that $\mathbb{E}[\alpha^2] < \infty$ and $\mathbb{E}[\beta^2] < \infty$. Moreover, assume that there exist $c > 0$ and $d > 0$ such that $|\alpha/\beta| \leq c$, $\mathbb{P}$-a.s., $|a/b| \leq c$, $\mathbb{E}[(\alpha - a)^2] \leq c^2 d^2$, and $\mathbb{E}[(\beta - b)^2] \leq d^2$. Then

$$|\mathbb{E}[\alpha/\beta] - a/b| \leq 2c(d/b)^2 + c|\mathbb{E}[\beta - b]|/|b| + |\mathbb{E}[\alpha - a]|/|b|. \quad (C.10)$$

**Proof.** Using the identity

$$\mathbb{E}[\alpha/\beta] - a/b = \mathbb{E}[(\alpha/\beta)(\beta - b)^2]/b^2 + \mathbb{E}[(\alpha - a)(\beta - b)]/b^2 + a \mathbb{E}[b - \beta]/b^2 + \mathbb{E}[\alpha - a]/b,$$

the claim is established by applying the Cauchy–Schwarz inequality and the assumptions of the lemma according to

$$|\mathbb{E}[\alpha/\beta] - a/b| \leq c \mathbb{E}[(\beta - b)^2]/b^2 + \{ \mathbb{E}[(\alpha - a)^2] \mathbb{E}[(\beta - b)^2] \}^{1/2}/b^2 + |a| \mathbb{E}[\beta - b]/b^2 + \mathbb{E}[\alpha - a]/b^2 \leq 2c(d/b)^2 + c|\mathbb{E}[\beta - b]|/|b| + |\mathbb{E}[\alpha - a]|/|b|.$$

**Proof of Proposition 4.** We proceed by induction and assume that the claim holds true for $n - 1$. Reusing the error decomposition (C.4), it is enough to bound the expectations of the terms $\mathbb{I}_{N}^{(2)}$ and $\mathbb{I}_{N}^{(3)}$ given in (C.5) and (C.6),
respectively (since $\mathbb{E}_n^{P_n}[I_N^{(1)}] = 0$). This will be done using the induction hypothesis, Lemma 3 and Proposition 2. More precisely, to bound the expectation of $I_N^{(2)}$, we use Lemma 3 with $\alpha \leftarrow \alpha_n$, $\beta \leftarrow \beta_n$, $a \leftarrow a_n$, and $b \leftarrow b_n$, where

$$\alpha_n := \frac{1}{N} \sum_{\ell=1}^{N} \{ \beta_{n-1}^{\ell} Q_{n-1} f_n (\xi_{n-1}^{\ell}) + Q_{n-1} (\hat{h}_{n-1} f_n + \hat{f}_n) (\xi_{n-1}^{\ell}) \}, \quad \beta_n := \frac{1}{N} \sum_{\ell=1}^{N} g_{n-1}^{\ell} (\xi_{n-1}^{\ell}),$$

$$a_n := \eta_{n-1} \langle z_{0:n-1} \rangle \{ Q_{n-1} f_n B_n \langle z_{0:n-1} \rangle h_n + Q_{n-1} (\hat{h}_{n-1} f_n + \hat{f}_n) \}, \quad b_n := \eta_{n-1} \langle z_{0:n-1} \rangle g_n.$$

For this purpose, note that $|\alpha_n/\beta_n| \leq \kappa_n$ and $|a_n/b_n| \leq \kappa_n$, where $\kappa_n$ is defined in (C.1). On the other hand, using Proposition 2 (applied with $p = 2$), we obtain

$$\mathbb{E}_n^{P_n}[|\alpha_n - a_n|^2] \leq d_n^2 \kappa_n^2 \quad \text{and} \quad \mathbb{E}_n^{P_n}[|\beta_n - b_n|^2] \leq d_n^2,$$

where $d_n^2 := c_n \tilde{r}_{n-1}/(d_n N)$. Using the induction assumption, we get

$$|\mathbb{E}_n^{P_n}[\alpha_n] - a_n| \leq \tilde{c}_n \kappa_{n-1} N^{-1} \tilde{r}_{n-1} \kappa_n \quad \text{and} \quad |\mathbb{E}_n^{P_n}[\beta_n] - b_n| \leq \tilde{c}_n \kappa_{n-1} N^{-1} \tilde{r}_{n-1}.$$

Hence, the conditions of Lemma 3 are satisfied and we deduce that

$$|\mathbb{E}_n^{P_n}[I_N^{(2)}]| = |\mathbb{E}_n^{P_n}[\alpha_n/\beta_n] - a_n/b_n| \leq 2 \kappa_n \frac{c_n \tilde{r}_{n-1}}{d_n N} \frac{\tilde{r}_{n-1}}{\tilde{r}_{n-1} N} + 2 \tilde{c}_n \kappa_{n-1} N^{-1} \tilde{r}_{n-1} \kappa_n.$$

The bound on $|\mathbb{E}_n^{P_n}[I_N^{(2)}]|$ is obtained along the same lines. \hfill \square

### C.5 Proof of Theorem 6

We first consider the bias, which can be bounded according to

$$|\mathbb{E}_k[I_{(k_0,k)}(f)] - \eta_{0:n} h_n| \leq (k - k_0)^{-1} \sum_{\ell=k_0+1}^{k} |\mathbb{E}_k \mu(\beta_n^{[\ell]})(\text{id}) - \eta_{0:n} h_n|$$

$$\leq (k - k_0)^{-1} N^{-1} \tilde{c}_n \kappa_{n-1} \left( \sum_{m=0}^{n-1} \| \tilde{h}_m \|_{\infty} \right) \sum_{\ell=k_0+1}^{k} \kappa_{N,n}^{\ell},$$

from which the bound (4.7) follows immediately.
We turn to the MSE. Using the decomposition
\[
\mathbb{E}_\xi[(\Pi_{(k_0,k_0),N}(f) - \eta_{0:n} h_n)^2] \leq (k - k_0)^{-2} \left\{ \sum_{\ell=k_0+1}^k \mathbb{E}_\xi[(\mu(\beta_n[\ell])(\text{id}) - \eta_{0:n} h_n)^2] + 2 \sum_{\ell=k_0+1}^k \sum_{j=\ell+1}^k \mathbb{E}_\xi[(\mu(\beta_n[\ell])(\text{id}) - \eta_{0:n} h_n)(\mu(\beta_n[j])(\text{id}) - \eta_{0:n} h_n)] \right\},
\]
the MSE bound in Theorem 2 implies that
\[
\sum_{\ell=k_0+1}^k \mathbb{E}_\xi[(\mu(\beta_n[\ell])(\text{id}) - \eta_{0:n} h_n)^2] \leq c^{mse}_{n} \left( \sum_{m=0}^{n-1} \|\tilde{h}_m\|_\infty \right)^2 N^{-1}(k - k_0).
\]
Moreover, using the covariance bound in Theorem 2 we deduce that
\[
\sum_{\ell=k_0+1}^k \sum_{j=\ell+1}^k \mathbb{E}_\xi[(\mu(\beta_n[\ell])(\text{id}) - \eta_{0:n} h_n)(\mu(\beta_n[j])(\text{id}) - \eta_{0:n} h_n)] \leq c^{cov}_{n} \left( \sum_{m=0}^{n-1} \|\tilde{h}_m\|_\infty \right)^2 N^{-3/2} \left( \sum_{\ell=k_0+1}^k \sum_{j=\ell+1}^k \kappa_{N,n}^{(j-\ell)} \right).
\]
Thus, the proof is concluded by noting that \( \sum_{\ell=k_0+1}^k \sum_{j=\ell+1}^k \kappa_{N,n}^{(j-\ell)} \leq (k - k_0)/(1 - \kappa_{N,n}). \)